A Study on Feature Analysis of Archival Metadata Standards in the Records Lifecycle

November 2013

Baek Jae Eun

A Study on Feature Analysis of Archival Metadata Standards in the Records Lifecycle

Baek Jae Eun

Graduate School of Library, Information and Media Studies

University of Tsukuba

November 2013

Abstract (English)

Following the development of the Internet and WWW, various digital information resources are being created and used in many different environments. The networked information environment has brought not only the popularization of digital resource but also some major problems. One of the major problems is maintaining digital resources for the future. Thus, we are faced with the fundamental problem of how to manage and preserve digital resources so that they can be used over the time.

Metadata schemas are well recognized as one of the important technological components for archiving and preservation of digital resources. There are several metadata standards for digital archiving and preservation, e.g. AGRkMS, EAD, ISAD(G), MoReq, PREMIS and so on. Each metadata standard has its data model and metadata element defined as a property of an entity included in the data model. Metadata standards have their own features in accordance with their primary application domain. However, a single standard is not enough to cover the whole lifecycle for archiving and preservation of digital resource. This means that we need to appropriately select metadata standards and combine them to develop metadata schemas to cover the whole lifecycle of resources (or records), i.e., from creation to archiving and preservation of resources.

The records lifecycle consists of several stages. Each stage of the lifecycle has some tasks to be carried out on the resource, e.g., creation, management, appraisal and so on. Metadata is used in those tasks of the lifecycle. Metadata elements are primarily defined as attributes of a resource. A metadata element is assigned its value during a particular task and may be used in other tasks of the records lifecycle. Thus, the requirements for metadata depend on the lifecycle stages and the tasks in each stage. It is crucial to select and combine metadata standards in accordance with the requirements of the application domain in every stage of the records lifecycle in order to define metadata schemas for archiving and preservation of the resources. However, the relationships between the metadata elements and resource tasks are not explicitly given as a part of the definition of the schemas. So, we use the lifecycle as a basis to analyze the feature of the different metadata standards and clarify the relationships between the metadata elements and the

records lifecycle stages.

In this study, we used metadata standards developed for archiving and preservation, i.e., ISAD(G), PREMIS. Also, we used AGLS Metadata, AGRkMS for records management and a set of metadata elements extracted from the decision tree for digital preservation proposed by the Digital Preservation Coalition in the UK.

The feature analysis of metadata standards in this study was carried out in two steps. In the first step of the study, we have clarified the features of major metadata element sets from the viewpoint of the records lifecycle. Through mapping and classification between metadata elements and the records lifecycle, we identified the relationships between metadata standards and the lifecycle stages. In the second step, we proposed a task-centric model and created mappings among the metadata elements in each stage of the lifecycle using the 5W1H categories.

In the first step of the study, we identified the stage where a value is assigned. And then, we identified the lifecycle stage(s) for each standard where many of the elements are assigned values. The stage(s) identified by this process is called ‘primary stage’ of the standard. For example, many of the AGLS metadata elements are assigned their values in an early stage of the lifecycle and updated in a later stage when the archival status is changed. From this study, we found that no single metadata standard can cover the whole lifecycle but also that an in-depth analysis of mappings between metadata standards in accordance with the lifecycle stages is required. We found that most metadata standards are primarily resource-centric and the different tasks in the resource lifecycle are not reflected in the design of metadata standard data models. Because one or more metadata standards are used in the whole lifecycle, the mappings of metadata elements have a crucial role in making the metadata standards interoperable. This means that we need to map metadata elements across lifecycle stages.

In the second step of the study, in order to clearly show a resource task in the lifecycle and help create mappings among the metadata elements, we proposed a Task model (task-centric model) as a framework model based on the lifecycle. In the proposed Task model, a task is linked to resources by a 5W1H attribute(s). We used the 5W1H categories (Who, What, Why, When, Where, How), to identify feature(s) of each element according to a resource task. Also, the 5W1H attribute is used to categorize metadata elements in the Task model. This categorization is used in the mappings between elements of different metadata schemas.

We determined a set of keywords used in the classification of elements into the 5W1H categories and created mappings between every pair of element sets. We examined a semantic definition of metadata element terms in the standards to find what categorization term typically appears in the definition. This classification was carried out manually because of the need to interpret the meanings and intention of the explanations.

We extracted detailed contextual information from the lifecycle which is useful to create mappings among metadata elements. Contextual semantics are implicit in the definition of metadata elements. Tasks performed on a resource are crucial contextual information sources. In addition, we compared the elements from the six different aspects of the 5W1H categories in the task-centric model.

Creating a unified framework to understand the features of metadata standards is necessary in order to improve metadata interoperability that covers the whole resource lifecycle. In this study, we approached this issue from the task-centric view of metadata, proposed a Task model as a framework and analyzed the feature of archival metadata standards.

In conclusion, the proposed model provides a new scheme to create metadata element mappings to make metadata interoperable. We identified the relationship of metadata standards and tasks in the records lifecycle. We also learned that using the records lifecycle and tasks will help with metadata interoperability for long-term preservation of digital resource.

レコードのライフサイクルを基礎としたアーカイバル メタデータ標準の特徴分析に関する研究

概要

インターネットと WWW の発展により、ディジタルリソースが多様な環境で作成され、発信されるようになった。ネットワーク上の情報環境の発展と普及によるディジタルリソースの一般化が進む中で、いくつかの大きな問題も明らかになってきた。特に、将来に渡ってディジタルリソースを利用可能な状態に維持し、管理し続けること、すなわちディジタルリソースの保存がそうした問題の 1 つである。

メタデータはディジタルリソースのアーカイビングや長期間の保存において重要な技術的要素として広く認められている。ディジタルアーカイビングや保存のためのメタデータ標準として、AGRkMS、EAD、ISAD(G)、MoReq、PREMIS などがある。アーカイブシステムのメタデータスキーマを設計するために、我々は目的に沿ったメタデータ標準を選択してカスタマイズしなければならず、さらに、異なるシステムのメタデータ間での相互運用性に関しても考慮しなければならない。

メタデータ標準は、一般に、基盤とするデータモデルと、データモデルに含まれている実体の属性として定義されるメタデータエレメント（記述項目）を持っている。メタデータ標準は、検索、管理、保存など目的と記述対象の特性に合わせて作られるため、標準毎に異なる特徴を持っている。しかしながら、ディジタルリソースのアーカイビングや保存のために、一つのメタデータ標準だけでレコードのライフサイクル（作成から管理、保存、そして再利用まで）の全体をカバーすることは難しい。これは、レコードのライフサイクル全体をカバーできるメタデータスキーマを開発するには、ライフサイクルの各ステージをカバーするメタデータスキーマに対する要求を十分に理解したうえで、メタデータ標準を組み合わせる必要があることを意味する。

レコードのライフサイクルは、作成、管理、評価、保存という、いくつかのステージで構成されている。ライフサイクルの各ステージでは，リソースに対

して何らかの操作が実行される。本研究では、これをタスク（Task）と呼ぶ。例えば、タスクには、Edit、Copy、Search、Discard、Collect、Access などがある。ライフサイクルの中では、各タスクの目的や内容に従ってリソースに対する処理が行われる。

メタデータエレメントはリソースの属性として定義される。メタデータはレコードのライフサイクル中の各タスクで利用される。そして、メタデータの内容はライフサイクルのステージと各ステージ内のタスクによって設定されたり、変更されたりする。しかし、リソースのタスクとメタデータエレメントの関係はメタデータの定義や記述の一部として明示的に与えられていない。そこで、本論文では、レコードのライフサイクルステージとメタデータエレメントの関係を明確にしてアーカイブのためのメタデータ標準の特徴分析を行った。

この研究では、アーカイビングと保存のために開発されたメタデータ標準である EAD、ISAD(G)、OAIS、PREMIS を分析対象として利用した。さらに、著者はアーカイビングや保存のためのメタデータ標準とともに、それ以外の目的を持つ異なるメタデータ標準を利用して分析することが、メタデータ間の違いや比較をより明確に表すために重要であると考え、異なる特徴や目的を持ついくつかのメタデータ標準を選び、分析対象として利用した。本研究では、リソースの検索のためのメタデータ標準である AGLS Metadata Standard、記録管理のための AGRkMS、イギリスの Digital Preservation Coalition (DPC)により提案されたディジタル保存のための決定木（Decision Tree）から判断のための属性を抽出して作成した属性記述項目の集合を利用した。

本研究では、アーカイブのためのメタデータ標準の特徴を分析するため、2 つの観点で研究を行った。第 1 の研究（研究 1）ではレコードのライフサイクルの観点から主なメタデータ標準の特徴を明確に分析した。この研究を通じて、レコードライフサイクルのタスクとメタデータ標準間の関係を確認することができた。これを基にして、第 2 の研究（研究 2）ではタスクモデル（Task Model）を提案して、タスク中心の観点（Task-centric view）からメタデータエレメントセットの特徴分析を行った。

複数のメタデータ標準を組み合わせてレコードのライフサイクル全体をカバーするには、メタデータの相互運用を可能にするメタデータエレメントのマッピングが重要である。また、メタデータスキーマの相互運用性を向上させるためには統一されたフレームワークを構築することが重要である。そのため、研

究 1 ではレコードライフサイクルのステージを基準とし、各ステージに対してメタデータエレメントをマッピングすることを試みた。

このメタデータのマッピングでは、メタデータ標準に従ったワークフローから「メタデータの作成、修正やアップデート」という点に注目し、各メタデータエレメントの値が決まるライフサイクルステージをプライマリステージ（Primary Stage）として定義した。例えば、行政機関が提供するリソースの発見と利用のために作られた標準である AGLS Metadata Standard では、エレメントの大部分がレコードのライフサイクルの Use & Manage ステージで値を割り当てられていることを確認し、Use & Manage ステージを AGLS のプライマリステージとした。

研究 1 で行ったレコードのライフサイクルモデルの観点からのメタデータ標準の特徴の分析において、レコードライフサイクルの中で各メタデータ標準が対応付けられるステージを識別することができた。そして、メタデータ標準の分析を通じて、著者は単一のメタデータ標準だけではライフサイクル全体をカバーすることができないことを明確にした。さらに、ライフサイクルのステージに従ってメタデータ標準間のマッピングに対する詳細な分析が必要であることを知った。

従来のメタデータ標準はリソースを記述対象とし、メタデータ標準が持つデータモデルのデザインにはレコードのライフサイクルやステージが反映されていない。しかし、著者は、研究 1 を通じて、多くのメタデータ標準は主にリソース中心（Resource-centric view）に定義される一方、その利用がレコードライフサイクルのステージと関係することを、プライマリステージに基づく分析によって確認した。言い換えると、こうしたメタデータ標準はリソース中心の観点から定義されるのみで、レコードのライフサイクル（ライフサイクルのタスク）との関係について定義されていない。そこで、研究 2 ではアーカイブのためのメタデータ標準の特徴分析のためにメタデータエレメントをタスク中心の観点から分析した。

メタデータエレメント間のマッピングを支援してレコードライフサイクルの中でリソースに対してなされるタスクを明確に表すために、メタデータエレメントをタスク中心の視点でとらえる、メタデータスキーマのためのタスクモデル（Task Model）を提案した。

タスクは権利、時間、目的、機関、人といったメタデータ記述に関わる実体

にリンクされている。これらの実体はタスクの中で何らかの役割を果たすエンティティであると言える。例えばタスクが行われる場所や組織、タスクを行うための理由や目的などがある。本研究ではタスクに関連付けられている実体（タスクに関連する人、場所、フォーマットなど）を表す一般化されたカテゴリとして、5W1H カテゴリ（Who, What, Why, When, Where, How）を利用することを提案した。さらに、5W1H カテゴリは各タスクに従って各メタデータのエレメントの特徴を明確にして分類するために利用した。

異なるメタデータ標準間でメタデータエレメントのマッピングと分類を行うために、本研究ではタスクモデルと 5W1H カテゴリを基盤として、それらの特徴を表すキーワードセットを定義した。マッピングと分類は 2 つのステップで行った。まず、各メタデータエレメントの値の内容を表すドキュメンテーション（定義、記述、ガイドラインなど）に含まれるキーワードを探し、その後、キーワードが該当するタスクモデルと 5W1H カテゴリに各エレメントを対応付けた。マッピングのためのエレメント同士の比較はあらかじめ決めた基準に基づいて行ったが、メタデータエレメントの意味解釈の必要性のために、マッピングと分類をすべて手動で実行した。

結論として、メタデータ標準間の関連を表すための統一的なフレームワークを作成することはレコードのライフサイクル全体でのメタデータの相互運用性を向上するために必要である。この研究ではメタデータの相互運用を改善するモデルの新しいツールとして、タスクモデルを提案した。さらに、メタデータエレメントの意味を分析的にとらえるための 5W1H カテゴリを提案した。そして、レコードのライフサイクルとタスクモデル、5W1H カテゴリを利用して、アーカイブのためのメタデータ標準の特徴分析することができた。

アーカイブのためのメタデータ標準の特徴分析を通じて、メタデータ標準はリソースに対して行われるタスクと関係があることを識別できた。タスクとメタデータの関係を明示的にとらえて分析することは、ライフサイクル全体の中でタスク毎に異なるメタデータ標準のエレメントを選択して利用するための新しい観点として有用である。以上のように、本研究では、レコードのライフサイクルを基礎としてタスクの視点からメタデータ標準の特徴を分析しなおすことが、ディジタルリソースのアーカイビングや保存のためのメタデータの相互運用性の向上に役立つことを示した。

Table of Contents

Abstract (English) -- i
Abstract (Japanese) --- iv
1 Introduction -- 1
2 Models and Standards for Archiving and Preservation - Literature Reviews -------- 7
2.1 Definitions and Descriptions -- 7
2.2 Record management - Recordkeeping, Archiving and Preservation ---------- 9
2.2.1 Recordkeeping System – DIRKS -- 9
2.2.2 Open Archival Information System – OAIS -------------------------------- 9
2.3 Selection of Digital Materials for Long-term Retention - Decision Tree Interactive Assessment -- 11
2.4 Lifecycle of Records - NARA Lifecycle -- 12
2.5 Metadata Standards and Tasks in the Lifecycle -------------------------------- 14
2.6 Properties and Characteristics of Archival Metadata -------------------------- 17
2.6.1 Metadata Standards for Archiving and Preservation --------------------- 18
2.6.2 Decision Tree for Long-Term Retention of DPC ------------------------- 24
2.7 Models for Metadata Interoperability -- 25
2.7.1 Application Profile --- 25
2.7.2 Vocabulary Mapping Framework -- 27
2.8 Related Works --- 27
2.8.1 Related Researches -- 28
2.8.2 Related Standards -- 32
3 From Resource-centric to Task-centric View of Metadata Schema----------------- 34
4 Feature Analysis of Metadata Schemas based on Lifecycle ------------------------- 38
4.1 Viewing Differences from Descriptive Elements ------------------------------ 38
4.2 Analysis Method --- 39
4.3 Mapping to Determine the Primary Stages in the Lifecycle ------------------- 40
4.4 Consideration -- 48
5 Facet Analysis of Archival Metadata Schemas for Metadata Interoperability ----- 49
5.1 Introducing a Task-centric View of Archival Metadata Standard------------- 49

5.1.1 Task-oriented View of Records Lifecycle – Task Model --------------- 49
5.1.2 Task-centric View of Metadata Schemas --------------------------------- 51
5.1.3 Combination of Task-oriented Model and Metadata Elements -------- 52
5.1.4 5W1H Categories -- 54
5.1.5 Discussion on Resource-centric and Task-centric Views of Metadata Elements --- 54
5.2 Analysis Criteria of Classification --- 55
5.2.1 Vocabulary of Systematic Classification by 5W1H Categories and Lifecycle Tasks -- 55
5.2.2 Classification Procedure -- 57
5.3 Mapping Metadata Schemas in 5W1H Categories and Lifecycle Tasks ----- 59
5.3.1 Classification of Descriptive Elements in 5W1H Categories ---------- 59
5.3.2 Mapping in the Task Groups --- 63
5.3.3 Analysis of Metadata Schemas using 5W1H Categories and Task Group --- 65
5.4 Consideration -- 68
6 Discussions --- 70
6.1 Related Researches on Metadata for Archiving and Preservation ------------ 70
6.2 Metadata Standards for Archiving and Preservation -------------------------- 72
6.3 Feature Analysis of Archival Metadata Standards ------------------------------ 73
6.4 Task-oriented Model and 5W1H Categories ------------------------------------ 74
6.5 Metadata Mappings based on the Task Model ---------------------------------- 75
6.6 Discussion Summary -- 76
7 Conclusion -- 78
Acknowledgements --- 81
Reference --- 84
List of Publications -- 91
Appendices --- 92
Appendix1. A Classification of Metadata Standards in the Records Lifecycle (ISAD(G) & Decision Tree) --- 94
Appendix2. A Classification of Metadata Standards in the Records Lifecycle (EAD & OAIS) --- 95
Appendix3. A Classification of Metadata Standards in the Records Lifecycle

(AGLS & PREMIS) -- 96
Appendix4. A Classification and Mapping between Metadata Standards in the 5W1H Categories -- 97
Appendix5. A Classification and Mapping between Metadata Standards and 5W1H Categories in the Task Model -- 108

List of Figures and Tables

1) Figures

Figure1. Scope of Record Management -- 8
Figure2. OAIS Functional Entities -- 10
Figure3. Decision Tree of DPC -- 12
Figure4. Lifecycle of NARA --- 13
Figure5. Lifecycle of This Research -- 14
Figure6. Relationship of Metadata Standards and Tasks in the Lifecycle --------------- 16
Figure7. Characteristics of Metadata for Archiving and Preservation -------------------- 17
Figure8. High-level Five Entity Model --- 20
Figure9. Model of the Levels of Arrangement of Fonds ----------------------------------- 21
Figure10. The PREMIS Data Model --- 22
Figure11. The Structure of an Information Package --------------------------------------- 23
Figure12. Question and Attributes in the DPC Decision Tree ----------------------------- 24
Figure13. Singapore Framework of Dublin Core --- 26
Figure14. The Mapping of User Task and Attributes in the FRBR ----------------------- 34
Figure15. From Resource-centric to Task-centric View of Metadata Schema ---------- 37
Figure16. Lifecycle and Metadata Standards for Archiving and Preservation ---------- 38
Figure17. Classification Criteria of Metadata Standards into the Records Lifecycle using the Step 1 and 2 --- 40
Figure18. An Example of using the Records Lifecycle for the Description Elements -- 44
Figure19. Stage of Lifecycle shown by Metadata Description Elements ---------------- 47
Figure20. Task Model -- 50
Figure21. Resource and Task-centric View of Metadata ----------------------------------- 51
Figure22. Task-centric View of Metadata and 5W1H -------------------------------------- 53
Figure23. Mapping of Metadata Elements in 5W1H Categories (Step 1) --------------- 57
Figure24. Mapping and Classification of Metadata Elements in the Task Model and 5W1H Categories (Step 2) -- 58
Figure25. Term Definition of AGLS Element -- 59

2) Tables

Table1. An Example of AGLS Metadata -- 41
Table2. An Example of Decision Tree -- 41
Table3. An Example of EAD -- 42
Table4. An Example of ISAD(G) --- 42
Table5. An Example of OAIS --- 43
Table6. An Example of PREMIS --- 43
Table7. Metadata Standards shown by Figures (%) --------------------------------------- 46
Table8. Classification Vocabulary with 5W1H Categories ------------------------------- 56
Table9. Classification Vocabulary with Task Model -------------------------------------- 56
Table10. AGLS: *Publisher* --- 60
Table11. AGRkMS: *Date Range, Start Date* and *End Date* ------------------------------ 61
Table12. DPC Attribute Set: *Multiple media formats* -------------------------------------- 61
Table13. EAD: *Title of the Unit* -- 62
Table14. OAIS: *Reason for Creation* -- 62
Table15. PREMIS: *Size* -- 63
Table16. Example of Mapping of Metadata Schemas for Archiving and Preservation -- 64
Table17. Metadata Standards in 5W1H Categories shown by Figures ------------------- 65
Table18. Metadata Standards in the Task Groups shown by Figures --------------------- 66
Table19. Metadata Standards in the Task Groups by Percentage ------------------------ 67

1 Introduction

An information society starts with generalization and dissemination of WWW and popularization of personal computers and the Internet since 1990s. The rapid growth of the Internet and WWW, a quantity of information resources could constantly produce and receive in the various forms. In our modern information environment, we cannot imagine our daily lives without digital resources and ubiquitous networks.

No longer are the main information resources (materials) ‘documents printed on paper’ or ‘material written on paper’. Currently, most resources are ‘documents created using a computer or the Web’ or ‘resources sent out on a network’. In other words, resources are of two types: non-digital and digital. A resource created and circulated in a digital form is common due to the change of environment, machines and technology. In this paper, we use the term digital resource to mean a digital resource which may be born digital or converted into digital.

Digital resources have their own problems of management and preservation. The increased usage of digital resources has brought us serious demands to preserve the digital resources over time, even though the media on which information resources are stored is continuously changing and it is well known that archiving and preservation of digital resources is not straightforward. The problem is not only the quantitative, but also how to preserve a digital resource in its original form for the next generation. There are also the problems of storage, preservation and reuse of digital resources in the future. In particular, digital resources made in a variety of forms on electronic media are quickly changed by the progress of information technologies. In other words, we need a number of solutions for long-term preservation and management of non-digital and digital resources for the future.

There are researches in various fields about archiving and preservation of digital resources, especially for the institutions known as memory institutions such as libraries and archives. Memory institutions that are responsible for the long-term management and preservation of digital resources are keen to develop systems for digital preservation. They - governments, industries and universities - are also developing and using policies, guidelines, management and technology strategies, for their selection and preservation of digital resources. Nevertheless, it is more and more difficult to maintain digital resources as time goes.

For the long-term preservation and archiving of a digital resource, many factors have to be taken into account to develop the policies and methods; evaluation and prioritization to select resources for preservation, laws and regulations for digital preservation, preservation technologies such as migration and emulation, metadata schemas for digital preservation. In general, preservation policies and strategies have to be clearly defined in accordance with the type of resources to be preserved and the purpose of preservation.

On one hand, a number of factors relevant to different aspects have to be examined in order to preserve digital resources. On the other hand, it is too complicated to examine all of the factors at the same time. In this paper, we study metadata for preservation and archiving, which is widely recognized as one very important issue for digital preservation [5].

A metadata standard is well recognized as one of the important components required in the creation, management, recordkeeping, archiving and preservation of digital resources. Metadata standards are usually designed for a specific purpose and used in different services, e.g., searching resources, rights management, and accessibility control. There are many major metadata standards used for management, recordkeeping, archiving and preservation of digital resources, e.g. Dublin Core, AGLS, AGRkMS, EAD, ISAD(G), METS, MoReq2, OAIS, PREMIS, and more.

Metadata schema for purposes such as finding aids, rights management and accessibility descriptions are used in accordance with the requirements of a particular stage of the resource's lifecycle. Metadata schema is related to different resource tasks throughout the whole resource lifecycle. They are created and revised by resource tasks and change according to the content and purpose of the tasks. Resources perform different tasks according to the stage of their lifecycle, which means that metadata associated with the resource needs to change. We need appropriate metadata schemas related to the lifecycle stage. We need guidelines to select appropriate metadata standards and to define profiles for the tasks and stages based on the metadata standards. However, most metadata standards do not explicitly mention the resource lifecycles or tasks. In other words, it is not explicitly defined when a descriptive element should be assigned or where its value should be revised in the lifecycle.

For example, PREMIS has five types of entities in its data model – intellectual entity, digital object, event, right and agent and elements. 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.

A major question is whether a single standard is sufficient for digital resource preservation. If we have to use multiple metadata schemas, we have to have an appropriate framework to enhance the interoperability between the schemas. In practice, multiple metadata standards are frequently used in a single system, e.g. descriptive metadata, administrative metadata and technical metadata. From another viewpoint, it is crucial to record information about a resource from the moment when the resource is created and to maintain the information in accordance with tasks required in every stage of the lifecycle of the resource. Thus, we naturally use more than one metadata schema in the record management and archiving process [5].

In the current information environment, where various types of resources coexist with heterogeneous formats of metadata standards, efforts have been made to achieve metadata interoperability in order to utilize multiple metadata standards. These efforts have generated different approaches to minimizing differences between the heterogeneous standards and maximizing consistency across them, including element mapping, crosswalks, application profiles, and the use of a metadata registry [12].

A single standard may or may not be suitable for a particular service. For interoperability and exchange of metadata standards, Application Profiles offer a framework for designing metadata applications [46].

Metadata vocabulary mapping is not new. There are notable examples such as VMF [24]. However, these mappings do not explicitly use the lifecycle to identify the semantics of the metadata elements. Metadata vocabulary mapping is primarily required for the interoperability of metadata.

The Vocabulary Metadata Framework (VMF) is used for the mapping of vocabularies from major metadata standards. VMF is designed as a tool to automate finding the 'best fit' mapping between terms in controlled vocabularies in different metadata schemes [36]. This means that, on one hand, we need to appropriately choose one or more metadata standard(s) and 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.

Based on the observation about metadata schemas for archiving and preservation of digital resources, we explain and propose a methodology to analyze metadata schemas in order to help selection and combination of metadata schemas used throughout the whole lifecycle, i.e. from creation to preservation and re-use. Specifically, we analyzed the relationship between a resource task and available metadata schemas for digital archiving and preservation.

A metadata standard is generally focused on resources from the viewpoint of the purpose of description. Mapping metadata standards using each stage of a lifecycle is not a suitable method. In order to analyze the features of archival metadata, we examined the relation between the metadata standards and the stages of a lifecycle. We propose a mapping method between metadata standards in order to link between the different metadata standards and the tasks within the stages of a lifecycle. We did a detailed analysis from the viewpoint of the task of a resource. This paper proposes a framework to characterize descriptive elements of metadata vocabularies and improve mapping among them.

First, we analyzed relationships between the lifecycle stages and the metadata standards by an analysis of patterns based on the lifecycle. From the crosswalk and mapping between metadata and the stage of a lifecycle, we examined the stages and identified a stage for every element where an initial value of the element is given, a stage where the value of the element is updated, and a stage where a particular metadata standard is most frequently used. In the first research, we showed that a descriptive element should be chosen appropriately and combined according to the task within the stage of a lifecycle. And we have learned that no single metadata standard covers the whole lifecycle.

Based on our first research, we proposed a Task model, a framework based on the resource lifecycle for a more detailed analysis of the element sets and mapping among them. Despite the fact that a metadata element is assigned value in a particular task, the relationship between the element and the task is not explicitly defined in conventional metadata standards. Descriptive elements are primarily defined as attributes of a resource and relationships between the resource and the tasks are not explicitly given as a part of the definition but may be given as a part of the usage guidelines. Our first study showed the need for metadata mapping over the lifecycle. However, contextual information used in every task is rarely used in the mapping of metadata elements which ignore the lifecycle.

In the second research, we proposed a task-oriented model based on a task-centric point of view for more detailed analysis of the element sets. We clarified the viewpoint of an 'Event' performed within a task, using the 5W1H attribute set (what, why, where, who, when, how) and, used it in order to categorize a metadata element in the context of each task where the element is used. The Task model and the 5W1H attribute set are important to narrow the scope of mapping and categorizing in order to perform efficient mapping between descriptive elements focusing on a task.

For this research, we used attribute sets from AGLS, AGRkMS, EAD & ISAD(G), PREMIS, the archiving system of OAIS, and a set of attributes extracted from the decision tree for a preservation process defined by the Digital Preservation Coalition (DPC).

In order to show the features of archival metadata standards, the author thinks that an analysis using various metadata standards shows a clearer difference when comparing of metadata. So we have chosen AGRkMS, EAD & ISAD(G), PREMIS form as typical standards in their particular domains. Although the AGLS, OAIS, and DPC attribute sets are not designed as metadata schema for archiving or preservation, we have included them as comparable objects in order to show the characteristics of archival metadata standards. Also, in order to analyze the relationships between a resource task and the metadata standards, we used the records lifecycle of NARA.

We examined the semantic definitions of each element to find what categorization terms typically appear in the definitions, and then we classified every element into 5W1H categories. This paper shows the two mappings and classifications.

This paper is organized as follows. Section 2 describes and arranges a fundamental concept - metadata standards for archiving and preservation of digital resources, records lifecycle model, and literature reviews, as the background. Section 3 explains the relation of a task and the metadata standard, and the definition - role, scope etc. - of a resource task. Section 4 shows the feature analysis of archival metadata standards from a viewpoint of a resource lifecycle, according to the first research. Section 5 explains about the feature analysis for interoperability of a metadata standard, and proposes the basic models - the 5W1H categories and the Task model – and, shows several example mappings among the standards, following the second research. In section 6 and 7, we have some discussion and our conclusions.

2 Models and Standards for Archiving and Preservation - Literature Reviews

2.1 Definitions and Descriptions

This chapter describes the definitions of terms used in this paper - Record, Record management, Recordkeeping, Archives, Preservation, a Task and the Records Lifecycle.

A record is "recorded information, regardless of medium or characteristics, made or received by an organization, and has value requiring its retention for a specific period of time" [37]. In this dissertation, 'resource' is used as a term which has a broader meaning of 'record' because some metadata schemas do not use the term 'record' but 'resource', e.g. AGLS.

In the lifecycle of resources at an organization, a record is created, used and managed by the policy, rules, guidelines given by the organizations. The records lifecycle is composed of several stages, such as creation, management, appraisal, preservation and so on. The records lifecycle is a model that shows tasks performed on a resource, according to specific stages. In a stage of the records lifecycle, a process or operation is performed on a resource in accordance with the content and purpose of each task. We call these processes or operations 'tasks'. A task can be an action such as Edit, Copy, Search, Discard, Collect, Access.

