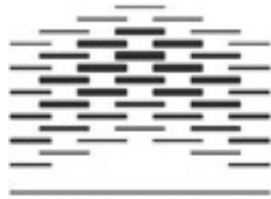




TALLINNA ÜLIKOOL



OSLO AND AKERSHUS
UNIVERSITY COLLEGE
OF APPLIED SCIENCES



Padam Gharti

Enriching Metadata Contents in Repositories: A case study in Institution Repositories of HIOA(ODA).

Supervisor: Michael Preminger

Oslo University College

Faculty of Journalism, Library and Information Science

DECLARATION

I certify that all material in this dissertation which is not my own work has been identified and that no material is included for which a degree has previously been conferred upon me.

.....Padam Gharti.....(Signature of candidate)

Submitted electronically and unsigned

ACKNOWLEDGEMENT

I am heartily thankful to my supervisor Michael Preminger, for his encouragement, guidance and support. The advice that he gave me truly helped for smooth progress of this thesis. He really enabled me to develop better understanding of the subject. His support and cooperation is much indeed appreciated.

I also express deepest attitude to all professors in DILL program who have given interesting lessons for me.

Besides, I am also fortunate to acknowledge to the individuals who I have interviewed for this thesis at HiOA. It is hardly possible to list name of all our well wisher here but we sincerely acknowledge the precious encouragement they offered us. I am indebted to all DILL students whose work have directly or indirectly influenced in this thesis.

I am grateful to all my former lecturers and all instructors at this program. My special thanks go to all my colleagues who have made this program so wonderful to be with. I owe my gratitude to many more people whose names are not mentioned here. Above all, I am thankful to my parents who are always supporting and encouraging me.

ABSTRACT

This thesis aims to explore the current state of metadata practices in Institution Repositories of HiOA in terms of the awareness of Institution Repositories, metadata and metadata types, the availability of guidelines, the chief source of information on enriching contents, the metadata scheme used, and the opinions of staff on the problems and factors regarding enriching metadata contents for digital collection in ODA, the institutional repository of HiOA. The study is limited to ODA as a case study.

A descriptive survey is used as the research method. Data was collected by using an online questionnaire. A survey link was distributed to informants who play different roles in the ODA projects. The research is based on qualitative methodology because it focuses on investigating the point of views from Hioa librarians, outside experts and authors as well as to analyzing the semantic of metadata elements being used in current digital publication at Hioa database. A total of 17 returned questionnaires were analyzed.

The presented case study reports an activity focuses on the study of a metadata element which generated an benefit that can be useful for different purposes such as research, learning, easy retrieval from IR. It is conceived as an open access initiative and supported by the institutional repository on which we rely in order to ensure the long-term preservation, the visibility and the interoperability of the solution.

In managing the metadata and making it richer different problems and challenges have been faced and investigated by these memory institutions and further research should be carried out to study other aspects of metadata implementation. The findings in the thesis could be a useful reference for the ODA projects.

Keywords: Content enrichment, Institution Repositories, metadata practices, ODA.

TABLE OF CONTENT

ACKNOWLEDGEMENTS	3
ABSTRACT.....	4
CHAPTER 1: INTRODUCTION	8
1.1 Background.....	8
1.2 Problem Statement.....	9
1.3 The aim of the study and the research questions.....	10
1.4 Research Methodology.....	11
1.5 Scope of the Study.....	11
1.6 Outline of the Thesis.....	12
CHAPTER 2: LITERATURE REVIEW.....	13
2.1 Content Enrichment.....	13
2.1.1 The Enrichment Approach.....	15
2.1.2 Metadata and Enrichment.....	15
2.2 Institution Repositories.....	16
2.3 Metadata practices in Institutional Repositories.....	18
2.3.1 The Importance and characteristics of metadata.....	19
2.3.2 Digital and Digitized Documents.....	21
2.3.3 Metadata Schema and Factors Affecting the Choice of a Metadata Scheme.....	21
2.4 Challenges of Metadata Practices.....	24