Record management is "the systematic control of all organizational records during the various stages of their lifecycle: from their creation or receipt, through their processing, distribution, maintenance and use, to their ultimate disposition. The purpose of records management is to promote economies and efficiencies in recordkeeping, to assure that useless records are systematically destroyed while valuable information is protected and maintained in a manner that facilitates its access and use" [19].

Created record is used and managed in record management. This step is called Recordkeeping. Recordkeeping is defined as "the making and maintaining complete, accurate and reliable evidence of business transactions in the form of recorded information" [58]. A system that performs record management is a recordkeeping system. A recordkeeping system is "a manual or automated system that collects, organizes, and categorizes records, facilitating their preservation, retrieval, use, and disposition" [67]. Records must be appraised, stored and preserved for long-term archive. These steps are called Archiving and Preservation.

"An archives is a place where people can go to gather firsthand facts, data, and evidence from letters, reports, notes, memos, photographs, and other primary sources" [40]. Also, an archive is defined as a service "to transfer records from the individual or office of creation to a repository authorized to appraise, preserve, and provide access to those records" [57]. In archive step, record is managed by archives system. "An archive system provides a full service, offsite, business records storage solution, which empowers you to manage the document lifecycle from Source-to-Shred " [1].

Archival service performs to preserve resources for long-term in the archive step. "Preservation encompasses the activities which prolong the usable life of archival records. Preservation activities are designed to minimize the physical and chemical deterioration of records and to prevent the loss of informational content" [39]. "Preservation is the means by which archives are protected for the use of present and future generations. It is a word commonly used by record offices, libraries and museums to describe the ways in which their collections are safeguarded and kept in good physical condition. This can be done through a variety of measures aimed both at minimizing the risk of loss of records and slowing down, as much as possible, the processes of physical deterioration which affect most archive materials" [53].

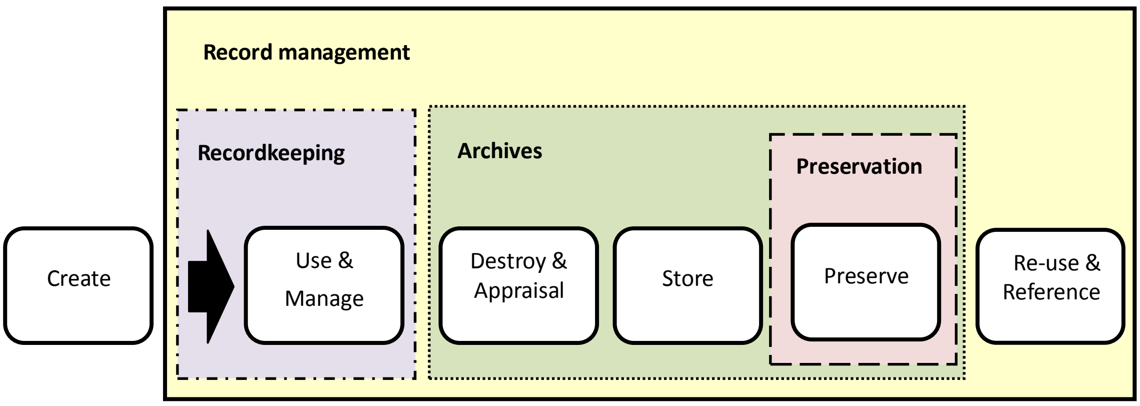


Figure1. Scope of Record Management

2.2 Record management

- Recordkeeping, Archiving and Preservation

2.2.1 Recordkeeping System - DIRKS

"The Designing and Implementing Recordkeeping Systems (DIRKS) is about building more efficient and accountable business practices through the design and encouragement of good recordkeeping across an organization" [59]. The DIRKS is composed of a methodology and manual.

The DIRKS methodology is a clear and simple statement contained and outlined in the Australian Standard on Records Management, AS ISO 15489-2002. The DIRKS methodology provides advice on how to identify appropriate recordkeeping strategies [59].

"The DIRKS Manual is a product developed by State Records to explain in a significant amount of detail how public offices can implement the methodology, in order to improve their recordkeeping practices" [60]. The DIRKS methodology is composed of eight steps, and the eight steps outlined in the DIRKS methodology is explained in detail in the DIRKS Manual.

Eight steps in the DIRKS methodology are Step A - Preliminary investigation, Step B - Analysis of business activity, Step C - Identification of recordkeeping requirements, Step D - Assessment of existing systems, Step E - Identification of strategies for recordkeeping, Step F - Design of a recordkeeping system, Step G - Implementation of a recordkeeping system, Step H - Post implementation review.

2.2.2 Open Archival Information System - OAIS

Open Archival Information System (OAIS) is an international standard for preservation of digital resources and is reference model of archival systems, defining concepts and responsibilities essential for ensuring preservation of digital information. The feature of OAIS is its categorization of information packages by their function (Submission Information Package, Archival Information Package, Dissemination Information Package) [14], [26].

An information package consists of "the digital object that is the focus of preservation, along with metadata necessary to support its long-term preservation and access." There are comprised of three information package: the Submission Information Package, the Archival Information Package, and the Dissemination Information Package [11]. "The SIP is sent from the information producer to the archive, the AIP is the information package actually stored by the archive, and the DIP is the information package transferred from the archive to a user in response to an access request" [48].

The AIP is the version of the information package that is stored and preserved by the OAIS. Within the AIP is an Information Object called the Preservation Description Information (PDI). The PDI contains additional information about the Content Information and is needed to make the Content Information meaningful for the indefinite long-term. The OAIS reference model identifies four types of PDI: Reference Information, Provenance Information, Context Information, Fixity Information [10], [48], [54].

"The OAIS reference model is a conceptual framework for a digital archive. The model establishes terminology and concepts relevant to digital archiving, identifies the key components and processes endemic to most digital archiving activity, and proposes an information model for digital objects and their associated metadata" [47].

The OAIS reference model is "a particular focus on digital information, both as the primary forms of information held and as supporting information for both digitally and physically archived materials" [50]. The OAIS reference model is designed as a conceptual framework and, outlines the functions required to access information objects

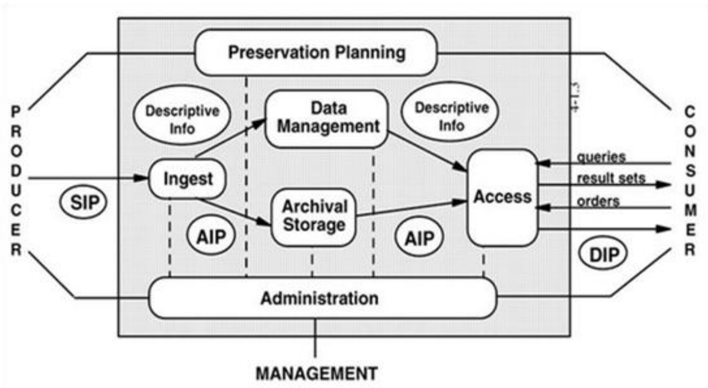


Figure2. OAIS Functional Entities
(From *Reference Model for an OAIS* of CCSDS)

and guarantee efficient long-term preservation [14]. “The reference model provides a framework for the understanding and increased awareness of archival concepts needed for long-term digital information preservation and access” [34].

2.3 Selection of Digital Materials for Long-term Retention - Decision Tree Interactive Assessment

The Digital Preservation Coalition (DPC) promotes information sharing and activities for long-term access of digital resources to reduce the obstacles in the way of preservation of resources. DPC has been working for preservation of digital resources from various viewpoints, and has suggested the guidelines for digital preservation in the Digital Preservation Handbook (DPH).

The DPC handbook provides an internationally authoritative and practical guide to the subject of managing digital resources over time and the issues in sustaining access to them. It will be of interest to all those involved in the creation and management of digital materials.

DPH shows a *decision process* for the selection of digital materials for long-term retention, which is called the Decision Tree. Clearly defined selection policies (decision processes) will enable cost savings in terms of time taken to establish whether or not to select and also potential costs further down the track of needing to re-assess digital resources which are either in danger of becoming or are no longer accessible [17].

The Decision Tree may be used as a tool to construct and test such a policy for each organization. The decision process represented in the tree should be addressed by each policy for selection of digital materials for the long-term. The decision process shows an evaluation process for the resources in the form of *Questions and Choices* [17]. The *Questions and Choices* assist in the ultimate decision to accept or reject long-term preservation responsibility.

The decision tree is composed of three sections - *Rights & Responsibility*, *Technology & Metadata*, *Documents & Costs*. Each section is expressed as a sub-tree of the whole process. The decision tree is composed of questions and answers - a question is a node and an answer is an edge coming from the node. An advice may be attached to a node as an answer to the question. And an advice may be attached to a node as an answer to the question.

As mentioned earlier, DPC has the character of 'Process of selection and evaluation' of digital materials for long-term retention, although the DPC attribute set is not designed as a metadata standard. We need to evaluate a resource and to find suitable technologies and strategies for long-term preservation. Therefore, such processes are necessary to support tasks for digital archiving and preservation. The selection process (policy) is also needed and used in the records lifecycle.

We used the decision tree (DPC attribute set) as a metadata attribute that represents the selection stage in lifecycle. We explain the extraction of the metadata attribute from the decision tree, in section 2.6.2.

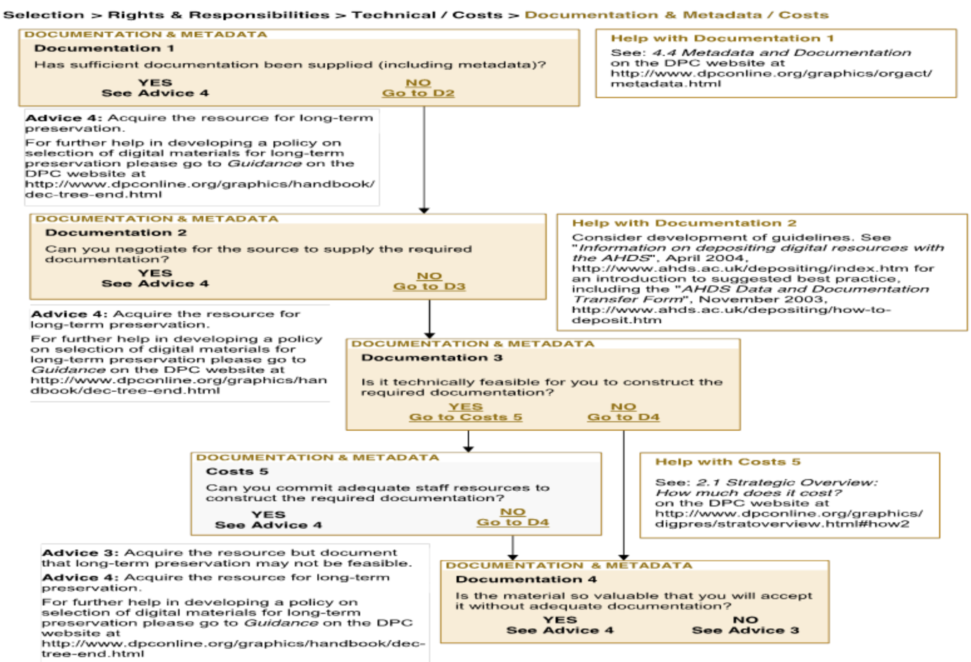


Figure3. Decision Tree of DPC

(From *Decision Tree for Selection of Digital Materials for Long-term Retention* of DPC)

2.4 Lifecycle of Records - NARA Lifecycle

Huge numbers of documents and records are created and disseminated everyday by various organizations and institutions. All of those resources are created, used, preserved and destroyed in accordance with the management processes determined by the individual organizations [41]. Each resource has a lifetime composed of a set of stages known as a records lifecycle.

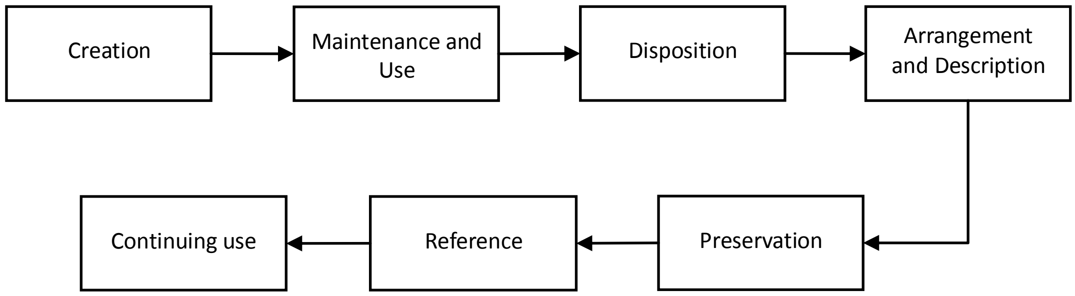


Figure4. Lifecycle of NARA
(From *What's a Record* of NARA)

The model of the records lifecycle used in this paper is based on that of the National Archives and Record Administration (NARA) of the US government. As shown below (Figure 4), the NARA's records lifecycle has seven stages defined independently from any resource types, such as digital resources, official documents, archives and national records, and also from any media types such as pictures, maps, photos, and videos. The paragraphs below explain the stages of the NARA lifecycle.

1) Creation

   Records are created by persons or departments that belong to various organizations and institutions.

2) Maintenance and use

   While in use, the record is collected, arranged and stored with similar records.

3) Disposition

   Records are kept according to the record schedule in the organization. And a record is evaluated at this stage. The records appraised are permanently preserved in the National Archives.

4) Arrangement and description

   Administrative information (metadata) is given to the records according to the management policies of the National Archives.

5) Preservation

Records should be preserved without losing anything. Meanwhile, there are additional reasons to change the media.

6) Reference

Supply the records preserved to provide search and reference services.

7) Continuing use

Proper management and continuing use of preserved records is promoted.

In this study, we merge the last two stages of NARA's lifecycle into one and define the resource lifecycle model as shown in Figure 5 because both of the last two stages, Reference and Continuing Use, mean use of the archived resources. This resource lifecycle model was used for feature analysis and we used this resource lifecycle model to define the Task model of the resource lifecycle.

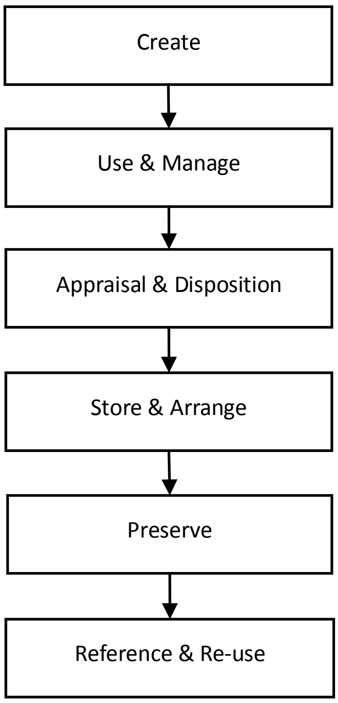


Figure5. Lifecycle of This Research

2.5 Metadata Standards and Tasks in the Lifecycle

A resource is affected by tasks in the lifecycle. The lifecycle includes several stages such as create, use, archive, preserve and so on. Tasks performed on the resources differ by lifecycle stage. For example, a resource search can be performed at all stages but revision is primarily done only in the creation stage in the case of records management.

There are several purposes of metadata schemas, such as resource discovery, recordkeeping, archival, preservation, and resource management. The metadata is used according to the content (or purpose) of each task in the lifecycle. For example, archival and preservation metadata schemas are used primarily to manage resources in accordance with the resource lifecycle.

As a resource is used in different tasks throughout the whole lifecycle, it is obvious that we need a metadata to clarify what resource attributes should be described in accordance with the tasks. That is, we need appropriate metadata schemas that correspond to the lifecycle stages. Figure 6 shows the relationship between a task and metadata schemas according to the lifecycle.

A resource is handled according to different tasks in each stage of lifecycle, and described by various metadata elements. For example, resources created in the 'Create' stage is described using AGLS Metadata elements for searching, using and management, such as Title, Creator etc. In addition, a resource in the 'Preserve' stage is described using PREMIS for long-term preservation, with used metadata elements in previous stages, e.g., AGLS Metadata, AGRkMS, EAD and so on. The relationship between metadata standard and task are shown in detail in chapter 5.

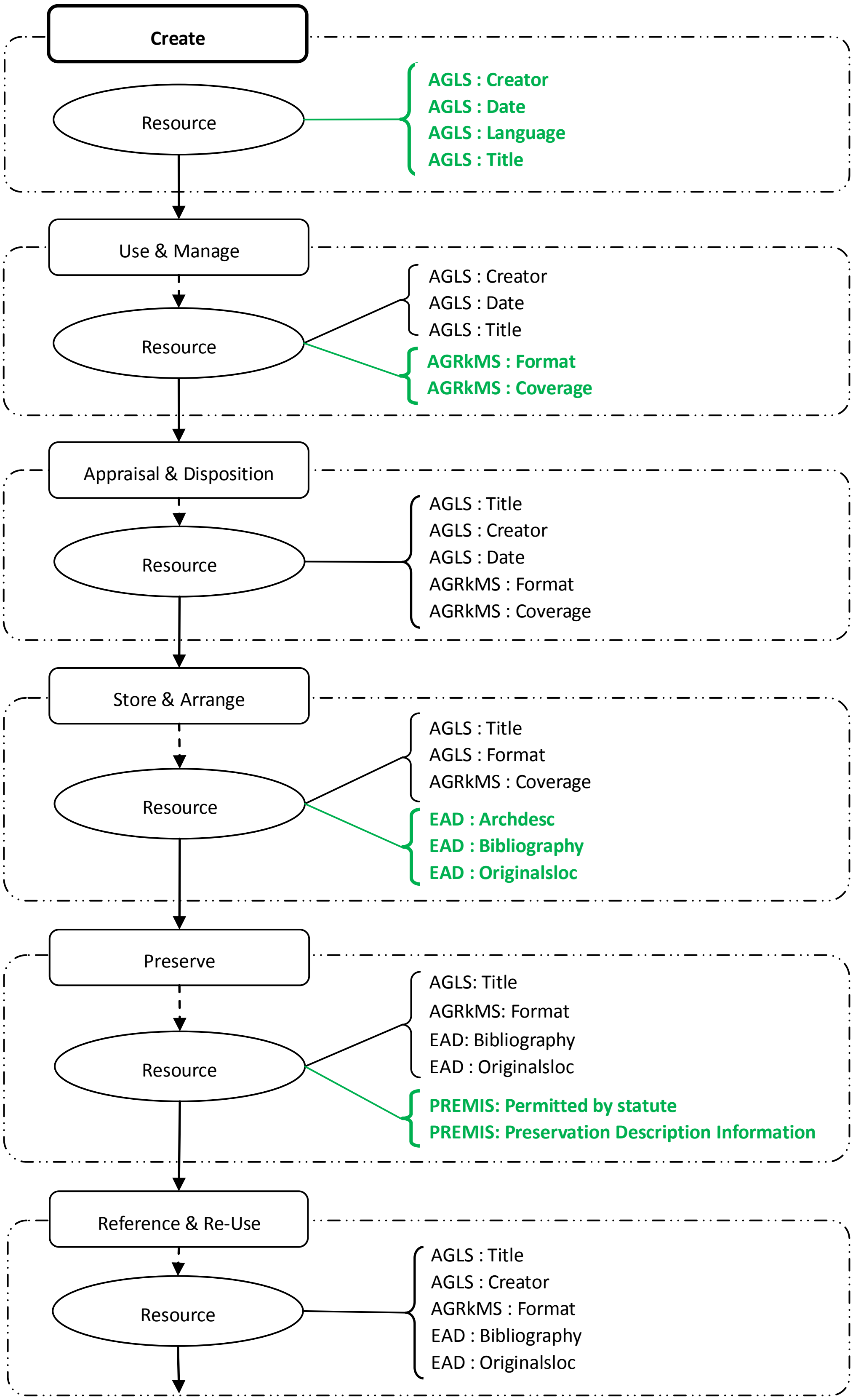


Figure6. Relationship of Metadata Standards and Tasks in the Lifecycle

2.6 Properties and Characteristics of Archival Metadata

Each metadata standard for archives has its own set of elements and controlled vocabularies. A typical metadata description contains elements such as title, creator, related resource, history of resource. Technical metadata explains the technical features of a resource, such as data for management, format, media, hardware and so on. The paragraphs below show details of the descriptive and technical metadata.

As a typical metadata of archives and preservation, we analyze the features using four metadata – EAD, ISAD(G), OAIS, PREMIS. ISAD(G) contains descriptive elements of resources in an appropriate granularity, i.e., fond, sub-fond, series, file, and item. EAD and OAIS have elements to describe intellectual content, structural features, administrative and technology information. Intellectual content obviously needs descriptive metadata and technology information is in technical metadata. Structural and administrative information have both descriptive and technical features. PREMIS has many elements to describe the technical features and structure of the digital resources. Figure 7 shows the features of these four standards [4].

In this analysis for the metadata elements, we have shown that, on one hand, these metadata schemas have common features, but on the other, they have different features determined by their objectives and purposes. This means that it is crucial to select and use appropriate metadata standards and combine them appropriately when designing a metadata schema for a specific archival system. In other words, the crucial metadata

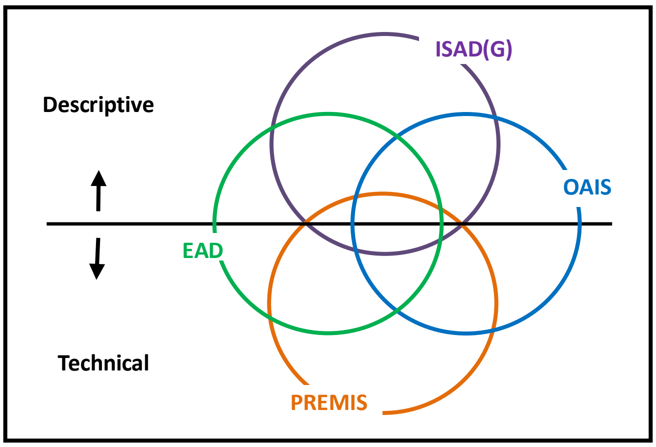


Figure7. Characteristics of Metadata for Archiving and Preservation

issues for the archival system are to create mappings between the lifecycle stages and the metadata standards and to create mappings between the metadata elements of different schemas used in the system. Therefore, a unified framework to enhance interoperability of metadata standards is crucial for digital preservation and archiving [5].

2.6.1 Metadata Standards for Archiving and Preservation

Describing a resource is "essentially about describing information resources using a standard framework or set of principles." Metadata is concerned with digital information management, as an essential component of the evolving networked information environment [2], and it is used to describe information that characterizes data.

Metadata is an essential component of any good recordkeeping system, digital preservation methods. Metadata also includes a wide variety of structured information that can be used to identify, as used in the current context of recordkeeping [45].

"Metadata properly facilitate the long-term access of the digital resources by explaining the technical environments needed to view the works, including applications and version numbers needed, decompression schemes, and other files that need to be linked to them, among others" [8].

Archival metadata is defined as the information to describe, manage and identify the structure of digital resources in order to preserve the resources over time [25]. Also preservation metadata provide much needed information required to manage the long-term preservation of digital resources and is a strategy to provide sufficient technical information about the resources [2], [8].

In this study, we used widely known metadata standards for recordkeeping, record management, archiving and preservation. In order to propose a new model to clarify the features of those standards, we have chosen AGRkMS, EAD, ISAD(G), OAIS and PREMIS from these standards as typical standards in their particular domains – i.e., AGRkMS for record keeping, EAD for archives, and OAIS and PREMIS for digital preservation. OAIS does not define a metadata element set in itself. We used the element set of CEDARS preservation metadata as the CEDARS set was drafted in close consultation with the OAIS reference model, to the extent that the elements borrow the concepts, terminology, and organization embedded within the OAIS framework.

In addition to these standards, we included AGLS and the DPC attribute set which are not archival metadata standards but contain metadata elements used in the lifecycle – AGLS for resource discovery, DPC for appraisal. The next paragraphs briefly introduce these metadata standards referred to in this study.

1) AGLS

Australian Government Locator Service (AGLS) Metadata standard is to refer descriptive information about resources, and it is known as resource discovery metadata. AGLS Metadata was designed to facilitate, discover and search resources by users online and, was used to improve the visibility and discoverability of Australian government resources in the online environment.

AGLS Metadata Standard provides a set of metadata properties, policies and guidelines defined for a particular application or implementation, and metadata property set consists of 60 properties. AGLS Metadata Standard associated usage guidelines to improve the visibility, manageability and interoperability of online information and services. “This is for use by any organization or individual creating or managing information sources or services that are locatable via the Internet. In particular, it is intended for information about resources and services on the World Wide Web”.

AGLS Metadata aims to improve the search of both digital and non-digital resources supplied by the Australian Government, and resources include documents, images, sound, video, physical objects, people and services [42].

2) AGRkMS

Australian Government Recordkeeping Metadata Standard (AGRkMS) describes the “information about records and the contexts in which they are captured and used.” This is information that the National Archives of Australia recommends be recorded in records management systems and business systems to be consistent with Records Management [31] and Metadata for Records [32], [33].

AGRkMS is based on the AGLS standard and sets out the type of recordkeeping metadata [43]. AGRkMS differs from the first standard in that it is based on a multiple-entity model, allowing for the description of five separate entities - Record, Agent, Business, Mandate and Relationship. “These entities recognized in the multiple-entity model represent the major components that are present in everyday organizational business, including recordkeeping.” It defines a basic set of 26 metadata properties and an additional 44 sub-properties that may be used to describe these entities [44], [45]. Figure 8 shows, at a high level, the five entities and how they are related in the AGRkMS schema, and the relationship entity is the key [45].

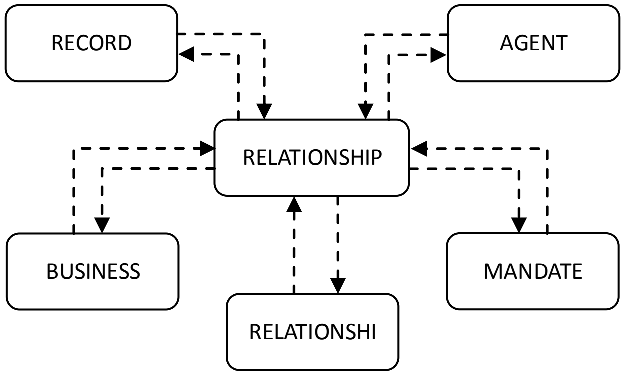


Figure8. High-level Five Entity Model
(From *DRAFT AGRkMS Implementation Guidelines Version 1.0*)

3) EAD

Encoded Archival Description (EAD) is an XML standard used to encode archival finding aids in a networked (online) environment. Finding aids reflect the hierarchical nature of archival collections and that provide a structure for describing the whole of a collection - inventories, indexes, or guides that are created by archival and manuscript repositories to provide information about specific collections. In addition to the content description of digital resources, EAD has the elements for structural description [5], [57], [65].

“EAD Elements section of the tag library contains descriptions of 146 elements and the EAD tag set is used both to describe a collection as a whole, and also to encode a detailed multi-level inventory of the collection. EAD is a metadata schema for archiving digital resources, keeping compatibility with ISAD(G) and one of the guiding principles of EAD is to maintain compatibility with ISAD(G)" [65].

The EAD aims "to create a data standard for describing archives, similar to the MARC standards for describing bibliographic materials. Such a standard enables archives, museums, libraries, and manuscript repositories to list and describe their holdings in a manner that is machine-readable and therefore easy to search, maintain, and exchange" [65].

4) ISAD(G)

The General International Standard Archival Description (ISAD(G)) was originally designed for archived resources in traditional archives and is not specific to digital resources. ISAD(G) is applied to descriptions of all kinds of resources in archives, and it expresses the type of a resource, the source organization of the resource, storage information of the resource and the history of the resource. ISAD(G) also describes information about collection, storage period, usage, copy condition, description element for context of resource, etc. [4].

ISAD(G) provides general guidance for the preparation of archival descriptions, and "defines the concept of hierarchical structure and states which data elements should be included at each level" [63].

ISAD(G) has 26 elements of which six are mandatory and rules. All elements of ISAD(G) "covered by these general rules are available for use, but only a subset

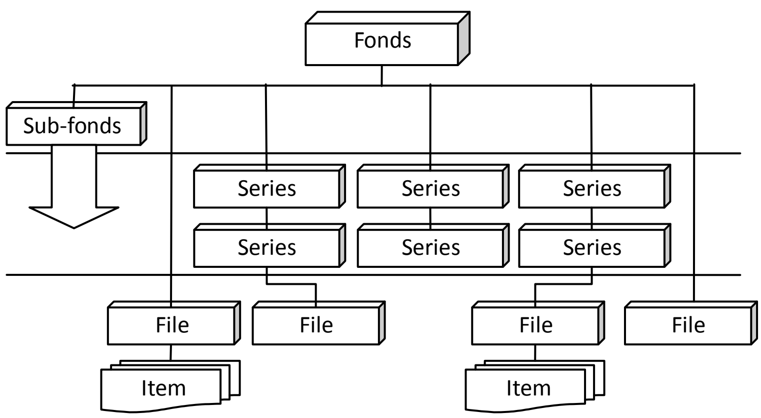


Figure9. Model of the Levels of Arrangement of Fonds
(From *International Council on Archives*)

need to be used in any given description. The rules are organized into seven areas of descriptive information for use at all levels of an archival catalogue (Identity Statement Area, Context Area, Content and Structure Area, Condition of Access and Use Area, Allied Materials Area, Note Area, Description Control Area)".

Figure 9 shows a hierarchical model of the levels of arrangement for the Fonds. There are levels of description, appropriate to each level of arrangement. "There may be a fond - level description, a series-level description, a file-level description and/or an item-level description, such as a sub-fonds or sub-series" [29].

5) PREMIS

The Preservation Metadata and Implementation Standard (PREMIS) is a metadata schema for preservation of digital resources and "is designed to be an effective and inexpensive implementable tool that provides the metadata or information needed to preserve digital information assets for the long term." PREMIS define a data model of instances which are subject to metadata description for preservation and the data dictionary.

The PREMIS data dictionary is the international standard for metadata to support the preservation of digital objects and it defines preservation metadata as the information a repository uses to support the digital preservation process [52].

"The PREMIS data dictionary has 22 metadata semantic units or data elements (19 contain nested sub-elements) divided across entities." Each semantic unit

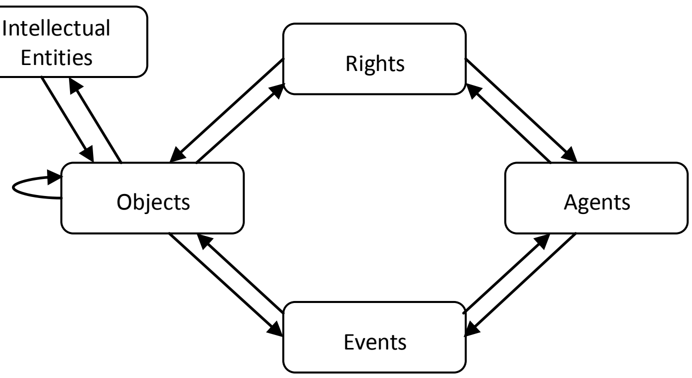


Figure10. The PREMIS Data Model
(From *PREMIS Editorial Committee*)

defined in the Data Dictionary is a property of one of the entities in the data model, and the PREMIS data model consists of five entities – intellectual entity, digital object, agent, rights and events [51], [69]. The PREMIS data model shows in the Figure 10.

6) Preservation metadata elements for the CEDARS project

The CEDARS (CURL Exemplars in Digital ARchiveS) approach adopts the OAIS information model (concepts and terminology) as an underlying framework for their metadata. "The CEDARS metadata also is supplied by the Resource Description element, which for the CEDARS project, is implemented as a Dublin Core record. This record can be supplemented by any other existing metadata records (e.g., MARC) associated with the digital object." [47]

The CEDARS metadata scheme treats Reference Information as metadata for resource discovery and includes descriptive, administrative, technical, and legal information.

"The CEDARS metadata element set is intended to enable the long-term preservation of digital resources. The metadata elements are intended to be applicable to a broad class of digital objects, and divides Provenance Information into three subcategories - History of Origin, Management History, and Rights Management" [47].

Figure 11 shows "the highest level of the Cedars metadata structure. The highest

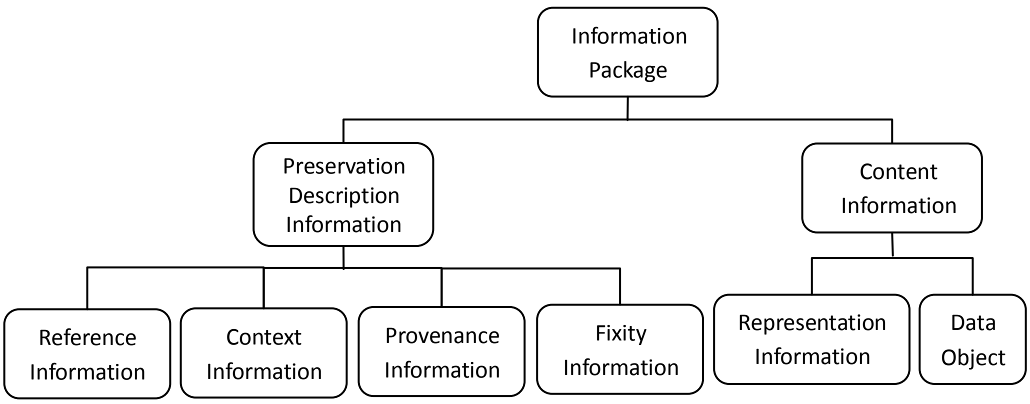


Figure11. The Structure of an Information Package
(From *The Cedars Project Team and UKOLN*)

level object in the OAIS model is an Information Package"[64].

2.6.2 Decision Tree for Long-Term Retention of DPC

As previously mentioned, DPC provides a selection decision tree for long-term preservation. The decision tree is composed of questions and answers in three sections – a question is a node and an answer is an edge coming from the node. An advice may be attached to a node as an answer to the question.

The decision tree does not have attributes as a metadata schema because it is not designed as a metadata standard but it has a set of questions as a tool to help choose a preservation strategy. The questions contain crucial semantic attributes to help choose an appropriate technology or method for preservation at every decision point. Therefore, a semantic attribute in a question can be transformed into a metadata attribute. Thus, the answers to a question are the value of an attribute or a class of values for the attribute.

In this paper, the DPC decision tree, from which we extract metadata attributes, is regarded as a metadata standard like other standards described in the previous section. For the conversion of the decision tree into a metadata schema, we extracted phrases from the questions and organized them into descriptive elements. The method of extracting phrases from the question statements is as follows:

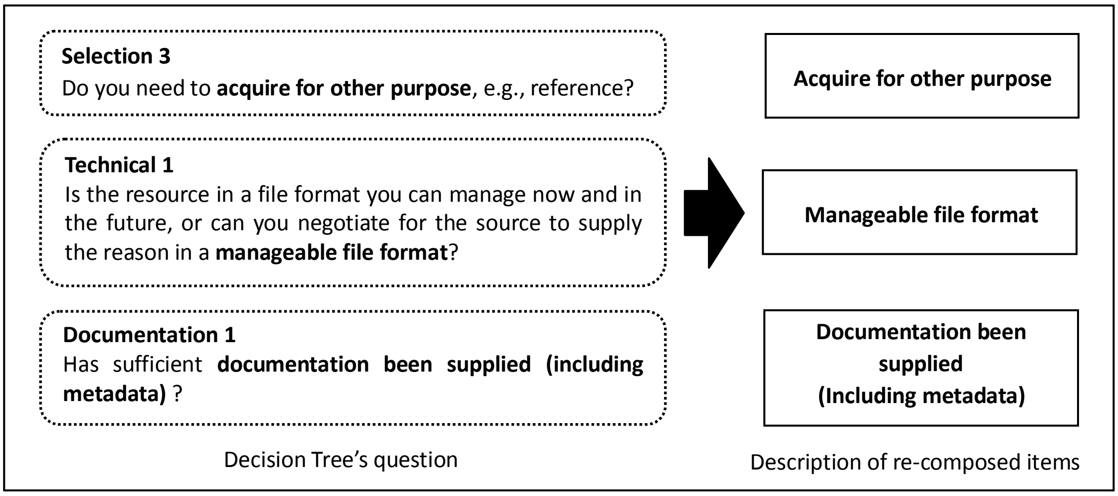


Figure12. Question and Attributes in the DPC Decision Tree

1. Identify the semantic feature in each question that is a node of the decision tree one at a time.
2. Extract a key word or a phrase from the question.
3. Reorganize the extracted key words and phrases from description elements of a metadata schema.

In this way, we got 27 attributes from the set of questions in the decision tree. For example, from a question in selection 3 of the decision tree in Figure 12 'Do you need to acquire for other purpose? ', we extracted the metadata element 'Acquire for other purpose'.

2.7 Models for Metadata Interoperability

A metadata schema for a domain should be 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 component 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 [6].

To define archival metadata schema for the system that created based on a specific purpose, we need to select and combine properly the metadata in accordance to requirement of archival system, i.e., it needs to define the application profiling.

Long-term preservation of digital resource is difficult using single schema in various archival metadata that have each characteristic. In other words, this means that each schema properly selects according to a specific application and metadata interoperability among other system needs. Thus, we performed to metadata mapping and classification from unified viewpoint to select properly the metadata in various metadata, for long-term preservation of archival metadata in our study.

2.7.1 Application Profile

"An Application Profile is defined as a schema which consists of data elements drawn from one or more namespaces, combined together and optimized for a particular local application” [28].

An Application Profile describes a set of metadata elements, policies, guidelines and vocabularies that have been defined for a specific application, particular domain, implementation, or object type. But “an Application Profile is not complete without documentation that defines the policies and best practices appropriate to the application” [68].

“An application profile is an assemblage of metadata elements selected from one or more metadata schemas and combined in a compound schema”[18]. Metadata elements in the application profile are drawn from elsewhere, from distinct namespace schemas and cannot create new elements not defined in existing namespaces [28], [46].

“The purpose of an application profile is to adapt or combine existing schemas into a package that is tailored to the functional requirements of a particular application, while retaining interoperability with the original base schemas”[18]. For example, The Dublin Core Metadata Initiative provided a framework for designing a Dublin Core Application Profile (DCAP). A DCAP is a document (or set of documents) that specifies and describes the metadata used in a particular application, and is designed to promote interoperability within the constraints of the Dublin Core model [15].

Figure 13 shows Singapore Framework for DCAP. “The Singapore Framework for Dublin Core Application Profiles is a framework for designing metadata applications for maximum interoperability and for documenting such applications for maximum reusability.”

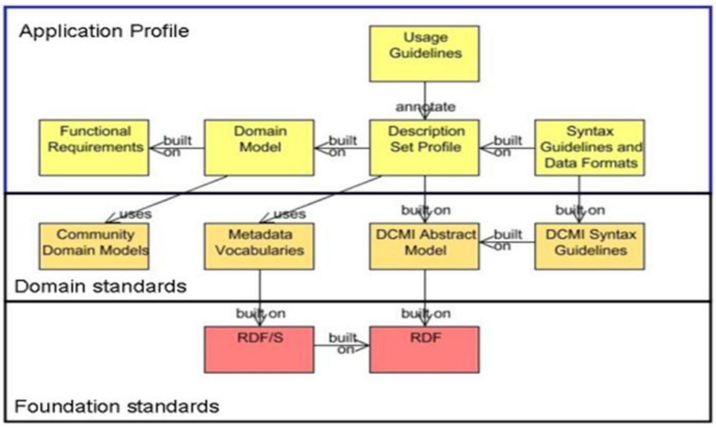


Figure13. Singapore Framework of Dublin Core
(From *Dublin Core Metadata Initiative*)

The framework defines a set of descriptive components that are necessary or useful for documenting an Application Profile, and forms a basis for reviewing Application Profiles for documentary completeness and for conformance with Web-architectural principles [46]. A fundamental issue of the DCMI Application Profiles from the standpoint of this study is that it does not covers interoperability issues along the records lifecycle or archival process.

2.7.2 Vocabulary Mapping Framework

In the metadata community, Vocabulary mapping is a crucial technology in the Semantic Web environment. "The Vocabulary Mapping Framework (VMF) is to provide an important technology for mapping the vocabularies of metadata standards" [24].

The VMF Project is "to create an extensive and authoritative mapping of vocabularies from major content metadata standards, creating downloadable tool to support interoperability across communities." The project is intended to be an expansion of the RDA/ONIX framework for resource categorization [24].

The VMF was developed to improve metadata interoperability based on Semantic Web technology. The VMF provides mapping among some major standards, e.g., Dublin Core, RDA, and LOM, and "the scope of VMF is not limited to these schemes and standards, but these are the initial focus, and many of them have representatives in the VMF project" [35]. The VMF aims to provide to automatically compute the best fit mapping between terms in controlled vocabularies in different metadata schemas and message (in the standard and, in principle, proprietary) [35], [36].

2.8 Related Works

This section presents several studies related to our study. The paragraphs below show related studies in metadata for archiving and preservation, metadata for semantic mapping, metadata interoperability, and records reference model, and add to related standards

2.8.1 Related Researches

1) Metadata for Digital Preservation: A Review of Recent Developments

Michael Day (2001) describes recent developments relating to digital preservation metadata and introduces digital preservation problems, and the importance of metadata for preservation strategies. Specially, the paper explains features of ‘Library-Based Projects’, and projects that relate to preservation, archives and metadata formats for recordkeeping. It also describes the taxonomy of the Information object class defined by ‘ The OAIS Reference Model ’ and some developments in the records domain and archives [16].

We referenced various definitions, descriptions, projects and metadata for recordkeeping from the paper. To review the digital preservation and research on the importance of metadata for preservation can help us make it clear for our study background.

2) Create Once, Use Many Times: The Clever Use of Recordkeeping Metadata for Multiple Archival Purposes

Joanne Evans et al (2005) analyses and explores the development of metadata for multiple archival purposes and relevance to future archival systems using the Clever Recordkeeping Metadata Project (CRKMP).

CRKMP examines the subject to create and share metadata automatically between business systems, record keeping systems, and archival systems. The paper offers a good example of metadata use in the whole records lifecycle. This project explains the interoperability, and the theory of the Records Continuum as a conceptual framework [21].

The theory of the Record continuum is used as a conceptual explanation. And recordkeeping metadata, ISAD(G), EAD and Australian Recordkeeping Metadata Schema etc are also refers to. The relation of the records continuum and metadata for recordkeeping and archives is not mentioned in the paper. Through this paper, we refer to the role, definitions, description of recordkeeping system or record management system, and interoperability. We learned the importance of recordkeeping system for integrated systems, and metadata interoperability through the CRKMP.

3) Practical Issues in Applying Metadata Schemas and Controlled Vocabularies to Cultural Heritage Information

Murtha Baca (2003) focuses on the selection of appropriate metadata schemas for Cultural Heritage Information. It describes the metadata mapping and crosswalks among various element sets such as CDMA, EAD, MARC, and VRA Core. And the paper focuses on the combination of controlled vocabularies and classification systems [3].

Our study used the definition of metadata mapping and crosswalks from the paper. We referenced the description about 'Selection of metadata schemas', 'Metadata mapping and crosswalk', through sample mappings of each metadata schema for museum, bibliographic, archival and Web resources.

4) Metadata Elements for Object Description and Representation: A Case Report form a Digitized Historical Fashion Collection Project

Marcia Lei Zeng (1999) discusses the application of existing metadata formats to a historical fashion collection and develops a catalog for digitized historical fashion collection objects. Three schemes – AACR, Dublin Core, and Visual Resources Associations (VRA) core – were used in this study. The paper describes how to choose, compare and use the different elements of metadata schema for the creation of catalog [70].

Metadata interoperability is an important aspect in our research. So, we referenced the explanations and concepts about metadata interoperability, and examined metadata mapping methods in this paper. The significant difference is that our study is based on the resource lifecycle which is an essential aspect of metadata for archive and preservation.

5) A Methodology for Sharing Archival Descriptive Metadata in a Distributed Environment

Ferro and Silvello (2008) discuss how to exploit widely accepted solutions for interoperability. It shows a methodology for creating sharable archival description metadata which exploits the synergy between the OAI-PMH protocol and the DC metadata format. Also, the paper presents a methodology for mapping EAD metadata into DC metadata records without losing information [22].

Definitions of archives and archival description, descriptions about EAD, OAI-PH and DC etc, we referenced these descriptions and the proposed methodology for our mapping.

6) Metadata Interoperability and Standardization : A Study of Methodology Part 1

Chan and Zeng (2006) studies interoperability problems with multiple metadata schemas, such as having the same subject domain and resources of the same type. It then explains three levels – Schema level, Record level, Repository level - from the same interoperability viewpoint. The six methods - derivation, application profiles, crosswalks, switching-across, framework and registry - are explained to show metadata interoperability with examples [12].

Metadata interoperability, Application profiles etc are very important aspect for the mapping and classification in our study. We referenced definition and description about the metadata interoperability, application profiles, crosswalks and metadata interoperability projects of different levels in the paper.

7) The Semantic Mapping of Archival Metadata to the CIDOC CRM Ontology

Bountouri and Gergatsoulis (2011) describes the semantics mapping of EAD to the CIDOC Conceptual Reference Model ontology and also defines this mapping. The research presents the relationship between the semantic hierarchies and the mapping of EAD to three hierarchies (Hierarchy of Linguistic Objects, Hierarchy of Physical Objects, and Hierarchy of Information Objects). Also, it expresses the mapping using a tree-based hierarchical structure [9].

Although, this is not related to our research directly, it helps us learn and be able to integrate various viewpoints and methods of mapping.

8) Semantic Interoperability across Digital Image Collections: A Pilot Study on Metadata Mapping

Park (2005) explains the issues of semantic interoperability of concept representations across digital collections and presents a semantic mapping between cataloger-defined names and DC metadata elements. The comparison and analysis was conducted using 20 digital image metadata templates and 659 metadata item records in a pilot study. They were mapped using CONTENTdm software and represented the usage of DC metadata elements by three digital image collections and figures [48].

Our study performed the classification using semantic mapping. We referred to the 'semantic mapping and, the mapping between cataloger-defined names and DC metadata elements' in the paper.

9) A Survey of Techniques for Achieving Metadata Interoperability

The survey by Haslhofer and Klas (2010) describes the metadata used in current information systems and its concepts. And then, metadata interoperability and its problems are explained. Especially, the metadata is divided into four blocks using four viewpoints - metadata, model, meta model, meta-meta model [27].

According to each of these four blocks, various metadata standards and metadata mappings and their techniques are explains in a study of metadata interoperability from different viewpoints. The mappings that we have created among the metadata standards improve interoperability of the metadata standards. This survey paper gives hints to compare and mapping between metadata schemas performed in the study described in the paper.

10) Interdisciplinary Contents Management Using 5W1H Interface for Metadata

Keiko Shimazu et al (2006) studies a metadata exchange interface for interdisciplinary content-sharing. The paper shows the interface module which converts tag-labels using 5W1H categories. In this paper, the interface for the metadata abstraction module for contents-circulation across various disciplines was designed using the concept of 5W1H, a representative result of communication study in the field of sociology. 5W1H, which stands for each initial letter, was proposed as the standard solution [56].

Our study uses the Task model and 5W1H categories to identify the contexts of the resources which are the objectives of metadata descriptions. This is a unique feature of this study in contrast with those works surveyed in survey papers and those listed in the paragraphs above. Especially, we examined the usage of 5W1H, and the metadata abstraction module using 5W1H - the metadata labels (of Dublin Core) to 5W1H, the labels of noun types to 5W1H.

11) A Metadata Lifecycle Model for Digital Libraries: Methodology and Application for an Evidence-based Approach to Library Research.

Chen et al (2003) describes and proposes the Metadata Lifecycle Model (MLM). The paper introduces MLM as a methodology of whole process of metadata provision for digital libraries. The MLM involves a ten-step process by which digital library projects can design and implement metadata provision. The purpose of the model is to achieve a consistent method for developing metadata for digital library projects, and to conduct a content-based analysis for digital collections [13].

In our study, we proposed and used the records lifecycle model and the Task model to carry out a feature analysis of metadata elements. Through the metadata lifecycle model that is provided in this paper, we discovered the various views of lifecycle models and we referred to them. We also learned about the metadata analysis which uses the MLM.

2.8.2 Related Standards

In addition to the standards mentioned in section 2.6, the following standards are often used for archiving and preservation. They are not used for the comparison in this research as METS is a container oriented standard and MoReq2 is a comprehensive model for records management.

1) METS

The Metadata Encoding and Transmission Standard (METS) schema is “a standard for encoding descriptive, administrative, and structural metadata regarding objects within a digital library, expressed using the XML schema language of the World Wide Web Consortium" [66].

The METS is based on experience by EAD and, expresses the structure and the contents of a digital resource. It can also be used as an information package of the Open Archival System (OAIS) which determines the reference system for preservation of a digital resource [4]. The METS provides "a framework for incorporating various components from various sources under one structure and also makes it possible to glue the pieces together in a record. It thus provides a framework for combining several internal metadata structures with external schemas. It is a standard that provides a method to encapsulate all the information about an object – whether digital or not" [62].

2) MoReq2

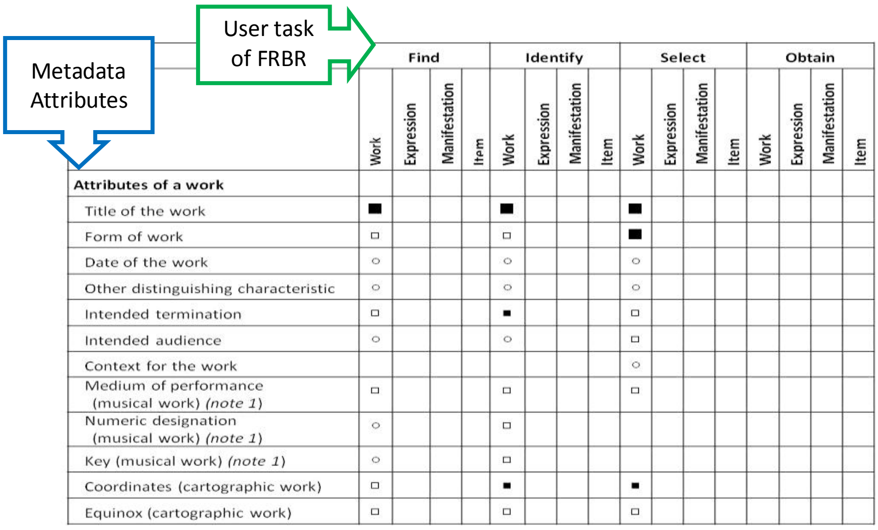
Model Requirements for the Management of Electronic Records 2 (MoReq2) builds on the earlier MoReq, published in 2001, by providing an evolutionary development that incorporates technological and other developments. The metadata model, MoReq2 is "intended for use throughout the European Union, though in practice it can be applied elsewhere" [38].

MoReq2 is an important standard for the management of electronic records. MoReq2 describes the capabilities of an electronic system that manage records, and is the specification that extends beyond pure records management into electronic document and records management (EDRM) and the management of other forms of content [61], [23]. MoReq2 consists of a formal specification of requirements for software systems that are capable of generic electronic records management system or services, accompanied by testing documentation and related information [55], [20]. The MoReq2 specification focuses mainly on the functional requirements for the management of electronic records by an Electronic Records Management System (ERMS). The MoReq2 metadata model is intended to be consistent, to the extent possible, with the following international standards, and is described in terms of a minimum set of metadata elements. These elements are those that the ERMS must be able to export, import, and process [20].

3 From Resource-centric to Task-centric View of Metadata Schema

A resource may be affected by a task performed in a lifecycle stage - for example, in the *Appraisal & Disposition* stage, a resource disposed may be revised in the appraisal process in accordance with the preservation policy of the given archive. Metadata should be able to record the change of the resource as the lifecycle stage proceeds. Thus, the metadata elements are assigned values or updated in the lifecycle stages. Most metadata standards are designed in accordance with the lifecycle stages where the metadata standards are applied. However, most metadata standards make no mention about the resource tasks. FRBR (Functional Requirements for Bibliographic Records) which is defined as a model for bibliographic description includes some generic tasks and metadata elements (i.e., metadata attributes) used in those tasks, e.g. title-of-work is required to find a work.

FRBR shows the four generic tasks - find, identify, select, and obtain – to explain the relationship between the attributes and tasks [30]. Figure 14 show the mapping of four generic tasks and *Work* (one of four elementary attributes) in the entity-relationship model. “Each task is in turn broken out into four sub-tasks defined in relation to the



	Find				Identify				Select				Obtain			
	Work	Expression	Manifestation	Item	Work	Expression	Manifestation	Item	Work	Expression	Manifestation	Item	Work	Expression	Manifestation	Item
Attributes of a work																
Title of the work	■				■				■							
Form of work	□				□				■							
Date of the work	○				○				○							
Other distinguishing characteristic	○				○				○							
Intended termination	□				▪				□							
Intended audience	○				○				□							
Context for the work									○							
Medium of performance (musical work) *(note 1)*	□				□				□							
Numeric designation (musical work) *(note 1)*	○				□											
Key (musical work) *(note 1)*	○				□											
Coordinates (cartographic work)	□				▪				▪							
Equinox (cartographic work)	□				□				□							

Figure14. The Mapping of User Task and Attributes in the FRBR
(From *IFLA Study Group on the Functional Requirements for Bibliographic Records*)

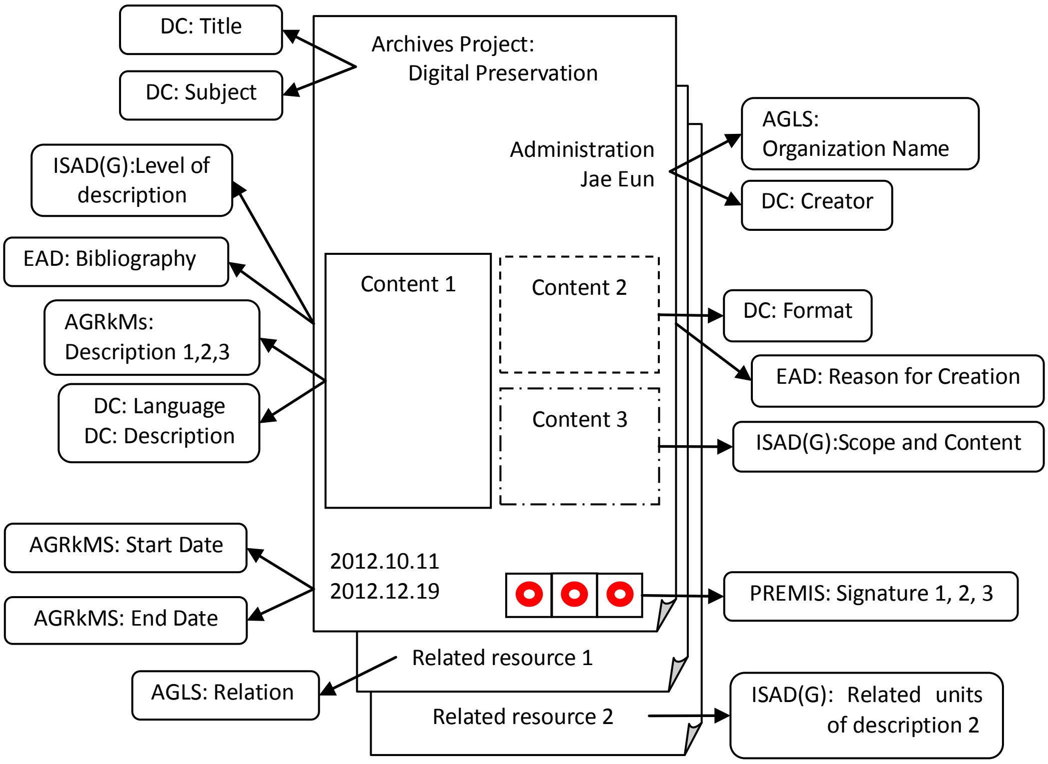
entity on which the task is focused (i.e., find work, find expression, find manifestation, find item, etc.). The symbols (■=High value, □=Medium value, ○=Low value) used in the tables indicate the relative value of each attribute or relationship in supporting a specific user task focused on a particular entity" [30].

FRBR User Tasks are included in a stage between *resource creation* and *use* in the lifecycle because of the nature of bibliographic description. User Task shows the relationship of metadata attributes and a task using the importance of metadata value, by applying to a task the metadata that describes a resource. FRBR User Task shows the metadata attribute is related to resource task. In addition, this means that we can show metadata attributes from task-oriented viewpoint.

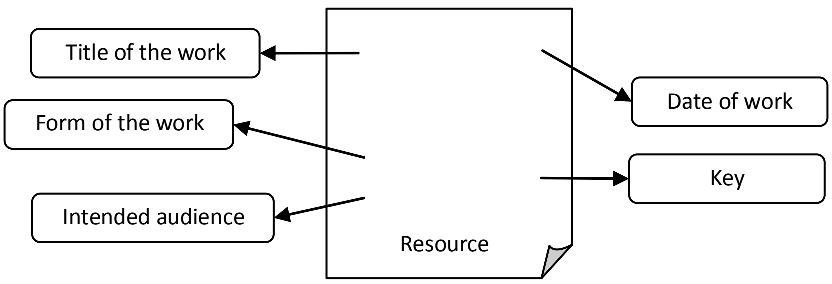
Figure 15a shows the metadata elements (title, creator, language, date, signature, relation and so on). This means that metadata elements are designed from a resource-centric view. As mentioned above, we use the relation between metadata and tasks to identify features of metadata schemas in this study. So, we examine metadata in each stage of the lifecycle from a task-centric view. Figure 15b shows the FRBR User Tasks in the lifecycle. Figure 15c shows some metadata elements and their related stages in the lifecycle.

For example, *ISAD(G): level of description* is an element that describes the level of a resource for archiving. If this element applied to a stage in the lifecycle, it should be included and used in the storing or archiving stage. For another example, *reason for creation* of EAD expresses the reason why the resource is created. This element applies to the creation stage in the lifecycle. In other words, a resource is examined in every stage of task of lifecycle for the tasks in the stage.