CHAPTER 3: RESEARCH METHODOLOGY	27
3.1. Research Objective and Research Questions.....	27
3.2 Research Population and Sampling.....	29
3.3 Data Collection Instrument.....	29
3.4 Online Questionnaire.....	30
3.5 Data analysis Methods.....	31
3.6 Limitations of the research.....	32
CHAPTER 4: DATA ANALYSIS AND DISCUSSION	33
4.1 Analysis of data collected by online questionnaires.....	34
4.1.1 Present practices in HiOA' institution repositories ODA.....	34
4.1.2 Contents for Metadata enrichment for collection in ODA.....	36
4.2 The usage of metadata elements in enriched ODA.....	38
4.2.1 The ways of customization of metadata elements in enriched ODA.....	38
4.2.2 Usage of qualified Dublin Core in ODA while enriching.....	39
4.3 The ways how repositories suited into the Scholarly communication.....	40
CHAPTER 5: CONCLUSION AND FUTURE RESEARCH	41
5.1 Treatment of Research Questions.....	41
5.2. Implications of the Research.....	43
5.3. Future Research Ideas.....	44
REFERENCES.....	46
APPENDICES.....	52

APPENDIX 1: DUBLIN CORE METADATA INITIATIVE-Dublin Core Qualifiers.....52

APPENDIX 2: THE ONLINE QUESTIONNAIRE.....54

LIST OF FIGURES AND TABLES

Table 2.2: Typology of IRs.....17

Table 4.1 The profile of Informants.....34

Fig 4.2 Usage of qualified Dublin Core in ODA.....39

Chapter1: Introduction

The chapter provides the background and statement of research problem as well as the aim of study and research questions. Afterwards, the scope of the study as well as the research methods is presented. Finally, an outline of the thesis is introduced.

1.1 Background

Content enrichment is a vital component of the new wave of publishing. Content that is better described will enable discoverability through improved indexing and linking (e.g. related articles, browsing by topic) and ultimately increased usage. That usage can be key to driving Impact Factor, and is becoming even more important with the growing adoption of article level metrics. Also, as we move from subscription models, usage is vital to support advertising based revenues. The future lies in being able to do more with content and its associated data. If content is enriched and well described, we open up new opportunities. Just one example is the ability to illustrate trends in research—i.e. which topics are becoming increasingly active, and which are becoming less so. With well-enriched content, these capabilities become easier to realize. Enriching article content has been accomplished by bringing discipline- specific research tools inside the articles. Adding article context satisfies the need that is frequently observed from the behaviour of researchers. For instance when reading an article on protein or gene sciences, the relevant protein or genomic information is fetched from the protein DataBank or GenBank for inspection(Elsevier, 2012).

Repositories can be of many shapes and sizes, from small specialist collections to national or international services. Many institutional repositories initially focussed on research outputs and some still limit their collections to this type of content(Lou McGill, 2010). Others have started to widen the original remit to include learning and teaching materials. Whilst institutional VLEs(A Virtual Learning Environment) have, to some extent acted as stores for learning and teaching materials, they tend not to support the search and retrieval functions required for a repository. Making this content more open, even within the institution, presents challenges for institutions with a commitment to open up their resources.

Institutional Repositories (IR) are appearing an increasingly significant type of special resource and service given by libraries (Li et al., 2011). Libraries are building repositories to archive the intellectual output of their faculty members, scholars and students (Shreeves,

2009). IR often contains presentations, historical research conducted at the university that has been converted into digital form, technical reports, working papers, electronic theses and discourse, and datasets (McDowell, 2007; Shreeves, 2009). A good trusted Digital Repository is one whose mission is to provide reliable, long-term access to managed digital resources to its designated community, now and in the future. Digital repositories may include a wide range of content for a variety of purposes and users. What goes into a repository is currently less an issue of technological or software ability, and more a policy decision made by each institution or administrator. (Trusted Digital Repositories, 2002).