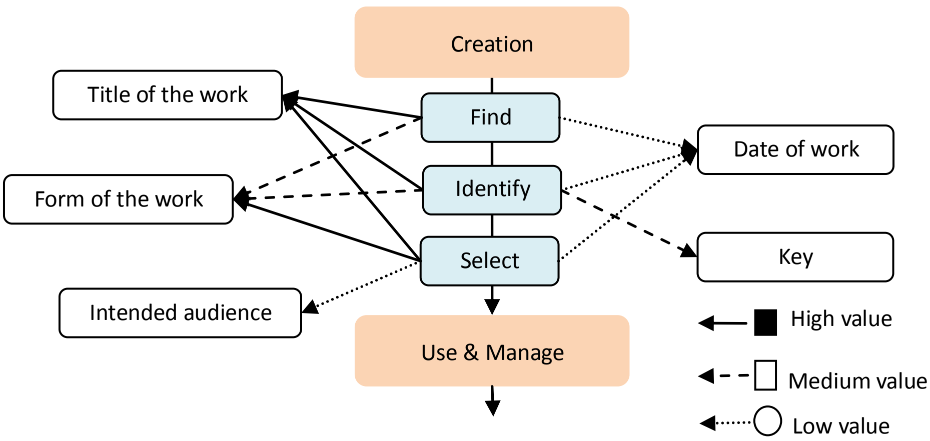
In general, a metadata standard is defined from a resource-oriented viewpoint in accordance with the purpose of the standard. On the other hand, each metadata element is used in a task at a lifecycle stage. The task-attribute relationship given in FRBR is a well-known example of the relationship. The task-attribute relationship is useful to clarify the feature of a metadata standard from the viewpoint of tasks performed in the resource lifecycle stages. Task oriented view of metadata standards is advantageous to define mappings between metadata standards along with the lifecycle stages.



a. Resource-centric View of Metadata Schema

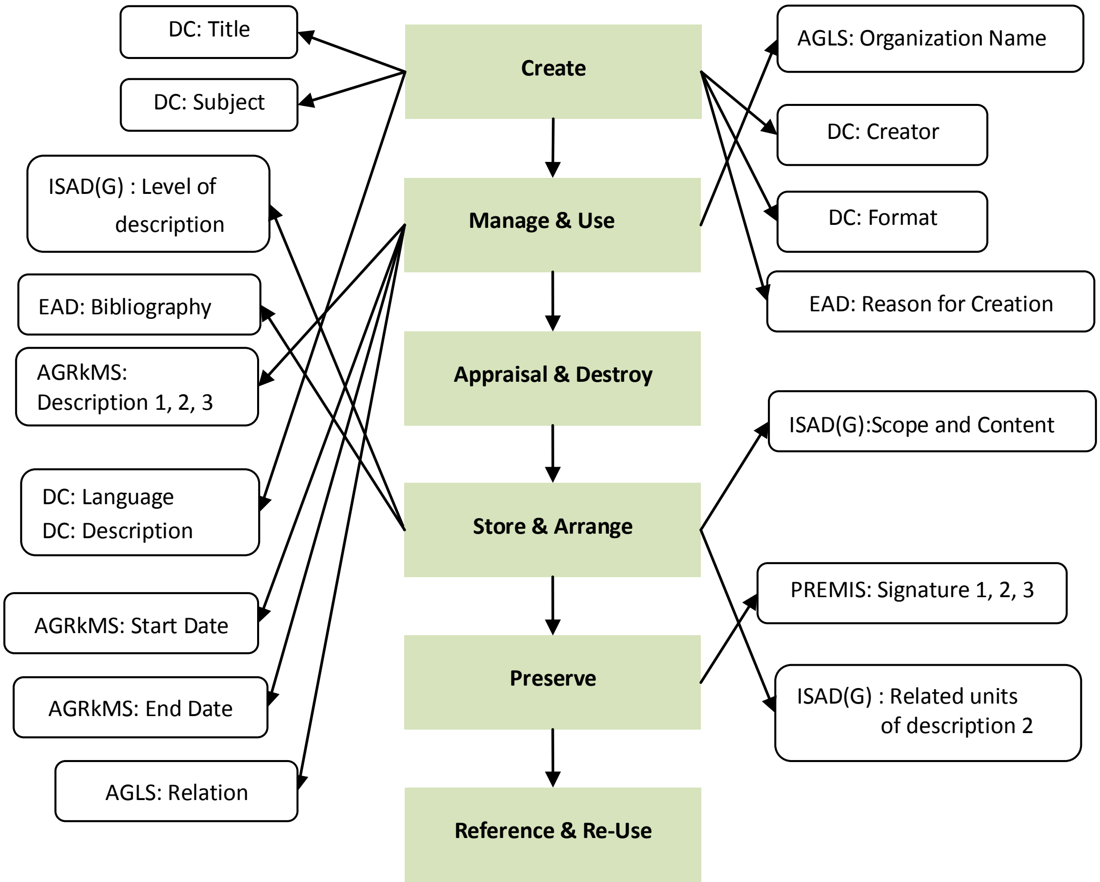


i. Resource-centric View of Metadata Attributes in the FRBR



ii. Task-centric View of Metadata Attributes in the FRBR

b. Resource and Task-centric View of Metadata Schema in the FRBR



c. Task-centric View of Metadata Schema

Figure15. From Resource-centric to Task- centric View of Metadata Schema

4 Feature Analysis of Metadata Schemas based on Lifecycle

In our first study, we showed a simple feature analysis based on the type of description elements and the relationships between the description elements and the lifecycle stages [5]. This section briefly shows the feature analysis of archival metadata schemas from the Viewpoint of Records Lifecycle.

4.1 Viewing Differences from Descriptive Elements

As the first research, the author performed element mapping and analysis of archival metadata from a viewpoint of lifecycle, in order to analyze the feature of metadata standards.

Mapping of metadata standards into the records lifecycle is examined to explicitly extract and compare the features of metadata schemas used in digital archives and preservation. For the mapping, it is necessary to extract descriptive elements from a metadata schema, and then to examine in which stage of the lifecycle the value of each element is determined.

During a workflow that takes place according to a metadata standard, a metadata element is created at some point and used in the whole records lifecycle. Therefore, the

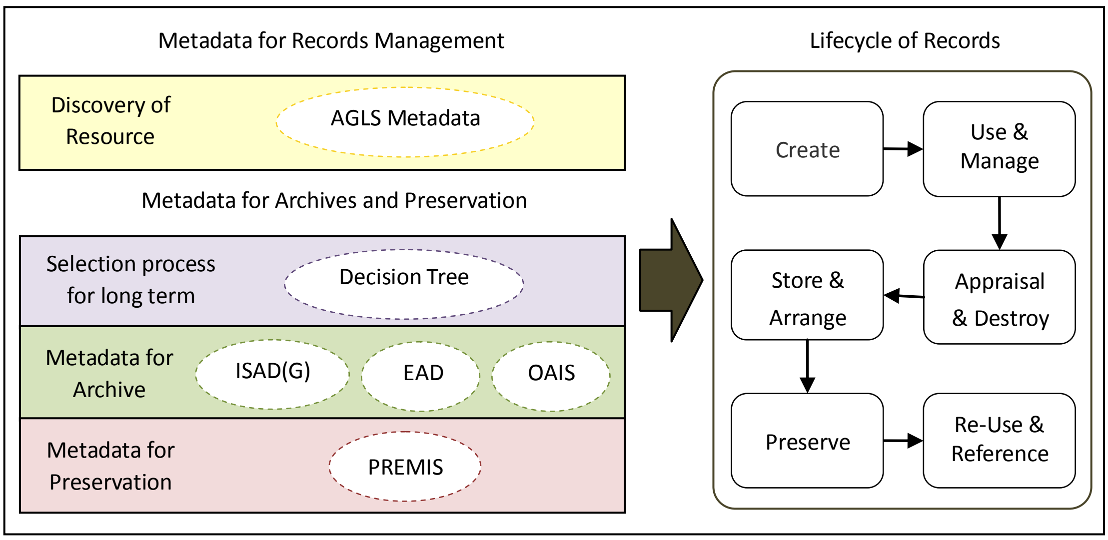


Figure16. Lifecycle and Metadata Standards for Archiving and Preservation

author used the viewpoint of ‘Creation and revision (updater) of metadata’. We define a Creation Stage and Update Stage of a metadata element in the records lifecycle as the stage where the metadata element is given an initial value and revised, respectively. The creation and update stages are called a primary stage of the metadata element.

A metadata element may have one or more primary stages. For example, as the value of creator element of a resource is determined when the resource is created, the primary stage of the creator element is the first stage of the lifecycle, i.e. “create” in Figure 16. Even if the creator element is very frequently used in the later stages, the primary stage is “create”. If the value is revised or updated in a later stage in the lifecycle, the stage is also a primary stage of the element.

4.2 Analysis Method

In order to analyze descriptive elements into a corresponding stage of the lifecycle, we carried out classification and mapping, using the following method.

(1) Analyze the feature of metadata standard.

For example, *preservationLevelDateAssigned* of PREMIS

Before analyzing descriptive element, PREMIS is metadata standard for preservation of digital objects and is use in the preservation stage of records lifecycle basically.

(2) Find and classify a corresponding keyword or a related meaning from the value of descriptive element.

For example, *preservationLevelDateAssigned* of PREMIS defined “The date, or date and time, when a particular *preservationLevelValue* was assigned to the object”. This element means not only the period which determines a preservation level, but also the period which changes the preservation demand and policy etc of repository. Thus, we decided this element as a preservation stage of lifecycle and classified it.

Mapping metadata standards into the records lifecycle is done in two steps:

Step1. Extract every metadata element from each metadata schema standard one by one, and determine the primary stages in the records lifecycle for the element.

Step2. For each metadata schema, determine its primary stage set in the lifecycle where the primary stage set means a set of stages in which the majority of the metadata elements are given their values or revised. This step requires over viewing of the metadata element sets across the stages of the lifecycle.

Following the steps shown above, we examined all of the six schemas (AGLS, ISAD(G), EAD, OAIS, PREMIS and the DPC's Decision Trees). The full result is shown Appendix 1, 2, 3 and the following sections explain the Step 1 and 2 in detail.

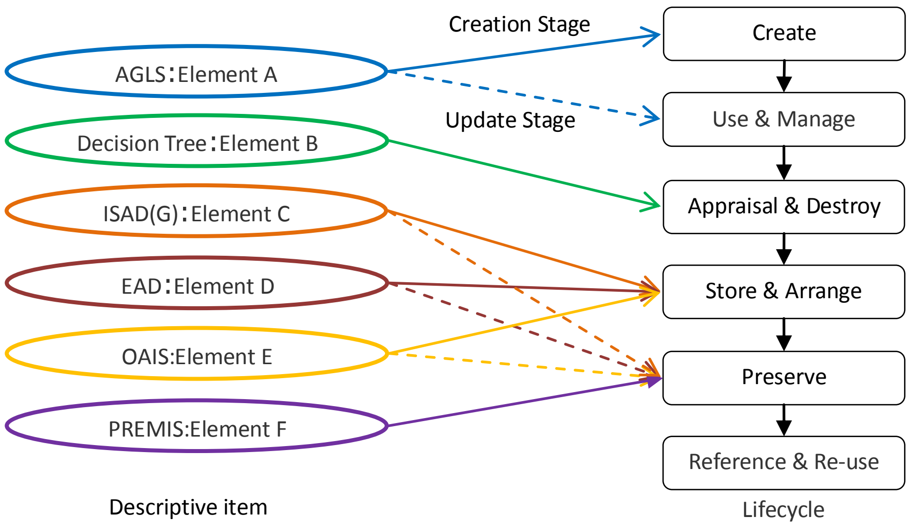


Figure17. Classification Criteria of Metadata Standards into the Records Lifecycle using the Step 1 and 2

4.3 Mapping to Determine the Primary Stages in the Lifecycle

1) Step 1: Extract Descriptive Elements of Metadata Standards for Records Management, Archives

This section shows analysis of a metadata element extracted from each metadata standard. Because every schema has many elements, this section shows the analysis using examples. Each element shown in the paragraphs below is given its primary stages in two aspects – Creation and Update. Creation shows a stage where initial value of the element is given and Update shows a stage(s) where the element value is changed or updated.

(1) AGLS Metadata

We select an element named *Availability* as an example. *Availability* is primarily used for non-digital resources, provided information on how the user may acquire physical accesses to a resource. Because this element explains the availability of resources in the real usage environment, we classify the stage of this element as *Use & Manage*. The value of the element is updated in *Appraisal & Destroy* and *Reference & Re-Use*. Table 1 shows the summary of the primary stages for *Availability*.

Table1. An Example of AGLS Metadata

Element of AGLS Metadata : Availability	
Point of view	Lifecycle Stages
Creation	Use & Manage
Update	Appraisal & Destroy, Reference & Re-Use

(2) Decision Tree

Acquire for other purpose is used as an example element of the DPC Decision Tree. As mentioned before, the descriptive element of the Decision Tree is re-composed by re-phrasing a question at a node. *Acquire for other purpose* explains appraisal for other purpose in resource selection. This element was classified in the appraisal stage, i.e., *Appraisal & Destroy*. As the Decision Tree is not a metadata scheme, Decision Tree does not include a revision of the element value.

Table2. An Example of Decision Tree

Element of Decision Tree : Acquire for other purpose	
Point of view	Lifecycle
Creation	Appraisal & Destroy
Update	Not Applicable

(3) EAD

Archdesc gives a description about a resource - contents, contexts, scopes and so forth. The element value is determined in *Create*. Then, it is to be updated in *Appraisal & Destroy*, *Store & Arrange* and *Preserve*. This is because each time a resource is processed in an archival system the description of the resource may be subject to change.

Table3. An Example of EAD

Element of EAD : archdesc	
Point of view	Lifecycle
Creation	Create
Update	Appraisal & Destroy, Store & Arrange, Preserve

(4) ISAD(G)

Level of Description is an element that expresses units of resource, which is divided into Fond, File, Item and so on. A unit of the resource may be changed if related resource(s) are added or removed.

A value for *Level of Description* is set in the *Create* stage of the Lifecycle, and updated in the step of *Use & Management* that confirms the related or subordinate resources, while using the resource. The value is updated in the steps in archival phases -*Appraisal & Destroy*, *Store & Arrange*, *Preserve* and *Reference* - where archives may change the values in accordance with their policy and changes in the time line.

Table4. An Example of ISAD(G)

Element of ISAD(G) : Level of Description	
Point of view	Lifecycle
Creation	Create
Update	Use & Management, Store & Arrange, Appraisal & Destroy, Preserve, Reference & Re-use

(5) OAIS

Change history before archiving describes the change history of a resource before it is deposited in an archive. The value of this element should be set in *Store & Arrange* and may be updated in *Preserve*.

Table5. An Example of OAIS

Element of OAIS : change history before archiving	
Point of view	Lifecycle
Creation	Store & Arrange
Update	Preserve

(6) PREMIS

Creating Application describes the applications used when a digital object was created. For this reason, the value of this element is determined in *Create*, and then, updated in *Store & Arrange* and *Preserve* where the digital object may be migrated to a new environment.

Table6. An Example of PREMIS

Element of PREMIS : creating Application	
Point of view	Lifecycle
Creation	Create
Update	Store & Arrange, Preserve

We took out every descriptive element from the metadata schemas, and mapped them to the records lifecycle stages in order to determine the primary stages of each element. Based on this investigation, we analyzed the relationship between each metadata standard and the lifecycle stages. Appendix 1, 2 and 3 shows the relationships between elements and the primary stages of the schemas. In these three tables, all elements of the metadata are shown where Roman and Italic fonts mean Creation and Update, respectively.

Figure 18 shows a summary of the analysis presented above. We have applied the analysis method above to all elements of the six schemas and summarized the result in the schema, mentioned in the next section.

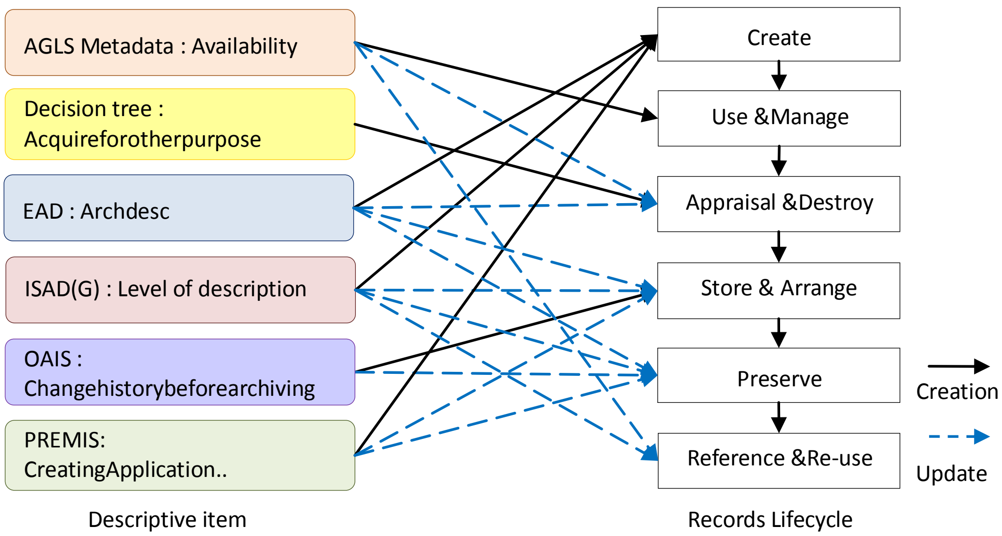


Figure18. An Example of using the Records Lifecycle for the Description Elements

2) Step 2: Determine Primary Stages for Metadata Standards

Based on mapping performed in the step 1 and feature of metadata standard, this section shows lifecycle stage that mainly expresses all metadata elements and each metadata standard.

(1) AGLS Metadata

AGLS Metadata is composed of a description about resources according to their contents for searching. In the lifecycle, we found that AGLS Metadata mainly expresses *Create*, *Use & Manage*, and *Reference & Re-Use*. This is a very natural result because the first two stages are not necessarily related to long-term archiving but to general resource discovery and management, and the last stage is for users who want to find and use resources in the archives. Also, archival metadata schemas have a small set of general descriptive metadata like the ones on AGLS.

(2) Decision Tree

DPC's decision tree was made as the selection policy of a resource. The element set created from the DPC's decision tree (chapter 2.3) is composed of descriptive elements about the evaluation of the resources. Therefore, these elements are used only in *Appraisal & Destroy* and *Store & Arrange*. This crispness is the feature of the decision tree compared with other metadata schema standards.

(3) EAD

EAD mainly has descriptive elements that express the appraisal of the resources, history, origin of resources, and relative information. As elements of EAD are mainly for evaluation and basic description for archives, many elements for *Appraisal & Destroy* and *Store & Arrange* and some elements for *Preservation* are included.

(4) ISAD(G)

ISAD(G) is similar to EAD, but it does not have so many elements for *Preservation* as EAD has. ISAD(G) has elements that express bibliographic information and administrative information for archives such as management, use of resources, history information, and so forth. Thus, ISAD(G) is linked to *Appraisal & Destroy*, and particularly to *Store & Arrange*. On the other hand, the first two stages of the lifecycle are also connected.

(5) OAIS

OAIS has elements to express collection and history of digital objects. On the other hand, it has many elements to express technological and structural contents. OAIS has many elements for re-using resources. This is because dissemination of archived resources is a part of the OAIS reference model. Thus, OAIS covers *Appraisal & Destroy*, *Store & Arrange*, *Preservation*, and *Reference & Re-Use*.

(6) PREMIS

PREMIS have many elements that express technological features for preservation of digital resources. Significant difference from other metadata schemas that are connected to more than one stage in the lifecycle is that PREMIS is concentrated into *Preservation*.

In the first study, we mapped the lifecycle stages to metadata elements extracted from the metadata standards. In this mapping, for every element extracted from metadata standards, we determined the primary stages where the element value is initially given or revised. Table 7 shows the statistics of the mapping.

In Table 7, the numbers show the percentage of elements of each standard whose values are initially given or revised in a corresponding stage of the lifecycle. For example, in the case of EAD, *Appraisal & Destroy*, *Store & Arrange* and *Preserve* stages are the primary stage for 14%, 33% and 20% of the elements, respectively. On the other hand, 24% elements are determined their values in the first two stages. This shows that EAD is oriented to resource organization in the archival storages rather than resource discovery and management in live resource repositories used in the early stages of the lifecycle. AGLS is primarily designed for resource discovery and access, which correspond to the first two stages of the lifecycle. In this study, however, the table shows AGLS is used in the whole lifecycle as a finding aid throughout the records lifecycle.

Table7. Metadata Standards shown by Figures (%)

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

As shown in Table 7, the primary stages are spread over the lifecycle but there is a peak in the *Use & Manage* stage. More importantly, appendix 3 shows that there is a clear split between *Create stage* and *Update stage*. This shows that the values initially given, are used for discovery in the first two stages of the lifecycle and the values may be revised for maintenance at archives. Thus, we can identify the overall features of the metadata standards shown in Figure 18 from the statistics shown in Table 7.

Every metadata schema is related to all stages of the lifecycle except the decision tree. Figure 19 shows the overall relationship between the schemas and the records lifecycle. The figure shows the high-density parts where many elements are connected to a specific stage. For example, AGLS has many elements connected to *Create*, *Use & Manage*, and *Reference & Re-use*. The paragraphs below show the analysis of each standard.

Figure 19 is useful to view the stages where crosswalks between metadata schemas are efficiently performed. This is because it helps us identify the correspondence between elements of similar meanings by showing the correspondence of elements to lifecycle stages. Thus, new viewpoint to enhance interoperability of the archival metadata schemas are given.

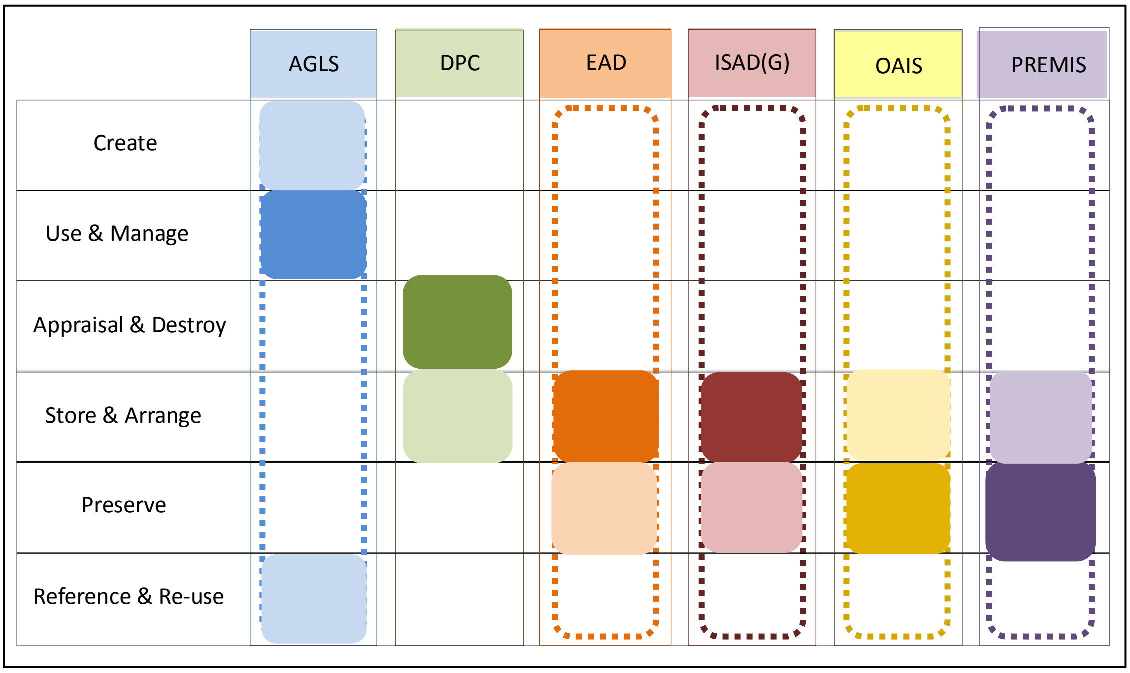


Figure19. Stage of Lifecycle shown by Metadata Description Elements

4.4 Consideration

Metadata standards for archiving and preservation of digital resources are various. However, each metadata standards has its own feature in accordance with its primary application. We have examined the metadata for preservation and archives of digital resources from the viewpoint of mapping between the metadata standards and the records lifecycle. In our research, we first started our study with a simple question “Is it possible to preserve resources long-term only by one metadata schema?” and another question “Is it possible to design a unified framework for metadata standards for archiving and preservation?” As a result the detailed examination of the metadata elements, we clarified the features of the standards from the viewpoint of relationships between the elements and the lifecycle stages.

The unified framework to identify the features of archival metadata standards proposed in the first study is useful to combine different archival metadata schemes in a single system because it is straight forward to find stages where mappings between different standards are heavily required. Thus, this unified framework is advantageous to enhance interoperability between the archival metadata standards.

Mapping between metadata schemas is a crucial issue because we are frequently required to unify metadata databases and because metadata mapping is required in the long-term preservation process. However, on the other hand, we know that metadata schema mapping is an expensive task. Our second research is to define a framework to help systematically map metadata elements for preservation.

5 Facet Analysis of Archival Metadata Schemas for Metadata Interoperability

5.1 Introducing a Task-centric View of Archival Metadata Standards

Mapping between metadata standards is an expensive but often unavoidable task to enable metadata use across organizations. As metadata elements are defined primarily for describing resources, each element expresses an attribute of a resource or a relationship between the resource and other resources. This means that there is no systematic way to use the resource lifecycle information in the mapping in spite of the fact that every metadata standard has lifecycle dependent features as shown in Chapter 4.

From the feature analysis discussed in the previous chapter based on the resource lifecycle, we have learned that we need to use not only the semantic description given in the definition of a metadata element but also the context information of the element which can be obtained from the lifecycle.

In our second study which is presented in this chapter, we introduce a task-centric view of metadata elements in order to create metadata mappings across the lifecycle stages. In the rest of this chapter, we describe a task-based model of the resource lifecycle, which we call the Task Model. Then, we define a task-centric view of metadata elements and we introduce 5W1H categories to characterize metadata elements for a task-oriented semantics analysis of the metadata elements [7].

5.1.1 Task-oriented View of Records Lifecycle – Task Model

The records lifecycle defines stages of records – from creation at offices to preservation in archives. In order to examine in detail the relationship between resource and each stage in records lifecycle, we propose the Task model. The Task model is defined in parallel to the records lifecycle.

The Task Model is a model that is created based on the records lifecycle. The Task model is proposed in this study in order to analyze metadata standards in detail from the

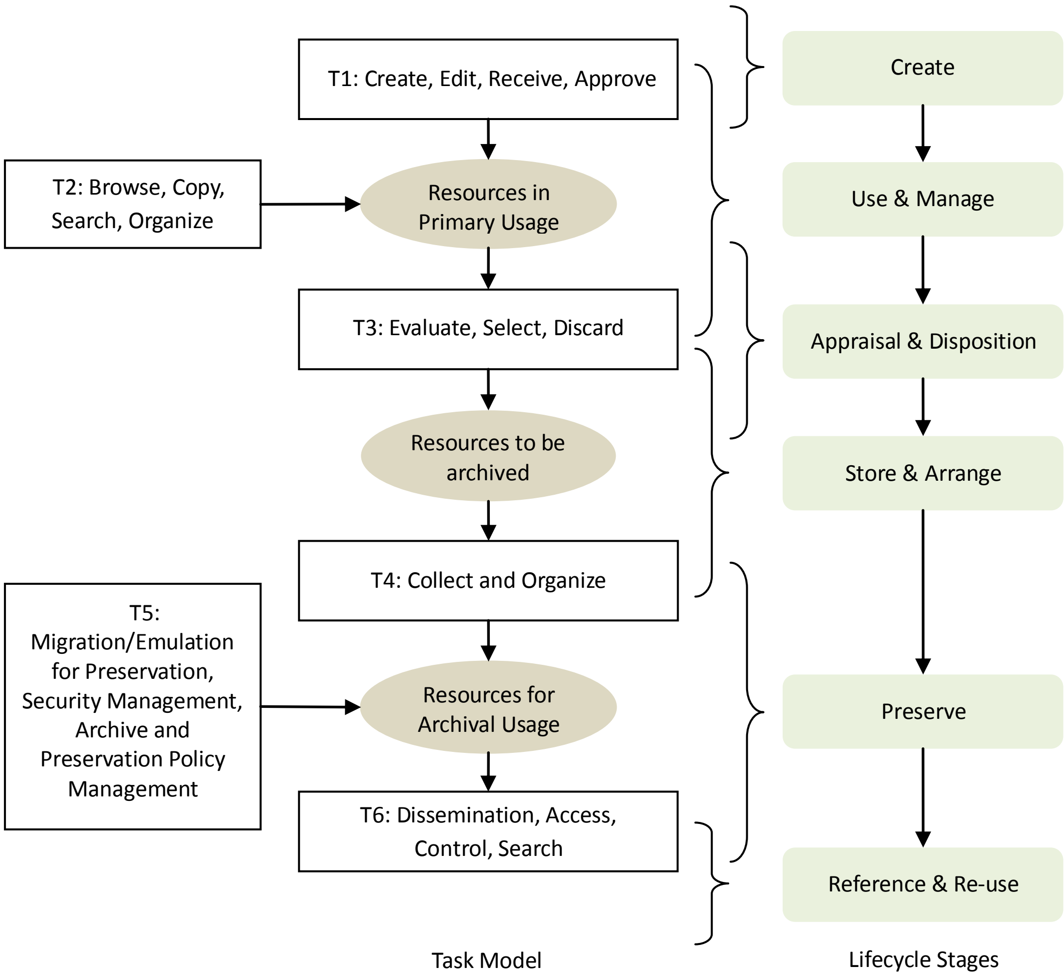


Figure20. Task Model

viewpoint of the tasks performed in each stage of the records lifecycle. In this model, task groups, which are composed of several tasks and linked to resources, are associated with the lifecycle stages. The records lifecycle briefly describes what tasks are performed in each stage but it is not clear how the resource attributes are used in the stages. On the other hand, the Task model is more descriptive than the lifecycle because each task in the groups indicates attributes of resources used in the task. .