Learning Center of HIOA

Oslo and Akershus University College of Applied Sciences (HIOA) is Norway's largest state university college, with a student body of approximately 16,000 students and 1,600 employees. HIOA has four faculties located at two campuses: Pilestredet and Kjeller. It has four faculties: Faculty of Health Sciences, Faculty of Education and International Studies, Faculty of Social Sciences, Faculty of Technology, Art and Design. It offers more than 50 Bachelor programs, 25 Master programs, three PhD programs and a large number of other courses. HIOA learning center is the heart of education and research for its students and faculty members. The library system contains links to reliable and credible sources and different electronic services for the users. The sources and services are accessible to users on and off campus. It has a department for digital services (the digital library) too. The department has 4-5 employees and is responsible for the learning center website, the digital publications, the management of shared electronic information resources, and the research documentation for HIOA (Rahman, 2011).

1.2 Problem Statement

ODA, Open Digital Archive, is HiOA's institutional open access repository. Repositories enable all researchers, regardless of what institution they are affiliated with, to access its work. ODA contains research produced by HiOA's employees, including scientific journal articles, scientific book chapters, scientific monographs, approved doctoral theses and approved theses by students from HiOA's master's degree programs. Putting digital content

into an institutional repository enables institutions to manage and preserve it, and therefore derive maximum value from it. A repository can support research, teaching, learning, and administrative processes (Lou McGill., 2010). These advances are becoming gradually available in electronic publication systems and, if incorporated in Open Access systems (e.g. journals, repositories) can contribute to even wider adoption by users and publishers. Such incorporation will enable, apart from intuitive reading capabilities, the efficient viewing of large data sets visualizations of videos and presentations, maps and/or images of cultural artifacts.

ODA contains:

- Peer-reviewed journal articles and other scientific documents registered in Cistin.
- Approved theses of students from HiOA's own master's degree programmes.
- Approved doctoral theses of students from HiOA's own research programmes.

The contents of the ODA is automatically harvested by national and global research resources, inter alia, NORA (Norwegian Open Research Archive) DRIVER (Digital Repository Infrastructure Vision for European Research), BASE (Bielefeld Academic Search Engine) and Google Scholar. Dissertations published in ODA are also automatically made visible in DART-Europe E-theses Portal containing more than 300,000 papers from more than 400 European universities. ODA contributes to the scientific activity of HiOA get increased visibility worldwide.

1.3 The aim of the study and research questions

This research attempts to explore the current state of metadata practices in Open Digital Archive(ODA) repositories, generally and with ODA as a case, The study is limited to ODA as a case study. ODA is currently using Dublin Core Metadata element set as a default metadata schema. It is important the metadata schema cover all the different file types to enrich the contents in Digital Journals. To achieve this aim, two following research questions are going to be regarded:

Research Question 1: *What are the Existing Dspace Facility in HioA and metada standards that are in use.*

Research Question 2: *Type of materials in Dspace the faculty, publishers and librarians would like to keep in the repository and they want to enrich.*

Research Question 3: *To identify the required qualified Dublin Core metadata elements for Dspace while enriching contents and issues on doing this.*

1.4 Research Methodology

In this study, ODA at HIOA is chosen as a case for investigation. I performed an informal survey about the needs of e-journals when it comes to content enrichment. I have taken a look at some proposed journals. Authors and publishers were asked a set of questions using google forms by email to ask them the necessity while enriching e-journals. The questionnaire consisted of both open-ended and closed-ended questions written in English. Data collected from questionnaires are qualitative data because all questions were designed to get the opinions and experiences of informants about many kinds of research issues. Afterward, constant comparative analysis (Hewitt-Taylor, J., 2001, p.42) is used to analyze data gathered from questionnaires.

1.5 Scope of Study

By exploring metadata practices for digital repositories. The study may enhance our understanding of the current state of metadata practices of digital repositories in HiOA. Enhancing article content is the key theme of the study. The aim is to increasingly enrich the value of research articles by including new and interactive content elements, mostly discipline-specific and key to the scientist's research and workflow. this study on content enhancements provide a number of benefits to authors and publishers, which in turn will can improve the user experience for journals and their readers. Due to time and technical constraints only informants involved in ODA repository are consulted for this study.