As shown in Figure 20, the Task Model defines the tasks performed in each stage of the records lifecycle. The Task Model is composed of six task groups (T1-T6) defined as follows,

Task 1: Creation tasks: Tasks used for initial creation including those for the approval process,

Task 2: Primary Usage tasks: Tasks for primary users to find and browse resources,

Task 3: Appraisal and Retention tasks: Tasks to select and discard resources,

Task 4: Archival Transformation tasks: Conversion and transformation tasks for archival storage,

Task 5: Preservation tasks: Maintenance tasks for archival storage, and

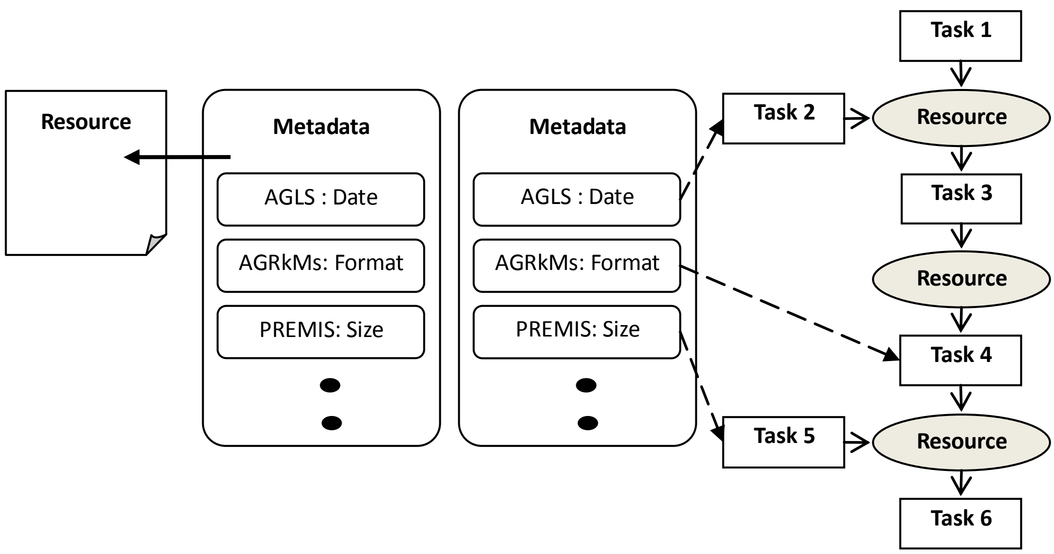
Task 6: Archival Usage tasks: Tasks to find and use archived resources

The lifecycle stages are shown to the right of the Task model in Figure 20. 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.

5.1.2 Task-centric View of Metadata Schemas

As a resource is used in different tasks throughout the whole lifecycle, it is obvious that we need a metadata model to clarify what attributes of a resource should be described in accordance with the task groups. However, in conventional resource-centric metadata models, it is not clear which metadata element is used in a particular task or stage.

Figure 21a and 21b show a resource-centric and a task-centric view of metadata standards. Figure 21a illustrates a metadata element which describes one resource using



a. Resource centric-view of Metadata b. Task centric-view of Metadata

Figure21. Resource and Task-centric View of Metadata

elements adopted from AGLS, ARGkMS and PREMIS. Figure 21b illustrates in which tasks the elements are used. For example, *date* from AGLS, *format* from AGRkMS, and *size* from PREMIS describes one resource. These metadata elements are linked to T2, T4 and T5 respectively in the Task model. Thus, different metadata element which describes a resource could apply and express each task in the Task model.

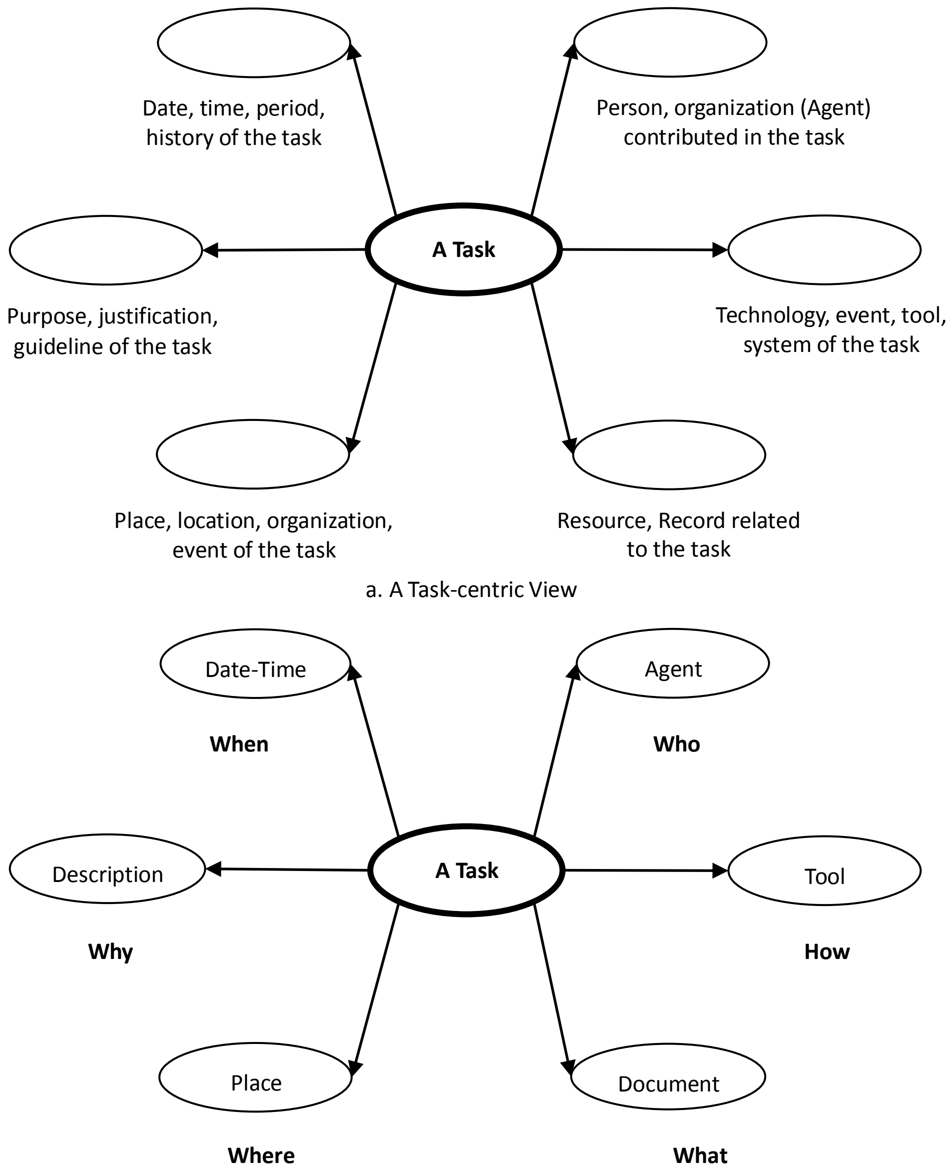
5.1.3 Combination of Task-oriented Model and Metadata Elements

This section describes in detail the metadata standards from the task-centric viewpoint. An execution of a task causes an event on resources. We describe the relationships of the metadata elements and an event (on resources).

Figure 22 shows a task-centric view of metadata. A task-centric view of metadata is to define metadata standards in the context of tasks. In Figure 22a, ‘A Task’ is linked to values of metadata elements, i.e., an entity, such as right, time, purpose, or person. This is the reverse of Figure 22b. As shown in Figure 22a, every single task is associated with those entities shown as a circle. These entities are agents that play some roles in the task, locations or institution where the task is performed, reasons and 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 Figure 22b links from the task are labeled using 5W1H categories, i.e., an input link to an entity is reversed as a metadata element of the entity. In this study, we propose to use 5W1H categories - who, where, when, what, why and how - as generalized categories to express the relationship of a related entity and task, as shown in Figure 22b. Figure 22b is derived from Figure 22a by categorizing the relationship from the task to the values. The paragraphs below show detailed explanation of this categorization.

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 has 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.



b. A Generalized Task-centric View – 5W1H

Figure22. Task-centric View of Metadata and 5W1H

5.1.4 5W1H Categories

5W1H attributes are used to identify categories of metadata elements. A metadata element category represented by a 5W1H attribute is called 5W1H categories in the rest of this paper. The paragraphs below show definitions of the 5W1H categories for this study.

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.

5.1.5 Discussion on Resource-centric and Task-centric Views of Metadata Elements

Tasks are carried out on a resource during the lifecycle, e.g., creation, edition, search, revision, appraisal, disposal, conversion, and so forth. Each of the entities linked from this task is a resource which appears in the lifecycle, i.e., a document, a person, a place, or a description. Entities such as documents and records are the primary objects managed by an archive and a record management system. Other entities are recorded as values of a metadata element as shown in Figure 22a.

Modern metadata standards have their own base data models, e.g., the PREMIS data model consisting of five classes of entities. However, in general, those data models are defined from a resource-centric standpoint but not a task-centric or lifecycle-oriented standpoint. 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 first research

The differences between data models of metadata standards have to be taken into account for mapping metadata elements. The underlying idea of this study is to use the generalized task-centric view of metadata for mapping metadata schemas instead of the conventional data entity-centric view.

5.2 Analysis Criteria of Classification

In the first study, we performed classification and mapping for the relation of metadata elements and the feature analysis of metadata standards within a resource task. This section describes classification and mapping of metadata elements.

5.2.1 Vocabulary of Systematic Classification by 5W1H Categories and Lifecycle Tasks

In the second study, we classify every metadata element - AGLS, AGRkMS, EAD, OAIS, PREMIS and the attribute sets of DPC - using the 5W1H categories and tasks in the lifecycle. We used explanation texts of each metadata element to find keywords. And keywords are used to classify all the metadata elements into the 5W1H categories and the lifecycle tasks. This classification was carried out manually because we had to interpret the meanings and intention of the explanations. We prepared a set of keywords for each task group and 5W1H categories and used the keywords to classify every element into a task and map it to 5W1H categories.

Tables 8 and 9 show the keywords for the 5W1H categories and the Task model, respectively. The keywords are manually extracted as typical words to express a category and a task, respectively. They are used as keywords for classification of metadata elements by tasks and by 5W1H categories. More than one keyword may appear in the definition of a metadata element.

The paragraphs below show the classification guideline,

1. Find keywords in the title, definition and guideline texts of a metadata element,
2. If no keyword is found, find a term (or terms) whose meaning is similar to a keyword,

3. If matching by 1 or 2 does not succeed find a keyword (-s) in a use-case example of the element.

For example, *Date of Publication* from OAIS explains the date of publication of a version of a specific digital object. *Date of Publication* has two keywords in its name, *date* and *publication* which are keywords for *when* and *what*, respectively. Thus, *Date of Publication* from OAIS is categorized both in *when* and *what*.

Table8. Classification Vocabulary with 5W1H Categories

5W1H Categories	Keywords (Example)
Who	Agent, Author, Creator, Institution, Name, Organization, People, Person etc
When	Date(s), Period, Time, Month, Day, Year etc
Where	Agent, Country, Institution, Location, Name, Organization, Place etc
What	Administration, Bibliography, Description, History, Policy, Relationship, Right etc
How	Action, Event, File format, Hardware/Software, Metadata scheme, Technique, Tool, Transference etc
Why	Purpose, Reason etc

Table9. Classification Vocabulary with Task Model

Task Group	Keywords (Example)
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

5.2.2 Classification Procedure

The classification workflow has two steps, classification by 5W1H categories (step 1) and classification by 5W1H categories in the lifecycle tasks (step 2). The following paragraphs describe the classification steps in detail. Figure 23 and 24 illustrates the steps.

Step 1 Metadata Mapping by 5W1H categories

1-1 Classification of descriptive elements: For every element of each metadata standard, examine whether the definition text of the element includes one or more keywords listed in Table 8 and, if found, classify the element to the corresponding category (-ies).

1-2 A mapping among metadata standards: In each 5W1H category, compare elements among the standards and create mappings. If a mapping table for any of the standards exists, it is also used to determine the mapping.

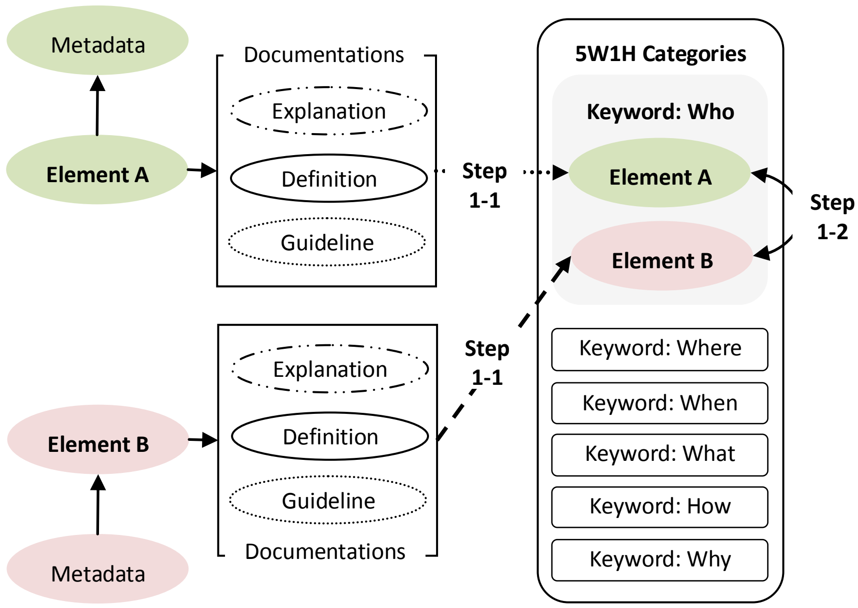


Figure23. Mapping of Metadata Elements in 5W1H categories (Step 1)

Step 2 Metadata Mapping in Lifecycle Tasks

2-1 Classification of descriptive elements by tasks: For every element of each metadata standard, examine whether the definition text of the element includes one or more keywords listed in Table 9 to classify the element to the corresponding task(s).

2-2 Classification of descriptive elements by 5W1H categories: For every element classified to a task, apply Step 1-1 to classify the element by 5W1H in each task.

2-3 A mapping among metadata standards: In each 5W1H category of each task, create mappings.

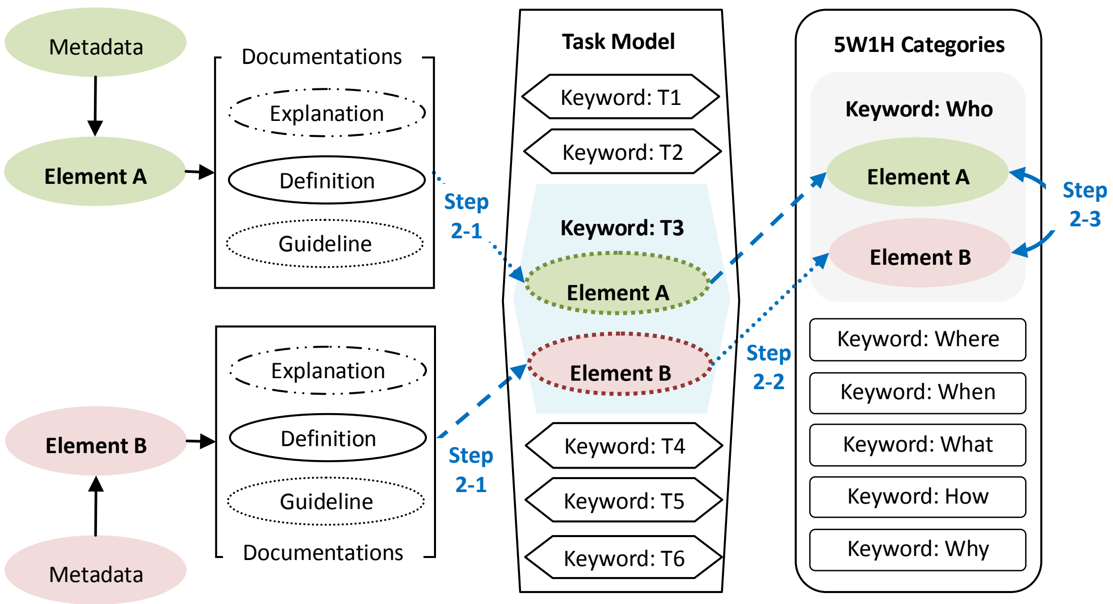


Figure24. Mapping and Classification of Metadata Elements in the Task Model and 5W1H categories (Step 2)

The paragraphs below explain the classification steps of the Description element of AGLS whose definition is shown in Figure 25. The definition text of Description property "an account of the resource" is insufficient to judge its category. So we use the guidelines text. The guidelines text includes the purpose and method of use, and the role of the element. Here, we find a phrase "*Description* of the content and/or purpose of the resource". We finally classify Description into *What*, matching words in this phrase to the keywords list of Table 8, e.g. "description", "content".

The guideline in Figure 25 says “resource discovery, remembering that search engines” as its purpose of use. The words “discovery” and “search” match with Use and Access listed in the keywords list of tasks, T1, T2, and T6 of Task Model. So we assign AGLS Description to these tasks.

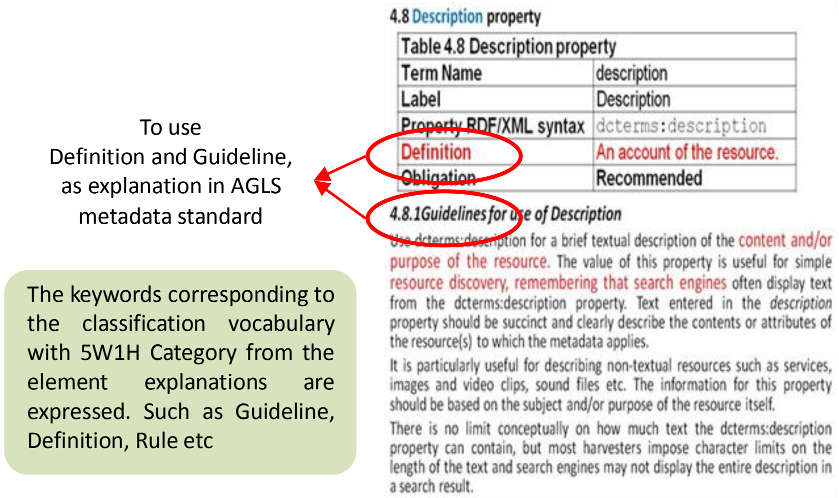


Figure25. Term Definition of AGLS Element

5.3 Mapping Metadata Schemas in 5W1H Categories and Lifecycle Tasks

This section shows 5W1H categories and task groups by example mappings among the elements of metadata standards chosen for the comparison. An example, the paragraphs and tables below show the classification and mapping examples of elements chosen from the metadata standards

5.3.1 Classification of Descriptive Elements in 5W1H Categories

(1) *Publisher* of AGLS Metadata

The *Publisher* element of AGLS means an entity responsible to make a resource available. AGLS says that this element may be used to provide details of the organization that provides access to the service. As shown in Table 8, vocabulary of 5W1H categories, agents such as organizations and institutions are often used as a location. Therefore, agent by Table 8 includes both *Who* and *Where*.

Corresponding elements of EAD and OAIS in these categories are shown in the table. These elements have similar keywords and meaning, like AGLS. So, we classified equally those both *Who* and *Where*. Elements of other standards have not corresponding elements to this element. It means that other standard elements have no corresponding vocabularies of this element.

Metadata which have no corresponding elements to *Publisher*, AGRkMS is the minimum metadata standard for record management. AGRkMS use general metadata element that describes resource, from AGLS. In PREMIS, the element which has relevance to intellectual entity is premised on using from other metadata standard. And, DPC does not contain element about intellectual contents that AGLS express, because of the attribute which is extracted from the evaluation process for preservation.

Table10. AGLS: *Publisher*

5W1H categories	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. The category of these elements is obviously *When*. Corresponding elements with *Date Range* element of AGRkMS includes AGLS, EAD, OAIS and PREMIS. Elements which correspond with Date

Range element of AGRkMS are elements of AGLS, EAD, OAIS and PREMIS, as shown in table 11.

Table11. AGRkMS: *Date Range*, *Start Date* and *End Date*

5W1H categories	Metadata Standards					
	AGLS	**AGRkMS**	DPC	EAD	OAIS	PREMIS
When	Date	**Date Range**		Date	Date of Publication	dateCreated ByApplication
		Start Date			Change History Before Archiving	
						Preservation LevelDateAssigned
		End Date				

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

Multiple media formats element of the DPC attributes means that a resource could have more than one media format regardless of digital or non-digital.

Here, *format* means a type of media of a resource and also a technology required to render a resource. Therefore, the former is categorized in *What* and the latter in *How*. Corresponding elements of the DPC attributes in these categories have AGLS, AGRkMS.

Table12. DPC Attribute Set: *Multiple media formats*

5W1H categories	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*. Corresponding elements of AGLS, AGRkMS, and OAIS in these categories are shown in the table 13.

Table13. EAD: *Title of the Unit*

5W1H categories	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. As shown in Table 8, reasons or purposes which create, manage, destroy and preserve resource includes in *Why*. This element is categorized in *Why*. Corresponding elements are shown in AGLS. *Description* of AGLS is included here as an element of a broader meaning.

Table14 OAIS: *Reason for Creation*

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

(6) *Size* of PREMIS

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

Extent of EAD as well.

Table15 PREMIS: *Size*

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

5.3.2 Mapping in the Task Groups

Followed by mapping of descriptive elements in 5W1H categories, this section shows the classification obtained by an application of Step 2 to the metadata standards. This section describes as example, which shows 'T3: Evaluate, Select, Discard' of the Task model. A part of the whole classified table shows Table 16.

T3 is associated with "Appraisal and Disposition" in the lifecycle (Figure 4) where the resource is selected and evaluated for archiving. We have classified elements of all metadata standards

By the keywords discussed in section 5.2.1 and shown in Table 9, T3 includes the keywords, such as appraisal, selection, destruction, approval etc. The result of classification that performed using these keywords, no element of PREMIS is included T3. PREMIS has no element directly related to T3, Because PREMIS is primarily designed for the 'Preservation' stage in the records lifecycle.

Table 16 shows a part of the all mappings among the elements classified to the task group T3. This mapping table shows the relationships between the elements classified into the 5W1H category in each task group. Format of *How* which expresses the format/environment (a technology/format that has a technical meaning) for performing a resource in T3, is mapped *Format* of AGLS, *Format* of AGRkMS, many elements of DPC and *Table Column Specification** of EAD etc. This description is a part of examples and Format corresponds to more metadata elements.

The classification table is shown appendix 4, because they are too large to include in this section.

Table16. Example of Mapping of Metadata Schemas for Archiving and Preservation

(T3: Evaluate, Select, Discard)

Task model	5W1H	AGLS	AGRkMS	Decision Tree of DPC	EAD	OAIS	PREMIS
T3	Who		Position				
					Sponsor		
					Publisher		
					Publication Statement		
					Author		
	When	Date	Date Range Start Date		Date		
					Publication Statement		
					Date of Unit		
			End Date				
	Where		Identifier Scheme				
			Position				
					Sponsor		
					Publisher		
					Author		
	What		Permission				
		Mandate					
		Right	Security Caveat				
			Right				
				Long term value Justify preservation			
					Ingest Process History	Processing Information	
						Appraisal Information	
			Identifier				
			Jurisdiction				
		Format	Format				
	How		Permission				
		Mandate		Negotiate for the source to supply			
		Right	Right				
		Format	Format	Digital version be selected for preservation			
						Table ColumnSpecification	
				Manageable file format			
			Document Form				
					Ingest Process History		
			Change history				
						Revision Description	
	Why			Long term value justify preservation			
				Acquire for other purposes			
			Document Form				

5.3.3 Analysis of Metadata Schemas using 5W1H Categories and Task Group

In this section, we analyze the result of classification and mappings shown in the previous sections. We created tables using the classification presented in the previous section. Table 17 and Table 18 show statistics of the classification of the elements into the 5W1H categories and task groups, respectively.

In Table 17, a number in a column shows how many descriptive elements of each standard are classified into each of the 5W1H categories. For example, the numbers of AGLS elements classified into *Who*, *When*, *Where*, *What*, *How* and *Why* are 4, 1, 4, 15, 7, 1, respectively.

The bottom row shows the total number of elements for every standard. Because an element can be classified into one or more 5W1H categories, the sum of the 5W1H rows may not be the same as the number of elements shown at the bottom of each column.

Table 17 shows that the most common portion of the AGLS elements is *What*, but in the case of PREMIS the most common portion is *How*. This means the descriptive element of AGLS expresses the meaning of ‘descriptive information about a resource’ or has the relating elements. And PREMIS mean there are many elements that express the meaning about ‘ a technical feature about resource ’.

Table17. Metadata Standards in 5W1H categories shown by Figures

Metadata / 5W1H	AGLS	AGRkMS	DPC	OAIS	EAD	PREMIS
	19	20	27	53	146	95
Who	4	0	0	4	18	2
When	1	1	0	2	9	7
Where	4	2	0	8	23	7
What	15	15	16	24	99	21
How	7	13	12	43	47	125
Why	1	1	2	3	0	2

Table18. Metadata Standards in the Task Groups shown by Figures

Task Group	5W1H	AGLS	AGRkMS	DPC	EAD	OAIS	PREMIS
Task 1	Who	3			15		
	When	1	1		5		1
	Where	3	1		16		
	What	6	3		26		
	How	3	1		11		2
	Why	1				1	
Task 2	Who	2			15	1	
	When	1	1		5		
	Where	1	2		18	1	
	What	11	14		29	3	
	How	5	11		9	2	11
	Why		1				
Task 3	Who				16		
	When	1	1		5		
	Where		1		15		
	What	10	7	5	19	1	
	How	6	8	11	10	1	
	Why	1	1	2			
Task 4	Who				17	3	
	When	1	1		9	2	1
	Where		2		12	7	1
	What	3	14		91	23	2
	How	3	11		46	42	31
	Why		1			2	
Task 5	Who				17	2	2
	When	1	1		8	2	5
	Where		1		21	7	5
	What	3	1	11	61	21	18
	How	3	1	1	32	42	119
	Why					3	2
Task 6	Who	2			16	1	
	When	1	1		5		
	Where	1			17	1	
	What	4			24	3	
	How	3			10	3	11
	Why						

We arranged corresponding metadata standard to 5W1H categories in each task group and, expressed by figures. Table 18 shows similar statistics sorted according to the task groups. This shows a feature of the metadata standards discussed in section 4. This table is sorted by the task groups but not by the lifecycle stages used in our first study.

Table 19 shows the overall distribution of elements in the task groups. Each row of this table shows values for each task group. A column shows values for a standard. Each value in a box contains a percentage of elements classified to a corresponding task group. This table shows a feature of the metadata standards analyzed from the viewpoint of the task groups. It shows a feature similar to but more refined than in our first study shown in section 4.

Table19. Metadata Standards in the Task Groups by Percentage

a. The highlighted metadata in task group (from each row)

Task Group	AGLS	AGRkMS	DPC	OAIS	EAD	PREMIS
Task 1	**21**	7		12	1	1
Task 2	25	**33**		13	4	5
Task 3	22	21	**60**	11	1	
Task 4	9	33	40	29	**45**	17
Task 5	9	5		23	44	**72**
Task 6	**14**	1		12	5	5

b. The highlighted task (from each column)

Task Group	AGLS	AGRkMS	DPC	OAIS	EAD	PREMIS
Task 1	21.	7		12	1	1
Task 2	**25**	**33**		13	4	5
Task 3	22	21	**60**	11	1	
Task 4	9	**33**	40	**29**	**45**	17
Task 5	9	5		23	44	**72**
Task 6	14	1		12	5	5

Each column shows the distribution of elements in the lifecycle. For example, AGLS could be used well in task 1, 2, 3 and 6, and PREMIS could be used in task 5. The boxes surrounded by bold lines show the highest value for each standard, and can be interpreted to imply a main task to which the standard is well suited.

Each row of the table shows the different weightings of a task for each standard. The highlighted boxes show the highest values in a row, which would mean the most suitable standard for each task.

Table 19a shows the highlighted metadata in the task groups, from each row (view of task). For example, Task 2 shows the highest value in AGRkMS, Task 4 shows the highest value in EAD. Table 19b shows the highlighted task from each column (viewpoint of metadata standard). For example, AGLS is high-lighted for task 2 and PREMIS shows the highest value for task 5. Percentage is rounded. The highlighted boxes have the highest number in each row.

5.4 Consideration

The fundamental point of this study is to see metadata standards from a task-centric 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.

We consider that the core contribution of this study is a shift of our viewpoint from a resource-centric view to a task-centric and lifecycle-centric view. It is often the case that information about tasks and lifecycle stages is not explicitly defined in the metadata elements. The contribution of this study is also the use of contextual information extracted from the records lifecycle model. We consider that the two models –Task and 5W1H categories – 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. The task-centric view proposed in this paper helps with access to archived information resources across repositories and over time.

Contextual semantics are implicit in the definition of metadata elements, which is one of the major barriers to creating mappings between metadata standards. A shift in the viewpoint of metadata elements, i.e. from resource-centric to task-centric, helps us find and use the contextual information in metadata mappings.

In this research, we proposed the 5W1H categories and the 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. Thus, we defined the Task model using the 5W1H categories for metadata mappings to improve metadata interoperability over the whole lifecycle. We learned that it is crucial to combine metadata standards for archiving and preservation of digital resources.

6 Discussions

This chapter re-examines the study presented in this paper from several viewpoints: comparison with related researches (section 6.1), metadata standards for archiving and preservation (section 6.2), feature analysis of metadata standards (section 6.3), Task-oriented model and 5W1H categories (section 6.4), metadata mappings based on the Task model (section 6.5), and discussion summary (section 6.6). Because the research is primarily based on qualitative analysis, this section contains a discussion section that re-examines the methods and results.

6.1 Related Research on Metadata for Archiving and Preservation

This dissertation presented a study on metadata standards for archiving and preservation from various viewpoints. In order to perform a “feature analysis of archival metadata standards” for long-term preservation of digital resources, the author introduced related research in section 2.8. This section discusses the differences and similarities between the author’s studies and related research in more detail.

1) ‘*Create Once, Use Many Times: The Clever Use of Recordkeeping Metadata for Multiple Archival Purposes*’ [21]. The paper analyses the development of recordkeeping metadata for multiple archival purposes and looks at the relevance to future archival systems. The Clever Recordkeeping Metadata Project (CRKMP) explains metadata interoperability and uses the Records Continuum theory as a conceptual framework. The paper did not show how to use the records continuum theory in detail, although it does describe some of its aspects. The author of this dissertation used the records lifecycle as a united framework in her studies. It is a point of similarity between the two studies that they both mention the records lifecycle and use the records continuum theory as a framework for analysis of archival metadata element or recordkeeping metadata elements. However, the paper did not mention the relationships between metadata standard and the records continuum theory, and did not provide a detailed description about the records continuum theory.

2) ‘*Metadata Elements for Object Description and Representation: A Case Report form a Digitized Historical Fashion Collection Project*’ [70]. The paper develops a

catalog for digitized historical fashion collection objects, and carries out a comparison between selected metadata elements (USMARC, DC, VRA) and the *desired elements*, which are proposed in the paper. The paper describes how to choose, compare and use the different elements of metadata schemas. This kind of mapping is similar to the semantic mapping that the author of this dissertation performed, where she carried out mapping using keywords extracted from the documentation of metadata elements. These two studies are similar in so far as they use parts of the element descriptions for metadata mapping.

3) '*A Survey of Techniques for Achieving Metadata Interoperability*' [27] describes the metadata used in current information systems and goes on to an examination of metadata interoperability and related problems. The paper gives suggestions on how to compare and map between metadata schemas. Metadata interoperability plays an important role in the archiving and preservation of digital resources. A study carried out by the author proposed a model to improve metadata interoperability and analyzed various features of metadata standards for long-term preservation of digital resources. The survey presented in the paper is not directly related to this dissertation, but it has helped to clarify the importance and purpose of metadata interoperability for research purposes.

4) '*Interdisciplinary Contents Management Using 5W1H Interface for Metadata*' [56] proposes a metadata exchange interface for interdisciplinary contents-sharing. In the paper, an interface for a metadata abstraction module for contents-circulation across various disciplines was designed using the concept of 5W1H. In addition, the study shows that elements of Dublin core can be converted into the 5W1H elements. The author uses the 5W1H categories to identify the context of the resources which are described using the metadata. The use of the 5W1H categories is a unique feature of these studies. The similarity of the two studies (the paper and the authors study) is to use the viewpoint of 5W1H. That is, the similarity between the two studies is found in the fact that they both convert and classify metadata elements using 5W1H.

5) '*A Metadata Lifecycle Model for Digital Libraries: Methodology and Application for an Evidence-based Approach to Library Research*' [13] describes and proposes the Metadata Lifecycle Model (MLM) as a methodology for the whole process of metadata provision for digital libraries. The MLM involves a ten-step process by which digital library projects can design and implement metadata provision. The purpose of the

model is to achieve a consistent method for developing metadata for digital library projects, and to conduct a content-based analysis for digital collections. The MLM and the records lifecycle (Task model in the authors study) are dissimilar when it comes to purpose, object and content. But there are similarities between the two studies (the paper and the authors study) when it comes to analyzing metadata schema from the viewpoint of the records lifecycle.

6.2 Metadata Standards for Archiving and Preservation

Metadata is one of the most important components in the archiving and preservation of digital resources. In general, every metadata schema has its base data model. Every metadata element is defined as a property (or an attribute) of an entity included in the data model. Metadata is used in the tasks of the records lifecycle. A metadata standard is characterized not only by its base data model but also by the tasks in the stages of the records lifecycle. However, the definitions and data models of metadata standards are generally not explicitly defined based on the resource lifecycle.

There are several metadata standards for digital archiving and preservation, i.e., EAD, ISAD (G), OAIS, PREMIS and so forth. Every standard has its own features in accordance with its primary application domain. Archival metadata standards are used primarily to manage resources in the later stages of the lifecycle.

Throughout her studies, the author has confirmed her beliefs that any single metadata standard is not sufficient to cover the whole lifecycle. This means that in order to define a metadata schema used in the lifecycle, metadata standards should be selected and combined suitably according to the requirements given at each lifecycle stage. The Dublin Core Application Profile gives us good guidelines to select and combine metadata standards but it does not provide guidelines on how to combine metadata standards in accordance with the resource lifecycle. Based on this understanding, the author clarified the need for an analysis of mapping between metadata standards in accordance with the records lifecycle. Mapping and performing crosswalks between metadata standards for data exchange are needed. In other words, selection of suitable metadata standards is crucial for the archiving and preservation of digital resources.

6.3 Feature Analysis of Archival Metadata Standards

Based on the view presented in section 6.2, the author carried out a study to clearly identify archival metadata features from the viewpoint of the records lifecycle, and proposed a methodology to analyze archival metadata schemas.

The author used the primary lifecycle stage, which is determined based on the value assignment to a metadata element as the key to characterize every metadata standard.

In this study, the author first identified the primary lifecycle stages for each metadata element, from which she identified primary stages of each standard. To give an example, ISAD(G) covers '*Store & Arrange*', PREMIS covers '*Preserve*' in the lifecycle. Next, the author analyzed metadata elements according to the tasks performed on the resource (a task-centric view) to clarify the relationships between the metadata elements and tasks.

The analysis using this viewpoint is the core contribution of this study – i.e. a shift from a resource-centric view to a task-centric view of metadata standards. In general, data models of conventional metadata are defined from a resource-centric standpoint but not a task-centric standpoint. However, the author concluded that a metadata element is affected by a task in the records lifecycle.

One of the most important findings that the author learned from this study is that a task-centric view of metadata standards is crucial to define a framework for organizing metadata schemas throughout the resource lifecycle and for interoperability of metadata schemas used at different stages of the lifecycle. In other words, the shift from a resource-centric view of metadata standards to a task-centric view is a core contribution of this study.

The Dublin Core Application Profile (DCAP) is a well-known framework to enhance metadata interoperability. It suggests to mix-and-match metadata vocabularies to develop an application metadata schema. The author therefore formed the hypothesis that any single archival metadata standard is not sufficient to cover the whole records lifecycle. The first study in the dissertation – a feature analysis of archival metadata standards – has proved that the author's hypothesis is true, which can be expected according to the DCAP. However, the fundamental difference between the author's analysis and DCAP is that this study includes a time line but DCAP doesn't. The second study in the dissertation proposed the Task Model, which was used to clarify the features of metadata elements and to create mappings among archival metadata standards. The mappings are the fundamental basis for the semantic interoperability of metadata. Thus, this study has shown a novel model to enhance interoperability of archival metadata which requires semantic linkages among metadata elements across lifecycle stages.

6.4 The Task-oriented Model and 5W1H Categories

Through section 6.2 and 6.3, the author identified the relationships between metadata standards and lifecycle tasks. Therefore, the author proposed a Task-oriented model (i.e. Task Model) to show metadata standards of resource-centric from the view of lifecycle tasks. The Task Model is created based on the records lifecycle to improve metadata interoperability over the whole lifecycle. The Task Model shows the relationships between the task groups, resources, and lifecycle stages.

A task creates an ‘Event’ performed on a resource. A resource is affected by the ‘Event’. Thus, an execution of a task causes an ‘Event’ on resources. Thus, the author used and described 5W1H categories to describe an ‘Event’, and to classify ‘A Task’.

The author proposed to use 5W1H categories to categorize tasks in detail and to classify the metadata elements according to each task. Thus, the author thinks that 5W1H categories are useful in analyzing the metadata elements as a new viewpoint based on tasks.

The model proposed to clarify the features of metadata standards is a major contribution of this study – i.e., the Task-centric model and 5W1H categories as a framework for feature analysis of archival metadata standards. The author believes that the Task model can be used to suitably select and combine elements from different metadata standards as needed according to lifecycle stage. That is, the Task model is proposed as a new tool of the model, which improves the interoperability of metadata standards in the lifecycle. The author thinks 5W1H category supports analytically understanding the meaning of a metadata element.

6.5 Metadata Mappings based on the Task Model

The approach used in this study to improve metadata interoperability in the resource lifecycle was to map the metadata elements based on the Task model, and to perform the classification of elements using the 5W1H categories in each task group.

In order to classify the metadata elements in the context of each task, this study determined a set of keywords based on features of the Task model and 5W1H categories. The author used these keywords to perform semantic mapping among the elements of the metadata standards chosen in this study, i.e., AGRkMS, PREMIS, EAD, and so forth. The mapping and classification in each task group was performed using the proposed keywords.

Metadata vocabulary mapping is not a new topic. It is primarily required for the interoperability of metadata, i.e. mapping between two elements from different metadata schemas. The author carried out metadata vocabulary mapping manually because it was necessary to interpret the meaning and purpose of the element definitions. That is, the author used contextual information extracted from the lifecycle in order to identify the meanings of the metadata elements.

One of the most important points in this study is the use of the information about context in the lifecycle, e.g., rules implicitly defined in the standards, relationships between use of elements and stages. The author has learned that it is necessary to use not only the semantic description given in the definition of a metadata element but also the context information of the element, which can be obtained from the lifecycle and the Task model.

The general metadata mapping was performed to find and classify semantic similarity among metadata elements. However, the author performed the mapping using not only the definitions of the elements but also contextual information of the elements. In addition, the author proposed to characterize the metadata elements in the context of each task, extracting the definition from six aspects using 5W1H categories. The author believes that the same contextual information in a task and in the 5W1H categories is useful to semantically link metadata elements.

6.6 Discussion Summary

One of the most difficult aspects of this research was the manual mapping and classification of the metadata vocabularies. The author has not yet applied the mapping table to test metadata interoperability in a practical environment due to a limitation of the resources available for her research.

In the study, evaluation of the mapping and classification by system (or tool) has not been carried out yet. Therefore, the author has not included an evaluation of mapping in this paper.

The author believes that evaluation of the semantic mappings between different metadata elements is necessary and important.

As the goal of this study is to propose a unified framework that improves the interoperability between metadata elements, creation of the mappings that cover several major standards and are carried out by manual but semi-formalized process, is sufficient to show the feasibility of the framework as the goal of this study. Evaluation of the mappings based on real metadata done by machines is left for future work. In addition, the author has left the development of software tools for task groups as an object of future study.

The author proposed the Task model and 5W1H categories as a framework. Therefore, the author has identified the relationships between the task groups and the metadata standards. In addition, she has found that metadata elements are affected by tasks and should describe a resource according to the Task model.

The author analyzed the features of archival metadata standards using two different approaches, i.e., the records lifecycle and the Task model. The outcome of studies that performed using two different approaches makes no odds. It is a natural result.

Through this study, the author learned that it is important to carry out appropriate mapping between metadata standards. In addition, the author is convinced that a combination of metadata standards for archiving and preservation of digital resource is important. The author identified the relationships between a metadata standard and a task through these models – Records Lifecycle, Task model and 5W1H categories. The author thinks that an analysis of the relationship between a task and metadata is useful for selecting and using the different metadata elements in the whole lifecycle.

Furthermore, the author believes that the models she has created improve the interoperability of metadata.

7 Conclusion

Digital resources are widely used in our modern society. The rapid growth of digital resources has not only the popularization of digital resource but also some major problems. One of these problems is to manage and maintain digital resources for future generations. Thus, we are facing fundamental problems of how to manage and preserve digital resources over time.

For archiving and long-term preservation of digital resources, proper policies and strategies (developing systems, guidelines, metadata schemas and so on) are necessary. Several standard methods for preserving digital resources have been developed and are in use. It is widely recognized that metadata is one of the most important components of archiving and preservation of digital resources. In this study, the author shows features of archival metadata standards throughout the whole lifecycle, in order to analyze metadata standards for digital archiving and preservation.

There are many metadata standards for archiving and preservation of digital resources, where each standard has its own feature in accordance with its primary application. In addition, metadata standards have a base data model, and a metadata element is defined as a property (or an attribute) of an entity included in the data model.

On the other hand, metadata standards are affected by tasks performed in the records lifecycle. Metadata has to be used in accordance with the tasks. However, in general, the data model is not explicitly linked to the records lifecycle or tasks, which means that users have to find appropriate metadata standards in accordance with the lifecycle stages.

It is crucial to select and combine metadata standards in accordance with requirements in an application domain and in the records lifecycle. This study identified and analyzed features of archival metadata standards to select, combine and use them appropriately throughout the resource lifecycle, for archiving and preservation of digital resource.

In order to analyze the features of the metadata standards, the author identified the primary records lifecycle stage(s) where a standard would be applied. As a result of this analysis, she clarified the features of the standards from the viewpoint of relationships between the elements and the lifecycle stages. In addition, she found that a metadata standard element is related to a task.

Based on this feature analysis, this study has proposed the Task Model to clarify tasks in the records lifecycle and to categorize metadata elements from the viewpoint of the tasks. Based on this, the author has proposed to categorize metadata elements using 5W1H categories coupled with the Task model derived from the resource lifecycle. In this study, metadata elements of the chosen standards are categorized using the 5W1H categories and mapped to each other. The mappings are grouped and sorted in accordance with the Task model.

Mapping between metadata schemas is often required throughout the preservation process because different schemes are used in different stages of the records lifecycle. Therefore, it is crucial to build a unified framework to enhance the interoperability of metadata schemas. 5W1H categories and the Task model are used as a unified viewpoint in this study. The author thinks that the proposed models help identify the contexts of descriptive elements and define crosswalks among standards. This study presents a basis for the interoperability of different metadata schemas used in digital archiving and preservation.

A major achievement of this study is the feature analysis of archival metadata schemas from the two viewpoints, a records lifecycle-view and a Task model-view. And the core contribution of this study is a shift from a conventional resource-centric view to a task-centric and lifecycle-centric view. Through this study, the author has learned that a metadata standard is related to a task in the records lifecycle. She also has learned that any single metadata standard for archiving and preservation does not cover the whole resource lifecycle.

The author has not yet applied the mapping table to test metadata interoperability in a practical environment due to the limitations of the resources available for this study. She understands that such a test is important to evaluate the mappings but has had to leave this for her future studies.

Another issue reserved for future study is to introduce the concept of application profiles into the task-centric model. This is because the metadata schemas expressed as application profiles are primarily resource-centric and task-oriented information is not explicitly described as a part of metadata schema. She thinks that a task-centric application profile for archival metadata may help with metadata interoperability and may help to select necessary metadata elements for each task.

It took the author a long time to accomplish the goal of this study. She was originally interested in studying digital archiving and preservation, and studied archival theory in her master’s course. Unfortunately, she lacked technical knowledge about metadata standards and metadata schemas. The author needed much time in order to gain a basic understanding of digital archiving and preservation including knowledge of metadata. Particularly, it was necessary to spend a lot of time to analyze the features of various metadata standards.

The author performed her general studies on long-term preservation and selection of digital resource, prior to starting on archival metadata. She surveyed and studied guidelines for long-term preservation of digital resources, and policies and guidelines for resource-selection. These researches were not directly used in this dissertation, but the author believes that they will greatly help her study on long-term preservation of digital resource in the future. The author will continue her studies on digital archiving and preservation. In addition, she hopes that her studies will produce useful insights on digital archiving and preservation in the future.

Acknowledgements

The author is grateful to my parents and brother who gave me the constant support and a prayer. The author is thankful to Prof. Shigeo Sugimoto who gave guidance till the completion of this dissertation without giving up.

In addition, the author thanks Jan Askhoj, a PhD student at her laboratory for his help and contribution in the discussion. The author thanks Profs. Liddy Nevile, Mitsuharu Nagamori and all members of our laboratory for their useful comments and support. In order to convey the appreciation to her family and friends in Korea and Japan, the author expresses her acknowledgements in Korean and Japanese below.

'박사과정' 이라고 하는 새로운 마라톤의 티켓을 저에게 주셔서, 새로운 풍경을 바라보며 새로운 길을 달리고, 바람을 느낄 수 있도록 큰 기회를 주신 하나님께 감사드립니다.

끝이 보이지 않았던 길, 이제까지 경험하지 못하고 계획하지 않았던 이 길을 달리면서, 많은 후회와 불만도 있었지만, 이런 저에게 포기하는 마음보다는 달콤한 사탕과 따뜻한 커피로 여기까지 인도해주셔서 너무 감사드립니다. 처음 이 길에 섰을 때 말씀드린 것처럼 아직 많이 부족하지만, 이 사회에 있어 중요한 사람이 될 수 있도록 앞으로 더욱 노력하겠습니다.

긴 시간동안 저의 손을 놓지 않고 함께 비와 눈을 맞아주면서, 길을 걸어주고 끊임없는 응원을 보내 준 가족에게 너무나 감사드립니다. 박사 공부에 대한 길을 열어주시고 많은 지원을 허락해 주신 우리 할아버지와 할머니 (Baek Mun Yong and Lee Jae Kyoung, Yoo Sik Woo and Jang Mu Hee), 쯔쿠바 대학를 찾아서 제일 먼저 알려주시고, 언제나 앞서서 모든 것을 준비해주신 아빠, 항상 저의 뒤에서 묵묵하게 많은 지원을 해주시면서 쓴 소리보다는 격려를 더 많이 해주시며 저를 믿어주신 아빠 (Baek Kwang Jin), 매일 새벽 부족한 딸을 위해 끊임없는 기도를 해주시고, 항상 올바르고 따뜻한 지혜로 언제나 저의 버팀목이 되어 주신 엄마 (Yoo Myoung Hee), 그리고 많은 칭찬과 가끔은 쓴소리로, 저를 든든하게 지치지 않도록 이끌어준 멋진 동생과 우니 (Baek Seung Woo and Uni), 우리 가족 모두에게 말로 표현 할 수 없을 만큼 너무나 감사드립니다.

박사과정, 쯔쿠바라고 하는 새로운 환경을 보여주시고 이 길에 들어설 수 있도록 지도해주신 덕성여대의 모든 교수님 (강미혜 교수님, 유재옥 교수님, 고 전명숙 교수님)과 숙명여대의 모든 교수님 (이춘실 교수님, 오경묵 교수님, 장윤금 교수님) 께 진심으로 감사드립니다.

쯔쿠바 대학의 박사과정의 길에 들어설 수 있게 도와 주신 이희재 교수님과 김성혁 교수님께 진심으로 감사드립니다. 큰 목표를 정하면 목표의 근처까지 갈수 있다고 말씀해 주신 이희재 교수님, 큰 고민을 가지고 있어도 달려가면 언제나 그 곳에서 교수님을 뵐 수 있었지만, 이제는 하늘에서 교수님께서 보셨을 때, 그리고 교수님을 다시 뵈었을 때 하나도 부끄럽지 않은 제자가 될 수 있도록 더 큰 목표를 향해서 더욱 열심히 걸어가겠습니다

때와 장소에 상관없이 메타데이터에 대해 설명을 해주신 김성혁 교수님, 찾아 뵈면 언제나 응원을 먼저 해주시고, 앞으로의 길에 대해 아낌없이 조언을 해주신 교수님, 진심으로 감사드립니다. 교수님께서 일러주신 길을 참고하며 앞으로 나아가겠습니다.

처음에는 끝이 안보이는 길과 터널이었지만 지금 이 자리에서 그 길을 돌아보니, 이 길을 저에게 주신 모든 분들, 저를 여기까지 인도해주신 모든 분들, 그리고 저를 위해서 기도해주신 모든 분들께 너무 감사드리고 또 감사드립니다.

처음 마라톤을 시작할 때의 기대와 설레임, 그리고 마라톤을 하면서 힘들었던 점과 어려웠던 점, 이 모든 것을 잊지 않겠습니다.

아직 많이 부족합니다. 많이 부족한 아이가 대단한 교수님들을 만나 그들에게 배울 수 있는 기회를 주시고, 그 속에서 같이 공부할 수 있게 이 길을 주셨던 것 처럼, 저 역시 이 길을 통해서 배우고 느낀 모든 것을 머리와 가슴에 담고 또 담아, 저를 필요로 하는 누군가에게 쓸모가 될 수 있도록 더욱 노력하겠습니다. 그리고 어딘가에 있을 새로운 길을 달릴 준비를 시작하겠습니다.

筑波大学での勉学と生活のいろいろな面で助けて下さったすべての方に心より感謝します。

偶然なきっかけで会ったにもかかわらず、快く指導教員になってくださった杉本先生、本当にありがとうございます。博士後期課程の学生として受け入れて下さって、長い間、見捨てることなく導いて下さったことに対し、本当に感謝いたします。

明け方、真っ赤なコメントでいっぱいのメールを受け取るときはいつもドキドキしましたが、もうそのドキドキさが無くなると思うとすごく寂しくなりそうです。いつも気になる点や、知らない部分があると、すぐ先生に質問しましたが、これからはどうすれは良いのかが凄く心配です。しかし、先生とお会いしてから英語で論文を書くことや、国際学会へ行って発表をすることもでき、そして、見聞を広げることができるように様々な機会をくださってありがとうございます。

知識不足な私のために、一対一の授業をして下さって、間違えている点は正確に指摘して下さった韓国語が上手な永森先生、ありがとうございました。いつも明るい笑顔で良いお話をしてくださって、映画という新しい視野を持つようにしてくださった西岡先

生、私の知識不足のために、多くの質問をいただきましたが、その質問を通じて研究をより頑張るきっかけを与えて下さった阪口先生と森嶋先生、心より感謝いたします。

杉本研究室で一緒に研究したすべての友人、いつも休む空間を作ってくれて、誰より早く多い称賛と激励をしてくれた馮暁暁さん、英語のチェックから論文のチェックまでいかなるときも不平を言うことなく本人のことのようにチェックしてくれたヤン･アシコイさん、メタデータに対することから、研究に対する質問と議論、コンピュータのことまで、いつも助けてくれた本間維さん、そして、研究をいつも共にしてくれた両角彩子さん、久保順子さん、三原鉄也さん、杉本研のすべての学生の方々に本当に感謝します。

最後に、1 年間、私に大きな助けを下さった「結城ロータリークラブ」の皆様と、自分の子供のようにいつも暖かく接してくださって、多くの経験ができるようにして下さった四宮英男氏、本当にありがとうございました。おかげさまで暖かい心をいっぱい学び、感じることができました。

時々、この場所が恋しくなると思いますが、ここで習った全てのものと、色々な方々から頂いた激励と指摘、考え方など、すべてを抱いて、これからも努力したいと思います。

Reference

[1] Archive Systems. *Business Records Storage*. Archive Systems. Retrieved April 30, 2013, from http://www.archivesystems.com/services/records-management-services/business-records-storage.aspx.

[2] Alemneh, Daniel Gelaw, Hastings, Samantha Kelly and Hartman, Cathy Nelson. (2002). A Metadata Approach to Preservation of Digital Resources: The University of North Texas Libraries' Experience. *Peer-Reviewed Journal on the Internet*, 7 (8/August). Retrieved February 15, 2013, from http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/rt/printerFriendly/981/902.

[3] Baca, Murtha. (2003). Practical Issues in Applying Metadata Schemas and Controlled Vocabularies to Cultural Heritage Information. *Cataloging & Classification Quarterly*, 36(3-4), 47-55.

[4] Baek, Jae-eun and Sugimoto, Shigeo. (2007). *Selection Guidelines for Preservation Method of Digital Resources*. Retrieved February 15, 2013, from http://www.tulips.tsukuba.ac.jp/limedio/dla m/M93/M932902/9.pdf.

[5] Baek, Jae-eun and Sugimoto, Shigeo. (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.

[6] Baek, Jae-eun and Sugimoto, Shigeo. (2011). Facet Analysis of Archival Metadata Standards to Support Appropriate Selection, Combination and Use of Metadata Schemas. In: *Proceeding of International Conference on Dublin Core and Metadata Applications, The Hague, Sep 2011. Netherlands*, 1-11.

[7] Baek, Jae-eun and Sugimoto, Shigeo. (2012). A task-centric model for archival metadata schema mapping based on the records lifecycle. *International Journal of Metadata, Semantics and Ontologies archive*, 7(4), 269-282.

[8] Besser, Howard. (2000). Digital longevity. *Northeast Document Conservation Center*, 155-166. Retrieved April 30, 2013, from http://besser.tsoa.nyu.edu/howard/Papers/sfs-longevity.html.

[9] Bountouri, Lina and Gergatsoulis, Manolis. (2011). The Semantic Mapping of Archival Metadata to the CIDOC CRM Ontology. *Journal of Archival Organization*, 9(3-4), 174-207.

[10] Brian F. Lavoie and OCLC. (2000). *Meeting the challenges of digital preservation: The OAIS reference model*. Retrieved February 15, 2013, from http://www.oclc.org/research/publications/library/2000/lavoie-oais.html.

[11] Brian F. Lavoie. (2004). *The Open Archival Information System Reference Model: Introductory Guide*. Retrieved February 15, 2013, from http://www.dpconline.org/docs/lavoie_OAIS.pdf.

[12] Chan, Lois Mai and Zeng, Marcia Lei. (2006). Metadata Interoperability and Standardization – A Study of Methodology Part 1. *D-Lib Magazine*, 12(6). Retrieved February 15, 2013, from http://www.dlib.org/dlib/june06/chan/06chan.html.

[13] Chen, Ya-Ning, Chen, Shu-Jiun, and Lin, Simon C. (2003). A metadata lifecycle model for digital libraries: methodology and application for an evidence-based approach to library research. In: *Proceedings of the World Library and Information Congress: 69th IFLA General Conference and Council, Berlin*. Retrieved January 20, 2013, from http://www.citi.sinica.edu.tw/ papers/sophy/117-F.pdf.

[14] Consultative Committee for Space Data Systems (CCSDS). (2002), (2012). *Reference Model for an Open Archival Information System (OAIS)*. Retrieved April 30, 2013, from http://public.ccsds.org/publications/archive/650x0m2.pdf.

[15] Coyle, Karen, Baker, Thomas and DCMI. (2009). *Guidelines for Dublin Core Application Profiles*. Retrieved April 30, 2013, from http://dublincore.org/documents/profile-guidelines/.

[16] Day, Michael. (2001). Metadata for Digital Preservation: A Review of Recent Developments. In: *Proceedings of the Research and Advanced Technology for Digital Libraries: 5th European Conference (ECDL 2001)*, *Darmstadt*, *Germany*, 161-172. Retrieved February 15, 2013, from http://link.springer.com/content/pdf/10.1007%2F3-540-44796-2_15.pdf.

[17] Digital Preservation Coalition. (2006). *Decision Tree for Selection of Digital Materials for Long-term Retention*. Retrieved February 15, 2013, from http://www.dpconline.org/advice/preservationhandbook/decision-tree.

[18] Duval, Erik et al. (2002). Metadata Principles and Practicalities. *D-Lib Magazine*, 8(4). Retrieved February 15, 2013, from http://www.dlib.org/dlib/april02/weibel/04weibel.html

[19] Emery, Priscilla. (2005). *Why Records Management?*. Retrieved March, 15, 2013, from http://www.realstorygroup.com/Feature/127-RM-101.

[20] European Commission. (2008). *MoReq2 Specification.* Retrieved February 15, 2013,

from http://ec.europa.eu/archival-policy/moreq/doc/moreq2_spec.pdf.

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

[22] Ferro, Nicola and Silvello, Gianmaria. (2008). A Methodology for Sharing Archival Descriptive Metadata in a Distributed Environment. In: *Proceedings of the 12th European conference on Research and Advanced Technology for Digital Libraries (ECDL '08)*, 268 – 279.

[23] Fresko, Marc. *MoReq2 NOUVEAU EST ARRIVÉ*. Retrieved February 15, 2013, from http://www.moreq2.eu/press-a-web/files-download/118_c980bb19777df125b9557abd0a2510d

[24] Dunsire, Gordon. (2010). The Vocabulary Mapping Framework and its potential for improving metadata interoperability in the Semantic Web. In: *Proceeding of the EUROVOC Conference*, *November 2010*, *Luxembourg*.

[25] Goto, Toshiyuki. (2007). Preservation metadata: Current situation and future perspectives. *Journal of Information Processing and Management*, 50(2), 74. (in Japanese).

[26] Harvard University Information Technology. (2008). *Digital Projects: Metadata Standards*. Retrieved February 15, 2013, from http://hul.harvard.edu/ois/digproj/metadata-standards.html.

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

[28] Heery, Rachel and Patel, Manjula. (2000). *Application profiles: mixing and matching metadata schemas*. Retrieved January, 20, 2013, from http://www.agi-imc.de/internet.nsf/0/f106435e0fd9ffc1c125699f002ddf31/$FILE/dubin_core.pdf.

[29] International Council on Archives. (2000). *ISAD(G): General International Standard Archival Description - Second edition*. Retrieved April 30, 2013, from http://www.ica.org/10207/standards/isadg-general-international-standard-archival-description-second-edition.html.

[30] International Federation of Library Associations and Institutions. (2009). *Functional Requirements for Bibliographic Records – Final Report*. Retrieved February 15, 2013, from http://www.ifla.org/files/assets/cataloguing/frbr/frbr_2008.pdf.

[31] International Organization for Standardization. (2001). *ISO 15489-1:2001: Information*

and documentation - Records management - Part 1: General.

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

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

[34] International Organization for Standardization. (2012). ISO 14721:2012: Space data and information transfer systems - Open archival information system (OAIS) - Reference model.

[35] JISC. (2008). *Extension of the RDA/ONIX Framework*. Retrieved January 1, 2013, from http://www.jisc.ac.uk/media/documents/projects/vmf_proposal.pdf.

[36] JISC. (2009). *The Vocabulary Mapping Framework (VMF): an introduction*. Retrieved February 15, 2013, from http://www.doi.org/VMF/documents/VocabularyMappingFrameworkIntroductionV1.0(091212).pdf.

[37] Marywood university. *Records Management and Archives: Definitions*. Retrieved January 20, 2013, from http://www.marywood.edu/archives/definitions.html#r.

[38] MoReq Collateral Website. *What are MoReq and MoReq2?*. Retrieved January 20, 2013, from http://www.moreq2.eu/faqs.

[39] National Archives and Records Administration. *What's a Record?*. Retrieved May 1, 2013, from http://www.archives.gov/about/info/whats-a-record.html.

[40] National Archives and Records Administration. *What's an Archives?*. Retrieved January 20, 2013, from http://www.archives.gov/about/info/whats-an-archives.html.

[41] National Archives and Records Administration. *Preservation Programs at the National Archives*. Retrieved January 20, 2013, from http://www.archives.gov/preservation/internal/.

[42] National Archives of Australia. (2006), (2010), (2011). *AGLS Metadata: Australian Government Implementation Manual (Version 3.0)*. Retrieved February 15, 2013, from http://www.naa.gov.au/Images/AGLS%20Manual%203.0%20-%20most%20recent%20revision%2030%20September%202011_tcm16-49605.pdf.

[43] National Archives of Australia. (2008). *Australian Government Recordkeeping Metadata Standard (Version 2.0)*. Retrieved May 1, 2013, from http://www.naa.gov.au/Images/

AGRkMS_Final%20Edit_16%2007%2008_Revised_tcm16-47131.pdf.

[44] National Archives of Australia. (2010). *AGLS Metadata Standard Part 1 – Reference Description (Version 2.0)*. Retrieved May 1, 2013, from http://www.agls.gov.au/pdf/ AGLS%20Metadata%20Standard%20Part%201%20Reference%20Description.PDF.

[45] National Archives of Australia. (2011). *Australian Government Recordkeeping Metadata Standard Implementation Guidelines (Version 2.0)*. Retrieved May 1, 2013, from http://www.naa.gov.au/Images/AGRkMS%20Implementation%20Guidelines_tcm16-501 56.pdf.

[46] Nilsson, Mikael, Baker, Thomas, Johnston, Pete and DCMI. (2008). *The Singapore Framework for Dublin Core Application Profiles*. Retrieved January 20, 2013, from http://dublincore.org/documents/singapore-framework.

[47] OCLC/RLG Working Group on Preservation Metadata. (2001). *Preservation Metadata for Digital Objects: A Review of the State of the Art*. Retrieved March, 9, 2013, from http://www.oclc.org/resources/research/activities/pmwg/presmeta_wp.pdf.

[48] OCLC/RLG Working Group on Preservation Metadata. (2002). *Preservation Metadata and the OAIS Information Model: A Metadata Framework to Support the Preservation of Digital Objects*. Retrieved February 15, 2013, from http://www.oclc.org/content/dam /research/activities/pmwg/pm_framework.pdf.

[49] Park, Jungran. (2005). Semantic interoperability across digital image collections: a pilot study on metadata mapping. In: *Proceeding of Canadian Association for Information Science (CAIS) 2005 annual conference, Ontario, London*. Retrieved April 30, 2013, from http://www.cais-acsi.ca/proceedings/2005/park_J_2005.pdf.

[50] Peccia, Nestor et al. (1998). *CCSDS Standards – A Reference Model for an Open Archival Information System (OAIS)*. Retrieved February 15, 2013, from http://track.sfo.jaxa.jp/ spaceops98 /paper98/track3/3b002.pdf.

[51] PREMIS Editorial Committee. (2008). *PREMIS Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group*. Retrieved January 20, 2013, from http://www.loc.gov/standards/premis/v2/premis-2-0.pdf.

[52] PREMIS Editorial Committee. (2012). *PREMIS Data Dictionary for Preservation Metadata, version 2.2*. Retrieved March, 9, 2013, from http://www.loc.gov/ standards/premis/v2/premis-2-2.pdf.

[53] Public Record Office of Northern Ireland. (2006). *Preservation of Records*. Retrieved January, 20, 2013, from http://www.proni.gov.uk/preservation_of_records_-_proni_

preservation_policy.pdf.

[54] Sawyer, Don. (2000). The Open Archival Information System and the NSSDC. *National Space Science Data Center News*, *Dec 2011*. Retrieved April 30, 2013, from http://nssdc.gsfc.nasa.gov/ nssdc_news/dec00/oais.html.

[55] Schram, Jef et al. (2009) *MOREQ2*. Retrieved February 15, 2013, from http://www.dlmforum.eu/index.php?option=com_jotloader&section=files&task=download&cid=2_e8b4447991bbc46bdf47b4e098eab651&Itemid=10&lang=en.

[56] Shimazu, Keiko, Arisawa, Tatsya and Saito, Isao. (2006). Interdisciplinary contents management using 5W1H interface for metadata. In: *Proceeding of IEEE/WIC/ACM International Conference on Web Intelligence (WI'06)*, 909 – 912.

[57] Society of American Archivists. *Glossary of Archival and Records Terminology*. Retrieved April, 30, 2013, from http://www2.archivists.org/glossary/terms/a/archive.

[58] State Records of New South Wales. *Glossary of Recordkeeping Terms*. Retrieved March, 9, 2013, from http://www.records.nsw.gov.au/recordkeeping/government-recordkeeping-manual/introduction/glossary-of-recordkeeping-terms/glossary-of-recordkeeping-terms#r.

[59] State Records of New South Wales. *Recordkeeping systems - What is DIRKS?*. Retrieved January, 20, 2013, from http://www.records.nsw.gov.au/recordkeeping/dirks-manual/ introducing-the-dirks-methodology/recordkeeping-systems.

[60] State Records of New South Wales. *The DIRKS methodology and manual*. Retrieved March, 9, 2013, from https://www.records.nsw.gov.au/recordkeeping/dirks-manual/ introducing-the-dirks-methodology/dirks-methodology-and-manual.

[61] Stefan, Lucia. (2010). *Principles of Filing Service in View of the New MoREQ2 Standard*. Retrieved January 20, 2013, from http://www.slideshare.net/Stelucia/ecm-brno-lucia-stefan.

[62] Tennant, R. (2001). Different paths to interoperability. *Library Journal*, 126(3), 118-119.

[63] The Archive-Skills consultancy. (2003). *Info-shot: ISAD (G): The Basics*. Retrieved January, 20, 2013, from http://www.archive-skills.com/infobytes/ isad-g.pdf.

[64] The Cedars Project Team and UKOLN. (2000). *Metadata for Digital Preservation: The CEDARS Project Outline Specification*. Retrieved January, 20, 2013, from http://www.egov.vic.gov.au/pdfs/MD-STR~5.pdf.

[65] The Library of Congress. (2002). *Encoded Archival Description (EAD)*. Retrieved April, 30, 2013 from http://www.loc.gov/ead.

[66] The Library of Congress. *Metadata Encoding and Transmission Standard*. Retrieved

March, 9, 2013, from http://www.loc.gov/standards/mets.

[67] United States Environmental Protection Agency. *Frequent Questions about Recordkeeping Systems*. Retrieved April, 30, 2013 from http://www.epa.gov/records /faqs/rks.htm#q1.

[68] Woodley, Mary S. et al and DCMI. (2001). *Dublin Core Metadata Glossary*. Retrieved April 30, 2013, from http://dublincore.org/documents/2001/04/12/usageguide/glossary.shtml#A.

[69] Yale University. (2006). Using PREMIS to support preservation of digital assets at Yale. In: *Proceedings of the Yale University Integrated Access Council [2nd] Preservation Metadata Task Force 2006*. Retrieved March 9, 2013, from http://www.library.yale.edu/cataloging/meta data/pmtf/YalePREMISintro.pdf.

[70] Zeng, Marcia Lei. (1999). Matadata Elements for Object Description and Representation: A Case Report from a Digitized Historical Fashion Collection Project. *American Society for Information Science*, 50(13), 1193-1208.

List of Publications

[1] Jae Eun Bak, Shigeo Sugimoto, “Selection Guidelines for Preservation Method of Digital Resources”, Digital Library Workshop, 2007. (Japanese)

[2] Jae-Eun Baek, Shigeo Sugimoto, “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, Vol. 10 Issue. 2, 2010, pp. 75-100.

[3] Jae-Eun Baek, Shigeo Sugimoto, “Facet Analysis of Archival Metadata Standards to Support Appropriate Selection, Combination and Use of Metadata Schemas” , Proceedings of 11th International Conference on Dublin Core and Metadata Applications (DC-2011), The Hague, Netherland, September 2011, pp. 1-11.

[4] Jae-Eun Baek, Shigeo Sugimoto, “Analysis of Metadata Standards of Record Management for Metadata Interoperability - From the viewpoint of the Task model and 5W1H” , Korean Society of Archival Studies, The Korean Journal of Archival Studies, No. 32, 2012, pp. 127-176. (in Korean)

[5] Jae-Eun Baek, Shigeo Sugimoto, “A task-centric model for archival metadata schema mapping based on the records lifecycle”, International Journal of Metadata, Semantics and Ontologies archive, Vol. 7 No. 4, 2012, pp. 269-282.

Appendices

Appendix1. A Classification of Metadata Standards in the Records Lifecycle (ISAD(G) & Decision Tree) -- 94
Appendix2. A Classification of Metadata Standards in the Records Lifecycle (EAD & OAIS) --- 95
Appendix3. A Classification of Metadata Standards in the Records Lifecycle (AGLS & PREMIS) -- 96
Appendix4. A Classification and Mapping between Metadata Standards in the 5W1H Categories -- 97
Appendix5. A Classification and Mapping between Metadata Standards and 5W1H Categories in the Task Model -- 108

1) Appendix 1 ~ 3

The tables in these appendices show the classification of each metadata element, according to lifecycle stage. Each row of the table shows a metadata element and, each column shows the lifecycle stage.

The tables show the primary stages of metadata elements according to lifecycle stage. In addition, bold letters show the primary stage and, italic letters show the read stage and revised stage in the lifecycle.

2) Appendix 4

The table shows the classification of six metadata standards according to the 5W1H Categories and the relationship between metadata elements, and how they correspond with each of the 5W1H Categories.

In the table, each row shows a descriptive element from a metadata standard while the columns show the 5W1H Categories. We have indicated repeated metadata elements with a star mark (*) on the side of the element.

3) Appendix 5

The table shows the mapping of metadata elements in each stage of the Task model (task group). In addition, the table shows the relationship between metadata elements, and how they correspond with each category.

In the table, each row shows a descriptive element from a metadata standard and the columns show the 5W1H Categories in each task group of the Task model. Bold letters shows the primary stage and letters with a star mark (*) show the read stage and revised stage of the task group.

Appendix1. A Classification of Metadata Standards in the Records Lifecycle (ISAD(G) & Decision Tree)

Metadata Lifecycle	**ISAD(G)**	**Decision Tree**
Create	Date(s) Language, scripts of material Level of description Name of creator(s) Title	
Use & Manage	Date(s) *Level of description* Scope and content	
Appraisal & Destory	Appraisal, destruction and scheduling information Archivist`s note *Date(s)* Date(s) of description Immediate source of acquisition or transfer Physical characteristics and technical requirements Rule or conventions	Long term value justify preservation Other purposes Multiple media formats Digital version be selected for preservation Documentation been supplied Negotiate for the source to supply Technically feasible for you to construct Material so valuable that you will accept Accept the costs and risks of trying to manage Cost effective for you to develop Cost-effective for you to transfer Ccept the costs and risks of trying to manage Commit adequate staff Manageable file format Technically feasible for you to transfer the material Available to you online or on a physical carrier Able to collect or receive the resource via a Enough available storage space Carrier that is acceptable for transfer and/or storage Transfer the resource to an acceptable carrier
Store & Arrange	Accruals Administrative, Biographical history Archival history *Archivist`s note* Conditions governing access Conditions governing reproduction *Date(s)* *Date(s) of description* Existence and location of copies Existence and location of originals *Extent and medium of the unit of description* Findings aids *Language, scripts of material* *Level of description* Note Publication note Related units of description Reference code *Rule or conventions* System of arrangement	Institutional remit/collection development policy Preservation responsibility Preservation responsibility been accepted elsewhere Higher degree of preservation commitment or access Acceptable arrangements for acquisition and/or transfer Re-evaluate acquisition The rights to transfer *Technically feasible for you to transfer the material* *Available to you online or on a physical* *Enough available storage space* *Documentation been supplied* *Negotiate for the source to supply* *Cost effective for you to develop*
Preserve	*Accurals* *Archival history* *Archivist`s note* *Date(s)* *Date(s) of description* *Level of description* *Note* *Reference code* *Rule or conventions*	
Reference & Re-use	*Date(s)* *Level of description* *Scope and content*	

Appendix2. A Classification of Metadata Standards in the Records Lifecycle (EAD & OAIS)

Metadata Lifecycle	EAD	OAIS
Create	Address, Addressline, Creation, Corpname Date, Famname, Geogname, Imprint Language, Language, Name, Namegrp Origination, Otherfidaid, P, Persname Publicationstm, Publisher, Ptr, Sponsor Subtitle, Title, Titlepage, Unitdate Unittitle	Reason for Creation
Use & Manage	*Address*, *Addressline*, *Corpname*, *Date* *Dimensions*, Extent, Extref, Extrefloc *Famname*, Genreform, *Geogname*, *Imprint* *Language*, *Langusage*, *Name*, *Namegrp* Occupation, *Otherfidaid*, *P*, *Persname* Prefercite, *Publicationstm*, *Publisher*, Physdesc *Ptr*, Scopeconten, *Sponsor*, *Subtitle* *Unitdate*, *Origination*	Existing Metadata Existing Records
Appraisal & Destory	Abstract, Accruals, Acqinfo, *Address* *Addressline*, Appraisal, Archdesc, *Archref* Author, C, C01 - c12, Container *Corpname*, *Date*, Descrules, *Famname* Frontmatter, *Geogname*, *Imprint*, *Language* *Langusage*, *Name*, *Namegrp*, Note *P*, *Persname*, Processinfo, *Ptr* Repository, *Sponsor*, *Subtitle*, *Unitdate*	Context Information, Custody History Documentation, *Existing Metada* History of Origin, Legislation Text Pointer Permitted by License Original Technical Environments Prerequisites, Procedures *Reason for Creation*, Reason for Preservation Related Information Objects Representation Information
Store & Arrange	*Abstract*, Accessrestrict, *Accruals*, *Address* *Addressline*, Altformavail, *Appraisal*, Archdesc *Archref*, Arrangement, *Author*, Bibliography Bibref, Bibseries, Bioghist, *C* *C01 - C12*, Chronitem, Chronlist, *Container* *Corpname*, Custodhist, Dao, Daodesc Daogrp, Daoloc, *Date*, *Descrules* *Dimensions*, Event, Eventgrp, *Extent* Extptr, Extptrloc, *Extref*, *Extrefloc* *Famname*, *Frontmatter*, Genreform, *Geogname* *Imprint*, Langmaterial, *Langusage*, Legalstatus Materialspec, *Name*, *Note*, *Occupation* Originalsloc, *Otherfindaid*, *P*, *Persname* *Physdesc*, Physfacet, Physloc, Phystech *Prefercite*, *Processinfo*, *Ptr*, *Publisher* *Publicationstm*, *Scopecontent*, *Sponsor*, *Language* Separatedmaterial, *Repository*, *Namegrp* Relatedmateri, Ref, Refloc *Unitdate*, Userestrict, Subject	Actions, Action History Actors, Administration History Change History Before Archivi, Contacts or Rights Holders *Context Information*, Copyright Statement *Custody History*, Date of Publication *Existing Metadata*, *Existing Records* Fixity Information, *History of Origin* Ingest Process History, *Legislation Text Pointer* Licence Text Pointer, Management History Name of Publisher, Negotiation History *Original Technical Environmen*, Permitted by License Place of Publication, *Prerequisites* *Procedures* Provenance Information *Reason for Creation* Reference Information *Related Information Objec* *Representation Informatic* Resource Description Rights Information Rights Management
Preserve	*Address*, *Archdesc*, *Archref*, *Author* *Bibliography*, *Bibref*, *C*, *C01 - C12* *Chronlist*, *Chronitem*, *Container*, *Corpname* *Custodhist*, *Date*, *Dao*, *Daodesc* *Daogrp*, *Daoloc*, *Descrules*, *Event* *Eventgrp*, Extref, *Extrefloc*, *Frontmatter* *Famname*, *Geogname*, *Imprint*, *Language* *Langusage*, *Materialspec*, *Name*, *Namegrp* *Note*, *Occupation*, *Persname*, *Processinfo* *Phystech*, *P*, *Ptr*, *Ref* *Refloc*, *Repository*, *Sponsor*, *Subject* *Unitdate*, *Origination*	*Actions*, *Action History* *Actors*, *Administration History* Authentication Indicator *Change History Before Archiving* *Contacts or Rights Holders*, Content Information *Context Information*, *Copyright Statement* *Custody History*, *Date of Publication* *Existing Metadata*, *Existing Records* *Fixity Information*, *History of Origin* *Ingest Process History*, Input Format *Legislation Text Pointer*, *Licence Text Pointer* *Management History*, *Name of Publisher* *Negotiation History*, Output Format Parameters, Permitted by Statute *Place of Publication*, Platform Policy History Preservation Description Information *Provenance Information*, *Reason for Preservation* *Reference Information*, *Related Information Objects* Render/Analyse Engines, *Representation Information* *Resource Description*, *Rights Information* *Rights Management*, Rights Warning Structure Information, Transformer Objects (TOs) Underlying Abstract Form Description
Reference & Re-use	*Address*, *Corpname*, *Date* *Famname*, *Geogname*, *Name* *Namegrp*, *Persname*, *Prefercite* *Sponsor*, *Unitdate*, Extref *Language*, *Imprint*, *Extrefloc* *Langusage*, *Occupation*, *P* *Ptr*, *Ref*, *Refloc*	*Actions*, *Actors* *Contacts or Rights Holders*, *Existing Metadata* *Input Format*, *Legislation Text Pointer* *Licence Text Pointer*, *Negotiation History* *Output Format*, *Parameters* *Permitted by License*, *Permitted by Statute* *Platform*, *Render/Analyse Engines* *Rights Information*, *Rights Management* Rights warning

Appendix3. A Classification of Metadata Standards in the Records Lifecycle (AGLS & PREMIS)

Metadata Lifecycle	**AGLS Metadata**	**PREMIS**
Create	Contributor Creator Date Format Identifier Language Publisher Rights Title	ObjectCharacteristics CreatingApplication Environment OriginalName
Use & Manage	Audience Availability Coverage *Date* Description *Format* Funcion *Identifier* Mandate Relation *Rights* Source Subject Type	*ObjectCharacteristics* *Environment* Relationship LinkingEventIdentifier LinkingIntellectualEntityIdentifier LinkingRightsStatementIdentifier EventType EventDateTime EventDetail LinkingAgentIdentifier LinkingObjectIdentifier CopyrightInformation LicenseInformation StatuteInformation LinkingObjectIdentifier LinkingAgentIdentifier
Appraisal & Destory	*Availability* *Date* *Rights*	
Store & Arrange	*Date* *Description* *Funcion* *Format* *Mandate* *Mandate* *Relation* *Rights* *Subject* *Type*	*ObjectCharacteristics* *OriginalName* Storage *Environment* *Relationship* *LinkingEventIdentifier* *LinkingIntellectualEntityIdentifier* *LinkingRightsStatementIdentifier* *EventDateTime* *LinkingAgentIdentifier* *LinkingObjectIdentifier* *CopyrightInformation* *LicenseInformation* *LinkingObjectIdentifier* *LinkingAgentIdentifier*
Preserve	*Date* *Description* *Format* *Identifier* *Mandate* *Relation* *Rights* *Type*	ObjectIdentifier ObjectCategory PreservationLevel SignificantProperties *ObjectCharacteristics* *OriginalName* *Storage* *Environment* SignatureInformation *Relationship* *LinkingEventIdentifier* *LinkingIntellectualEntityIdentifier* *LinkingRightsStatementIdentifier* EventIdentifier *EventType* *EventDateTime* *EventDetail* EventOutcomeInformation *LinkingAgentIdentifier* *LinkingObjectIdentifier* AgentIdentifier AgentName AgentType RightsStatement RightsStatementIdentifier RightsBasis *CopyrightInformation* *LicenseInformation* *StatuteInformation* RightsGranted LinkingObjectIdentifier *LinkingAgentIdentifier* RightsExtension
Reference & Re-use	*Availability* *Audience* *Date* *Format* *Function* *Identifier* *Rights* *Source* *Relation* *Type*	*ObjectCharacteristics* *Environment* *EventDateTime* *EventDetail* *LinkingAgentIdentifier*

Appendix4. A Classification and Mapping between Metadata Standards in the 5W1H Categories

5W1H model	AGLS	AGRkMS	Dicision Tree of DPC	OAIS	EAD	PREMIS
Who	Creator			Resource Description	Author Creation	
	Contributor					
	Publisher			Name of Publisher	Publication Statement	
					Publisher	
	Audience			Actors		
					Sponsor	
						signer
						messageDigestOriginator
				Contacts or Rights Holders		
		Position				
					Origination	
					Imprint	
					Subject	
					Corporate Name	
					Family Name	
					Personal Name	
					Name	
					Name Group	
					Abbreviation	
					Expansion	
					Emphasis	
					Item	
					Profile Description	
When	Date	Date Range Start Date			Date Date of the unit	dateCreatedByApplication
				Date of Publication		
				Change History Before Archiving		
						preservationLevelDateAssigned
						eventDateTime
						copyrightStatusDeterminationDate
						termOfGrant
						startDate
						endDate
					Chronology List Chronology List Item	
		End Date				
					EventEvent Group	
					Item*	
					Profile Description*	
					Imprint*	
Where	Creator*			Resource Description*	Author*	
	Contributor *				Creation*	
	Publisher			Name of Publisher *		
				Place of Publication		
					Publication Statement*	
					Publisher*	
	Audience *			Actors*		
					Sponsor*	
						signer*
		Identifier Scheme				
		Position*				
						messageDigestOriginator
		Location				contentLocationVal

Where						ue
						storage
						copyrightJurisdictio n
						statuteJurisdiction
						statuteInformationD eterminationDate
				History of Origin	Location of Originals	
				Custody History		
				Change History Before Archiving*		
				Contacts or Rights Holders *		
					Subordinate Area	
					Repository	
					Origination*	
					Imprint*	
					Subject*	
					Corporate Name*	
					Geographic Name	
					Name*	
					Name Group*	
					Address	
					Address Line	
					Abbreviation*	
					Expansion*	
					Emphasis *	
					Item*	
					Physical Location	
					Profile Description*	
What					Abbreviation *	
					Expansion*	
					Creation*	
					File Description	
					Series Statement	
					Origination*	
					Physical Location*	
	Function	Keyword				
	Subject					
	Title	Name Words		Resource Description*	Title	
					Subtitle	
					Title of the Unit	
					Subject*	originalName
		Name Scheme				
					Corporate Name*	
					Geographic Name	
					Family Name*	
					Personal Name*	
					Name*	
					Name Group*	
					Component	
					Component (1 Level) ~(12)	
					Physical Description	
					Dimensions	
		Extent			Extent	
		Medium			Scope and Content	storageMedium
	Format		Multiple media formats			
		Format			Digital Archival Object	
					Digital Archival Object Description	

What					Digital Archival Object Group	
					Digital Archival Object Location	
				Administration History	Administrative Information	
				Custody History *		
				Change History Before Archiving *		
				Management History		
				Ingest Process History	Processing Information	
					Biography or History	
					Subordinate Area*	
					Custodial History	
					Acquisition Information	
					Function	
					Appraisal Information	
					Accruals	
					Arrangement	
				Context Information		
				Related Information Objects		
					Conditions Governing Access	
					Legal Status	
					Restrictions on Use	
					Conditions Governing Use	
					Language of the Material	
					Language	
	Language	Language				
	Type	Category			Genre/Physical Characteristic	significantPropertiesType
					Material Specific Details	
					Physical Facet	
					Physical Characteristics and Technical Requirements	
					Edition	
					Edition Statement	
					Index	
					Index Entry	
					Note	
					Note Statement	
					Language Usage	
					Other Finding Aid	
					Profile Description*	
					Reference	
					Reference Location	
					Related Material	
					Title Proper of the Finding Aid	
	Source			Provenance Information		
				History of Origin*	Location of Originals*	
					Alternative Form Available	
	Relation					

What		Related Entity Assigned Entity ID Assigned Entity ID Scheme Relationship Role				
					Bibliography	
					Bibliographic Reference	
					Bibliographic Series	
					ID of the Unit	
						objectIdentifier
	Identifier	Identifier				
		Identifier String				objectIdentifierType
		Identifier Scheme*				objectIdentifierValue
	Availability					
	Description	Description		Reference Information	Archival Reference	
					Text Division	
				Resource Description*		
					Abstract	
					Archival Description	
					Archival Description Group	
					Organization	
	Contributor					
	Coverage	Coverage				
			Acceptable arrangements for acquisition and/or transfer			
			Re-evaluate acquisition			
		permissions				act
	Mandate		Institutional remit/collection development policy			
						preservationLevel
						preservationLevelRole
						restriction
		Caveat Category				
	Rights	Security Caveat				
		Caveat text				
		Security Classification				
		Rights	The rights to transfer			
			Acquisition, Preservation responsibility			
			Preservation responsibility (been accepted elsewhere)			
			Higher degree of preservation commitment or access			
				Rights Management		

What				Negotiation History		
				Rights Information		
				Copyright Statement		
				Rights Warning		
				Permitted by Statute		
				Legislation Text Pointer		
				Permitted by License		
						licenseTerms
				Licence Text Pointer		
						rightsBasis
		Rights Statement				
						licenseInformation
						copyrightInformation
						copyrightStatus
						licenseNote
						copyrightNote
						statuteNote
					Descriptive Rules	
		Rights Type				
		Rights Status				
						statuteInformation
						statuteNote
		Jurisdiction				
						significantProperties
		Disposal				
			Material so valuable that you will accept			
			Long term value justify preservation			
			Acquire for other purposes			
			Accept the costs and risks of trying to manage			
			Cost effective for you to develop			
			Cost-effective for you to transfer			
			Accept the costs and risks of trying to manage			
			Commit adequate staff			
		Contact		Contacts or Rights Holders *		
		Position*				
				Actions		
				Content Information		
					Change	
					Chronology List*	
					Chronology List Item*	
					Container	
					EAD Identifier	
					Emphasis*	
					Event*	
					Event Group*	
					File Plan	
					Front Matter	

What					Heading First Heading Second Heading	
					Item*	
					Number	
					Occupation	
					Other Descriptive Data	
					Paragraph	
					Personal Name*	
					Resource	
					Preferred Citation	
					Revision Description	
					Separated Material	
					Spanned Column Specification	
					Title Statement	
					Title Page	
How					Abbreviation *	
					Expansion*	
					Component * Component (1) ~(12) *	
					Geographic Name*	
					Name*	
					Name Group*	
						agentNote
						significantProperties*
						significantPropertiesValue
					Physical Characteristics and Technical Requirements*	significantPropertiesType*
					Controlled Access Headings	
					Index*	
					Index Entry*	
					Note*	
					Note Statement*	
					Language Usage*	
					Other Finding Aid*	
					Pointer	
					Pointer Group	
					Pointer Location	
					Reference*	
					Reference Location*	
						storage*
						contentLocationType
						contentLocationValue*
				Administration History *		
				Management History *		
				Ingest Process History *		
				Action History		
				Policy History		
						objectIdentifier*
		Identifier*				statuteCitation
						linkingEventIdentifier
						linkingEventIdentifierType
						linkingEventIdentifi

How	Identifier*					erValue
						linkingRightsStatementIdentifier
						linkingAgentIdentifier
						linkingObjectIdentifier
						rightsStatementIdentifier
						agentIdentifier
						eventIdentifier
						objectIdentifierValue*
						relatedEventIdentifierValue
						eventIdentifierValue
						agentIdentifierValue
		Identifier String*				
		Identifier Scheme*				objectIdentifierType*
						relatedEventIdentifierType
						eventIdentifierType
						agentIdentifierType
						agentName
						licenseIdentifier
						licenseIdentifierType
						licenseIdentifierValue
		Name*				dependencyIdentifier
		Name Words*				
		Name Scheme*				
		Keyword*				
	Availability *					
						preservationLevelValue*
						preservationLevelRole*
						storageMedium*
		Extent*			Extent*	Size
		Format*			Table Column Specification	
					Digital Archival Object*	
					Digital Archival Object Description*	
					Digital Archival Object Group*	
					Digital Archival Object Location*	
			Multiple media formats *			format
			Digital version be selected for preservation			
			Manageable file format			
			Carrier that is acceptable for transfer and/or storage			
						formatDesignation
						formatName
						formatVersion

How						formatRegistryName
						creatingApplicationVersion
						creatingApplicationName
						formatNote
						creatingApplication
						formatRegistry
						formatRegistryKey
						formatRegistryRole
						software
						swName
						swVersion
						swType
						swOtherInformation
						swDependency
						hardware
						hwName
						hwType
						hwOtherInformation
		Document Form				
						linkingEventIdentifier
						linkingEventIdentifierType
						linkingEventIdentifierValue
						objectCharacteristics
						compositionLevel
						inhibitors
						inhibitorType
						inhibitorTarget
						inhibitorKey
						signatureInformation
						signature
						signatureEncoding
						signer*
						signatureMethod
						signatureValue
						signatureValidationRules
						signatureProperties
						keyInformation
				Underlying Abstract Form Description		
				Transformer Objects (TOs)		
				1) Platform		
				2) Parameters		
				3) Render/Analyse Engines		
				4) Output Format		
				5) Input Format		
				Render/Analyse/Convert Objects		
				Semantic Information		
				Render/Analyse Objects (RAO)		
				Data Object		
				Original Technical Environments		environment environmentPurpose
				Prerequisites		

				Procedures		environmentNote
				Documentation		
				Fixity Information		fixity
						messageDigestOriginator
				Authentication Indicator		
						dependency
						dependencyName
						dependencyIdentifierType
						dependencyIdentifierValue
		permissions*				act*
	Mandate *		Negotiate for the source to supply			
						preservationLevel*
						restriction*
						relationshipType
		Category*				eventType
						agentType
						relationship
	Relation *					relatedEventIdentification
						relatedEventSequence
						linkingAgentIdentifierType
		Related Entity* Assigned Entity ID* Assigned Entity ID Scheme* Relationship Role*				relatedObjectIdentifierValue
						relatedObjectSequence
How						linkingAgentIdentifierValue
						linkingAgentRole
				Rights Management*		
				Negotiation History *		
				Rights Information *		
				Copyright Statement*		
				Rights Warning *		
				Permitted by Statute *		
				Permitted by License*		
		Rights*				linkingRightsStatementIdentifierType
	Rights *					linkingRightsStatementIdentifierValue
						linkingRightsStatementIdentifier
						rightsBasis*
						copyrightInformation*
						copyrightStatus*
						copyrightNote*
						licenseInformation *
						licenseTerms*
						licenseNote *
		Rights Statement*				
				Legislation Text Pointer *		
				Licence Text Pointer *		
						rightsGrantedNote
					Descriptive Rules *	
		Rights Type*				
		Rights				

How		Status*				
		Jurisdiction*				
						statuteInformation*
						statuteNote*
						rightsStatementIdentifierType
						rightsStatementIdentifierValue
	Source *			Provenance Information *		
				History of Origin*		
				Context Information*		
						objectCharacteristicsExtension
						significantPropertiesExtension
						eventOutcomeDetailExtension
						creatingApplicationExtension
						environmentExtension
						signatureInformationExtension
						agentExtension
						rightsExtension
					Extended Pointer Extended Pointer Location Extended Reference Extended Reference Location	linkingIntellectualEntityIdentifierType
						linkingIntellectualEntityIdentifierValue
						linkingObjectIdentifierTyp
		Disposal*				
		Integrity Check				messageDigestAlgorithm
						messageDigest
		Precedence				
	Description*	Description*				eventDetail
						eventOutcomeInformation
						eventOutcome
						eventOutcomeDetail
		Change History Property Name Prior Value Relationship ID				eventOutcomeDetailNote
						linkingObjectIdentifierRole
				Related Information Objects*	Archival Reference *	
				Existing Metadata	Resource Description*	
				Existing Records		
				Resource Description*		
						objectCategory
						linkingObjectIdentifierValue
			Documentation been supplied (including metadata)			
			Technically feasible for you to construct			
			Technically feasible for you to transfer the material			

How			Available to you online or on a physical carrier			
			Able to collect or receive the resource via			
			Enough available storage space			
			Transfer the resource to an acceptable carrier			
				Contacts or Rights Holders *		
				Actions * Content Information *		
				Representation Information		
				Structure Information		
					Change*	
					EAD Identifier*	
					Emphasis*	
					File Plan*	
					Front Matter*	
					Heading* First Heading* Second Heading*	
					Item*	
					Number*	
					Other Descriptive Data*	
					Resource*	
					Revision Description*	
					Spanned Column Specification *	
Why	Description*			Reason for Creation		
				Reason for Preservation		
			Long term value justify preservation*			
			Acquire for other purposes*			
		Document Form*				
				Provenance Information *		
						preservationLevelRationale
						formatRegistryRole*

Appendix5. A Classification and Mapping between Metadata Standards and 5W1H Categories in the Task Model

TASK Model	5W1H model	AGLS	AGRkMS	Dicision Tree of DPC	OAIS	EAD	PREMIS
T1	Who	Creator					
		Contributor					
		Publisher				Publication Statement	
						Publisher	
						Sponsor	
						Abbreviation	
						Emphasis	
						Expansion	
						Corporate Name	
						Family Name	
						Imprint	
						Item	
						Name	
						Name Group	
						Origination	
						Personal Name	
						Subject	
	When	Dates	Date Range Start Date			Date of the Unit	
						Publication Statement*	
						Date	dateCreatedBy Application
			End Date			Imprint*	
						Item*	
	Where	Creator*					
		Contributor*					
			Location				
						Sponsor*	
		Publisher*				Publication Statement*	
						Publisher*	
						Abbreviation*	
						Emphasis*	
						Expansion*	
						Address	
						Address Line	
						Corporate Name*	
						Geographic Name	
						Imprint*	
						Item*	
						Name*	
						Name Group*	
						Origination*	
						Subject*	
	What	Title	Name Name Words			Subject*	
						Title	
						Title of the Unit	
						Subtitle	
		Language	Language				
						Language of the Material	
						Language	
		Coverage					
		Description				Abstract	

T1	What	Mandate					
		Rights					
						Dimensions	
			Extent				
						Material Specific Details	
						Administrative Information	
						Language Usage	
						File Description	
						Item*	
						Abbreviation*	
						Emphasis*	
						Expansion*	
						Geographic Name*	
						Name*	
						Name Group*	
						Number	
						Origination*	
						Personal Name*	
						Series Statement	
						Title Page	
						Title Proper of the Finding Aid	
						Title Statement	
	How	Description*					
			Extent*				
						Language of the Material*	
						Table Column Specification	
		Mandate*					
		Rights*					creatingApplicationExtension
							creatingApplication (name, version)
						Language Usage*	
						Abbreviation*	
						Emphasis*	
						Expansion*	
						Geographic Name*	
						Item*	
						Name*	
						Name Group*	
						Number*	
	Why	Description*			Reason of creation		
T2	Who	Audience			Actors *		
		Contributor*					
			Position				
						Sponsor*	
						Publisher*	
						Publication Statement*	
						Abbreviation*	
						Emphasis*	
						Expansion*	
						Corporate Name*	
						Family Name*	
						Imprint*	
						Item*	
						Name*	
						Name Group*	
						Origination*	
						Personal Name*	
						Subject*	
	When	Date*	Date Range* Start Date*			Date*	
						Publication	

T2	When					Statement*	
						Date of the Unit *	
			End Date*				
						Imprint*	
						Item*	
	Where	Contributor*					
					Actors *		
			Position*				
			Location*				
			Identifier Scheme				
						Sponsor*	
						Administrative Information*	
						Publication Statement*	
						Publisher*	
						Repository	
						Imprint*	
						Abbreviation*	
						Emphasis*	
						Expansion*	
						Address*	
						Address Line*	
						Corporate Name*	
						Geographic Name*	
						Item*	
						Name*	
						Name Group*	
						Origination*	
						Subject*	
	What		Identifier				
			Identifier Scheme*				
			Name Scheme				
			Language				
						Acquisition Information	
		Format	Format				
						Scope and Content	
						Dimensions*	
			Extent				
						Physical Description	
		Type	Category			Genre/Physical Characteristic	
						Physical Facet	
		Relation					
			Related Entity				
						Bibliography	
						Bibliographic Reference	
						Bibliographic Series	
		Availability					
		Description	Description				
					Resource Description	Abstract*	
		Function	Keyword				
		Subject					
		Coverage	Coverage				
			permissions				
		Mandate*					
		Rights*			Permitted by Statute		
			Rights				
			Security Caveat				
			Caveat text				
			Security Classification				
			Caveat Category				
			Jurisdiction				
		Source					
					Actions		
			Contact				

T2	What		Position*				
			Disposal				
						Paragraph	
						File Description*	
						Name*	
						Name Group*	
						Number*	
						Item*	
						Abbreviation*	
						Emphasis*	
						Expansion*	
						Geographic Name*	
						Origination*	
						Personal Name*	
						Preferred Citation	
						Separated Material	
						Series Statement*	
						Subject*	
						Title Page*	
						Title Proper of the Finding Aid*	
						Title Statement*	
	How		Identifier*				
			Identifier Scheme*				
			Name Scheme*				
		Format*	Format*			Table Column Specification*	
			Extent*				
			Document Form				
			permissions*				
		Mandate*					
		Relation*					
			Related Entity*				
			Category*				
		Rights*	Rights*		Permitted by Statute *		
		Source*					
					Actions *		
			Description*				
			Change History				
			Jurisdiction*				
			Keyword*				
			Disposal*				
							environment
							software
							swName
							swVersion
							swType
							swOtherInformation
							swDependency
							hardware
							hwName
							hwType
							hwOtherInformation
						Name*	
						Name Group*	
						Item*	
						Abbreviation*	
						Emphasis*	
						Expansion*	
						Geographic Name*	
						Number*	
	Why		Document Form*				
T3	Who		Position*				
						Sponsor*	
						Publisher*	
						Publication Statement*	
						Author	

T3	Who					Abbreviation*	
						Emphasis*	
						Expansion*	
						Corporate Name*	
						Family Name*	
						Imprint*	
						Item*	
						Name*	
						Name Group*	
						Origination*	
						Personal Name*	
						Subject*	
	when	Date*	Date Range* Start Date*			Date*	
						Publication Statement*	
						Date of the Unit *	
			End Date*				
						Imprint*	
						Item*	
	Where		Identifier Scheme*				
			Position*				
						Sponsor*	
						Author	
						Publisher*	
						Publication Statement*	
						Name*	
						Name Group*	
						Abbreviation*	
						Emphasis*	
						Expansion*	
						Corporate Name*	
						Geographic Name*	
						Imprint*	
						Item*	
						Origination*	
						Subject*	
	What		permissions*				
		Mandate*					
		Rights*	Security Caveat*				
			Rights*				
				Long term value justify preservation			
				Acquire for other purposes			
				Accept the costs and risks of trying to manage			
				Commit adequate staff			
				Material so valuable that you will accept			
					Ingest Process History	Processing Information	
						Appraisal Information	
			Identifier*				
			Identifier Scheme*				
			Jurisdiction*				
			Contact*				
			Position*				
			Keyword*				
			Disposal*				
		Format*	Format*				

T3	What					File Description*	
						Name*	
						Name Group*	
						Series Statement*	
						Abbreviation*	
						Emphasis*	
						Expansion*	
						Geographic Name*	
						Item*	
						Number*	
						Origination*	
						Personal Name*	
						Revision Description	
						Subject*	
						Title Page*	
						Title Proper of the Finding Aid*	
						Title Statement*	
	How		permissions*				
		Mandate*		Negotiate for the source to supply			
		Rights*	Rights*				
		Format*	Format*	Digital version be selected for preservation			
						Table Column Specification*	
				Manageable file format			
				Carrier that is acceptable for transfer and/ or storage			
			Document Form				
				Technically feasible for you to transfer the material			
				Available to you online or on a physical carrier			
				Able to collect or receive the resource via			
				Enough available storage space			
				Transfer the resource to an acceptable carrier			
				Documentation been supplied (including metadata)			
				Technically feasible for you to construct			
					Ingest Process History *		
			Identifier*				
			Identifier Scheme*				
			Change History*				
			Jurisdiction*				
			Keyword*				
			Disposal*				
			Integrity Check				
						Name* Name Group*	
						Abbreviation*	

T3	How					Emphasis*	
						Expansion*	
						Geographic Name	
						Item*	
						Number*	
						Revision Description*	
	Why			Long term value justify preservation*			
				Acquire for other purposes*			
			Document Form*				
T4	Who				Name of Publisher		
					Contacts or Rights Holders		
					Actors		
						Publication Statement*	
						Publisher*	
						Author*	
						Sponsor*	
			Position*				
						Abbreviation*	
						Emphasis*	
						Expansion*	
						Corporate Name*	
						Family Name*	
						Imprint*	
						Item*	
						Name*	
						Name Group*	
						Origination*	
						Personal Name*	
						Profile Description	
						Subject*	
	When	Date*	Date Range* Start Date*		Change History Before Archiving	Date*	
					Date of Publication *		
							dateCreatedBy Application*
						Publication Statement*	
						Chronology List	
						Chronology List Item	
						Date of the Unit *	
			End Date*				
						Event Event Group	
						Imprint*	
						Item*	
						Profile Description*	
	Where				History of Origin	Location of Originals	
					Custody History		
					Change History Before Archiving		
					Name of Publisher *	Publisher*	
					Place of Publication	Publication Statement*	
					Contacts or Rights Holders *		
			Location*				

						Sponsor*	
						Repository*	
						Physical Location	
					Actors *		
						Author*	
							storage
			Identifier Scheme*				
			Position*				
						Name*	
						Name Group*	
						Abbreviation*	
	Where					Emphasis*	
						Expansion*	
						Address*	
						Address Line*	
						Corporate Name*	
						Geographic Name*	
						Imprint*	
						Item*	
						Origination*	
						Profile Description*	
						Subordinate Area*	
						Subject*	
						ID of the Unit	
		Identifier*	Identifier*				
			Identifier Scheme*				
						Subtitle*	
		Description*	Description*		Reference Information	Archival Reference	
						Text Division	
					Resource Description*		
						Abstract*	
						Archival Description	
						Archival Description Group	
						Organization	
		Function*	Keyword*				
		Subject*					
						Component Component (1 Level) ~(12)	
						Physical Description*	
		Format*	Extent*			Extent	
T4						Dimensions*	
			Medium*			Scope and Content*	storageMedium
	What		Format*	Multiple media formats			
						Digital Archival Object	
						Digital Archival Object Description	
						Digital Archival Object Group	
						Digital Archival Object Location	
					Custody History *		
					Change History Before Archiving *		
					Ingest Process History	Processing Information*	
					Administration History	Administrative Information*	
					Management History		
						Subordinate Area*	
						Biography or History	

						Custodial History	
						Acquisition Information*	
						Function	
						Appraisal Information*	
						Accruals	
						Arrangement	
					Context Information		
						Conditions Governing Access	
						Legal Status	
					Related Information Objects		
						Conditions Governing Use	
			Language				
						Language of the Material*	
						Language*	
		Type*	Category*			Genre/Physical Characteristic	
						Material Specific Details*	
						Physical Characteristics and Technical Requirements	
						Physical Facet*	
						Edition	
T4	**What**					Edition Statement	
						Note	
						Note Statement	
						Other Finding Aid	
						Profile Description*	
						Reference	
						Reference Location	
						Related Material	
						Title Proper of the Finding Aid*	
						Language Usage*	
					History of Origin*	Location of Originals*	
					Provenance Information		
						Alternative Form Available	
		Relation*	Category*				
			Related Entity*				
						Bibliography*	
						Bibliographic Reference*	
						Bibliographic Series*	
		Rights				Descriptive Rules	
			Rights	Acquisition, Preservation responsibility			
				Preservation responsibility been accepted elsewhere			
				Higher degree of preservation commitment or access			
				The rights to transfer			
					Rights Management		
					Negotiation		

T4	**What**				History		
					Rights Information		
					Copyright Statement		
					Rights Warning		
					Permitted by Statute		
					Permitted by License		
					Legislation Text Pointer		
					Licence Text Pointer		
			Security Classification				
			Security Caveat				
			Jurisdiction				
		Coverage	Coverage				
			permissions				
		Mandate		Institutional remit/collection development policy			
				Acceptable arrangements for acquisition and/or transfer			
				Re-evaluate acquisition			
				Cost effective for you to develop			
				Cost-effective for you to transfer			
				Commit adequate staff			
						Physical Location*	
						Index	
						Index Entry	
						Event*	
						Event Group*	
						Chronology List*	
						Chronology List Item*	
					Contacts or Rights Holders *		
					Actions		
					Content Information		
						Heading First Heading Second Heading	
			Name Scheme				
			Position*				
			Disposal				
							significantProperties
						File Description*	
						Change	
						Container	
						EAD Identifier	
						File Plan	
						Front Matter	
						Name* Name Group*	
						Abbreviation*	
						Emphasis*	
						Expansion*	
						Geographic Name*	

T4						Item*	
						Number*	
						Occupation	
						Other Descriptive Data	
						Origination*	
						Paragraph*	
						Personal Name*	
						Preferred Citation*	
						Resource	
						Revision Description*	
						Separated Material*	
						Series Statement*	
						Spanned Column Specification	
						Subject*	
						Title Page*	
						Title Statement*	
	How	Identifier*	Identifier*				
			Identifier Scheme*				
		Description*	Description*		Reference Information *		
			Resource Description*		Resource Description*		
			Change History*		Related Information Objects *	Archival Reference*	
					Existing Metadata		
					Existing Records		
					Context Information *		
						Component* Component (1) ~(12)*	
						Language of the Material*	
						Physical Characteristics and Technical Requirements*	
						Controlled Access Headings	
						Other Finding Aid*	
						Language Usage*	
						Reference*	
						Reference Location*	
						Pointer	
						Pointer Group	
						Pointer Location	
						Index *	
						Index Entry *	
		Rights*				Descriptive Rules*	
			Rights*		Rights Management *		
					Negotiation History *		
					Rights Information		
					Copyright Statement*		
					Rights Warning *		
					Permitted by Statute *		
					Permitted by License*		
					Legislation Text Pointer *		
					Licence Text Pointer*		

T4	How		Jurisdiction*				
		Format*				Table Column Specification*	
			Extent*			Extent*	
			Format*				formatNote
				Multiple media formats *			format
							creatingApplicationVersion*
							creatingApplicationName*
							storageMedium*
						Digital Archival Object*	
						Digital Archival Object Description*	
						Digital Archival Object Group*	
						Digital Archival Object Location*	
			DocumentForm *				
						Extended Pointer	
						Extended Pointer Location	
						Extended Reference	
						Extended Reference Location	
							creatingApplicationExtension*
							objectCharacteristicsExtension
							inhibitors
							inhibitorType
							inhibitorTarget
							inhibitorKey
			permissions*				
		Mandate*					
			Category*				
		Relation*					
			Related Entity*				
						Note*	
						Note Statement*	
					Provenance Information *		
					History of Origin*		
					Original Technical Environments		environment* environmentPurpose environmentNote
					Prerequisites		
					Procedures		
					Documentation		
					Management History *		
					Ingest Process History *		
					Administration History *		
					Action History		
					Policy History		
					Contacts or Rights Holders *		
					Actions*		
					Fixity Information		
					Authentication Indicator		

					Content Information *		
					Underlying Abstract Form Description		
					Transformer Objects (TOs)		
					1) Platform		
					2) Parameters		
					3) Render/Analyse Engines		
					4) Output Format		
					5) Input Format		
					Render/Analyse /Convert Objects		
					Semantic Information		
					Render/Analyse Objects (RAO)		
					Data Object		
						Name*	
						Name Group*	
			Name Scheme*				
			Keyword*				
			Disposal*				
			Integrity Check				
			Precedence				
T4	How					Heading* First Heading* Second Heading*	
							significantProperties*
							objectCharacteristics
							storage*
							dependency
							dependencyName
							dependencyIdentifierType
							dependencyIdentifierValue
							software*
							swName*
							swVersion*
							swType:
							swOtherInformation*
							swDependency *
							hardware*
							hwName*
							hwType*
							hwOtherInformation*
						Change*	
						EAD Identifier*	
						File Plan*	
						Abbreviation*	
						Emphasis*	
						Expansion*	
						Geographic Name*	
						Item*	
						Number*	
						Other Descriptive Data*	
						Resource*	
						Revision	

						Description*	
						Spanned Column Specification*	
	Why	Description*			Reason for Creation*		
					Provenance Information *		
			Document Form*				
T5	Who				Name of Publisher *	Publication Statement*	
						Publisher*	
					Contacts or Rights Holders		
						Author*	
						Sponsor*	
							messageDigest Originator
							signer
						Imprint*	
						Abbreviation*	
						Emphasis*	
						Expansion*	
						Corporate Name*	
						Family Name*	
						Item*	
						Name*	
						Name Group*	
						Origination*	
						Personal Name*	
						Profile Description*	
						Subject*	
	When	Date*	Date Range* Start Date*		Change History Before Archiving *	Date*	
					Date of Publication *		
							dateCreatedByApplication*
							preservationLevelDateAssigned
							eventDateTime
							statuteInformationDeterminationDate
							termOfGrant
						Publication Statement*	
						Date of the Unit *	
			End Date*				
						Event*	
						Event Group*	
						Imprint*	
						Item*	
						Profile Description*	
	Where				History of Origin		
					Custody History		
					Change History Before Archiving *		
			Location*				
						Author*	
					Name of Publisher *	Publication Statement*	

						Publisher*	
					Place of Publication		
					Contacts or Rights Holders *		
							messageDigest Originator*
					storage		
						Sponsor*	
						Repository*	
						Physical Location*	
							contentLocationValue
							signer*
							copyrightJurisdiction
	Where						statuteJurisdiction
						Name*	
						Name Group*	
						Reference*	
						Reference Location*	
						Imprint*	
						Abbreviation*	
						Emphasis*	
						Expansion*	
						Corporate Name*	
						Geographic Name*	
						Item*	
						Origination*	
						Profile Description*	
						Subordinate Area*	
						Subject*	
T5						Physical Description*	
		Format*	Format*			Digital Archival Object*	
						Digital Archival Object Description*	
						Digital Archival Object Group*	
						Digital Archival Object Location*	
							storageMedium*
						Physical Facet*	significantPropertiesType
						Physical Characteristics and Technical Requirements*	
						Alternative Form Available*	
						Archival Description*	
	What					Archival Description Group*	
						Component Component (1) ~(12)	
						Text Division*	
						Function*	
							act
		Mandate*					restriction
							preservationLevel
							preservationLevelRole
					Reference Information		
					Resource Description*		
						Custodial History*	
						Accruals*	

T5	What					Edition*	
						Edition Statement*	
						Other Finding Aid*	
						Reference*	
						Reference Location*	
						Note*	
						Note Statement*	
						Related Material*	
						Profile Description*	
						Title Proper of the Finding Aid*	
						Index*	
						Index Entry*	
							originalName
		Rights*			Rights Management		
					Negotiation History		
					Rights Information		
					Rights Warning		
					Legislation Text Pointer		
					Permitted by License		licenseInformation
							licenseTerms
							licenseNote
					Licence Text Pointer		
					Copyright Statement		copyrightNote
							copyrightInformation
							copyrightStatus
							rightsBasis
						Descriptive Rules*	
						Organization*	
							statuteInformation
							statuteNote
					Context Information		
					Related Information Objects		
					Provenance Information		
					History of Origin *		
					Custody History*		
					Change History Before Archiving *		
					Management History		
						Subordinate Area*	
					Ingest Process History	Processing Information*	
					Administration History	Administrative Information*	
					Contacts or Rights Holders *		
					Content Information		
						Event*	
						Event Group*	
						ID of the Unit*	
							objectIdentifier
							significantProperties*
						Change*	

T5	What					Container*	
						EAD Identifier*	
						File Description*	
						File Plan*	
						Front Matter*	
						Physical Location*	
						Abbreviation*	
						Emphasis*	
						Expansion*	
						Geographic Name*	
						Item*	
						Name*	
						Name Group*	
						Number*	
						Other Descriptive Data*	
						Origination*	
						Personal Name*	
						Resource*	
						Revision Description*	
						Series Statement*	
						Spanned Column Specification*	
						Subject*	
						Title Page*	
						Title Statement*	
	How						eventDetail
							eventOutcomeInformation
							eventOutcome
							eventOutcomeDetail
						Physical Characteristics and Technical Requirements*	significantPropertiesType*
		Format*					storageMedium*
							size*
			Format*				format*
						Table Column Specification*	
						Digital Archival Object*	
						Digital Archival Object Description*	
						Digital Archival Object Group*	
						Digital Archival Object Location*	
							formatNote*
							formatRegistryRole
							creatingApplicationVersion*
							creatingApplication*
							creatingApplicationName*
							linkingEventIdentifier
							linkingEventIdentifierType
							linkingEventIdentifierValue
						Component* Component (1) ~(12)*	
							act *
		Mandate*					
							restriction*
							preservationLe

							vel*
							preservationLevelValue
							preservationLevelRole*
		Rights*			Rights Management *		
					Negotiation History *		
					Rights Information		
					Copyright Statement*		
							copyrightInformation*
							copyrightStatus *
							copyrightNote*
					Rights Warning*		
					Legislation Text Pointer *		
					Permitted by License*		
							licenseInformation *
							licenseTerms*
							licenseNote *
					Licence Text Pointer*		
						Descriptive Rules*	
						Other Finding Aid*	
							rightsGrantedNote
T5	How						rightsBasis*
							rightsStatement
							linkingRightsStatementIdentifierType
							linkingRightsStatementIdentifierValue
							statuteInformation*
							statuteNote*
							rightsStatementIdentifierType
							rightsStatementIdentifierValue
					Reference Information *		
					Resource Description*		
					Existing Metadata		
					Existing Records		
					Context Information *		
					Related Information Objects*		
					Provenance Information *		
					History of Origin *		
					Original Technical Environments		environment* environmentPurpose* environmentNote*
					Prerequisites		

T5	How				Procedures		
					Documentation		
					Management History *		
					Ingest Process History *		
					Administration History*		
					Action History		
					Policy History		
					Contacts or Rights Holders *		
					Fixity Information		fixity
							messageDigest Originator*
					Authentication Indicator		
					Content Information *		
					Representation Information		
					Structure Information		
					Underlying Abstract Form Description		
					Transformer Objects (TOs)		
					1) Platform		
					2) Parameters		
					3) Render/Analyse Engines		
					4) Output Format		
					5) Input Format		
					Render/Analyse /Convert Objects		
					Semantic Information		
					Render/Analyse Objects (RAO)		
					Data Object		
						Note*	
						Note Statement*	
						Index*	
						Index Entry*	
						Pointer*	
						Pointer Group*	
						Pointer Location*	
							objectCharacteristics*
							objectIdentifier *
							objectCategory
							significantProperties*
							significantPropertiesValue
							significantPropertiesExtension
							creatingApplicationExtension*
							objectCharacteristicsExtension *
							environmentExtension
							signatureInformationExtensio

T5	How						n
							eventOutcomeDetailExtension
							agentExtension
							rightsExtension
							compositionLevel
							inhibitors*
							inhibitorType*
							inhibitorTarget*
							inhibitorKey*
							storage*
							contentLocation
							contentLocationType
							contentLocationValue*
							linkingIntellectualEntityIdentifier
							linkingRightsStatementIdentifier
							linkingAgentIdentifier
							linkingObjectIdentifier
							eventIdentifier
							eventIdentifierType
							eventIdentifierValue
							linkingIntellectualEntityIdentifierType
							linkingIntellectualEntityIdentifierValue
							linkingAgentIdentifierType
							linkingAgentIdentifierValue
							linkingAgentRole
							dependency*
							dependencyName*
							dependencyIdentifierType*
							dependencyIdentifierValue*
							software*
							swName*
							swVersion*
							swType:
							swOtherInformation*
							swDependency*
							hardware*
							hwName*
							hwType*
							hwOtherInformation*
							signatureInformation
							signature
							signatureEncoding
							signer*
							signatureMeth

T5	How						od
							signatureValue
							signatureValidationRules
							signatureProperties
							keyInformation
							relationship
							relationshipType
							relationshipSubType
							relatedObjectIdentifierValue
							relatedObjectSequence
							relatedEventIdentification
							relatedEventIdentifierType
							relatedEventIdentifierValue
							relatedEventSequence
							eventType
							eventOutcomeDetailNote
							linkingObjectIdentifierType
							linkingObjectRole
							linkingObjectIdentifierValue
							agentName
							agentType
							agentNote
							licenseIdentifier
							licenseIdentifierType
							licenseIdentifierValue
							statuteCitation
						Change*	
						EAD Identifier*	
						File Plan*	
						Name*	
						Name Group*	
						Abbreviation*	
						Emphasis*	
						Expansion*	
						Geographic Name*	
						Item*	
						Number*	
						Other Descriptive Data*	
						Resource*	
						Revision Description*	
						Spanned Column Specification*	
	Why				Provenance Information *		
					Reason for Creation *		
					Reason for Preservation		preservationLevelRationale
							formatRegistryRole*
		Publisher*				Publication Statement*	
						Publisher*	

T6	Who	Audience*			Actors*		
						Sponsor*	
						Imprint*	
						Abbreviation*	
						Emphasis*	
						Expansion*	
						Corporate Name*	
						Family Name*	
						Item*	
						Name*	
						Name Group*	
						Origination*	
						Personal Name*	
						Subject*	
	When	Date*	Date Range* Start Date*			Date*	
						Publication Statement*	
						Date of the Unit *	
			End date*				
						Imprint*	
						Item*	
	Where	Publisher*				Publication Statement*	
						Publisher*	
						Sponsor*	
						Repository*	
					Actors*		
						Name*	
						Name Group*	
						Item*	
						Abbreviation*	
						Emphasis*	
						Expansion*	
						Address*	
						Address Line*	
						Corporate Name*	
						Geographic Name*	
						Imprint*	
						Origination*	
						Subject*	
	What					Conditions Governing Access*	
						Legal Status*	
						Conditions Governing Use*	
						Bibliography*	
						Bibliographic Reference*	
						Bibliographic Series*	
		Availability*					
		Identifier*					
		Subject*					
		Source*					
					Permitted by Statute *		
					Resource Description*		
					Actions*		
						File Description*	
						Title Proper of the Finding Aid*	
						Abbreviation*	
						Emphasis*	
						Expansion*	
						Geographic Name*	
						Item*	
						Name*	
	What					Name Group*	

						Number*	
						Origination*	
						Paragraph*	
						Personal Name*	
						Preferred Citation*	
						Series Statement*	
						Subject*	
						Title Page*	
						Title Statement*	
		Availability*					
		Identifier*					
		Source*					
						Table Column Specification*	
						Controlled Access Headings*	
					Permitted by Statute *		
					Resource Description*		
					Actions*		
							environment*
							software*
							swName*
							swVersion*
	How						swType*
T6							swOtherInformation*
							swDependency *
							hardware*
							hwName*
							hwType*
							hwOtherInformation*
						Name*	
						Name Group*	
						Abbreviation*	
						Emphasis*	
						Expansion*	
						Geographic Name*	
						Item*	
						Number*	