1.6 Outline of the Thesis

The content of thesis is presented in five chapters in addition to table of content, figures and tables, reference and appendices.

Chapter one presents an introduction that describes the background of the study and the statement of the problem. It also includes brief introductions to the research objectives and research questions, the scope of the study, the methodology, and the significance of the study as well as an outline of this study.

Chapter two gives a review of the related literature on key concepts considered relevant and necessary for an understanding of the study of content enrichment related to institutional repositories and metadata practices for enriching them.

Chapter Three focuses on the methodology used in this research. The research population and sampling are presented. The procedures for constructing the data collection instrument are also described. Moreover, data collection and analysis as well as the limitations of study are included in this chapter.

Chapter Four reveals the findings of the study and this is contextualized with a discussion based on the literature review and theoretical background.

Chapter Five provides a summary of the research and conclusions from the findings. The implications of the study and recommendations for future research are also included.

Chapter 2: Literature Review

This chapter will review the literature relevant to this research. First it will explain what Institutional Repositories are, content enrichment in digital Journals in IR will then be discussed. Next there is a section on metadata in terms of its definition, roles and types, as well as metadata standards for descriptive and subject cataloging as well as criteria to adopt metadata schemas. challenges for the creation of metadata are also included.

However, it is not intended to be exhaustive survey. The useful sources for literatures on this study are books, online databases and Internet. Books were consulted for insightful concepts and theory. Several online databases were mainly used for gathering relevant articles which present the current states, conceptual and practical information and trends from the authors' viewpoints. More specifically, Springer Link databases, D-Lib Magazine, Cataloging & Classification Quarterly, Emerald Insights, JSTOR, Informa etc were used. To find documents related to topic, some search engines were used including Google Scholar, and BYBSYS at Oslo University College Library, Google, Google Scholar, and online journals were also worth to access for the relevant sources. According to the scope of the study, the related concepts can be divided into three key themes: Content Enrichment, Institutional Repositories, Metadata Practices. Several keywords have been used for searching documents like "content enrichment", "Institutional Repositories", "metadata practice". Further, other related keywords were also formulated along the way in order to obtain more literatures .

2.1 Content Enrichment

Before the advent of modern publishing technology. In the past science was simply recorded through hand-writing and until the arrival of the published medium remained isolated and largely unquestioned with print technology came new methods of recording and communicating research. But this two was limited article reflect, references and supplementary information intangible and there was little connection with the authors research process what opportunity to gain their deeper inside? In today's world information loose more quickly and our level of knowledge advanced more rapidly so too the communication of research (Aalbersberg, IJsbrand Jan, et al, 2012).

Content enrichment is a vital component of this new wave of publishing. While the majority of access to content is now online, PDF has been remarkably tenacious as the primary format that people use for reading and sharing.

As stated by Jason Markos (2015), with the content enrichment taken into place, the user will be able to realize 3 key objectives:

Increase relevance, discoverability, and usage

Content that is more suitably described will enable to recognize the potential through improved indexing (Search Engine Optimization) and linking (e.g. related articles, browsing with topic) and eventually expand the practices of usage. That usage can be key to motivating Impact Factor(IF), and is becoming even more significant with the growing adoption of article level metrics. Also, as we move from endorsement models, usage is vital to support advertising based earnings.

Improved Feasibility

Feasibility is key, not just from the arms race point of view (groups 'X' is doing it, so we'd better do it), but also for securing authors, which makes conceivably more important. As authors become customers, society publishers will need to demonstrate more value to them by making their work easier to find and utilize. There are a number of ways that this can be achieved, but one strategy is to better capture their hard work represented by the data accompanying their articles.

Facilitate new value concept

The future resides in being able to do more with content and its corresponding data. if Content is enriched and well described, then we open up for new possibilities. One example is the ability to emphasize tendency in research—i.e. topics which are becoming increasingly active, and which are becoming less so. With well-enriched content, these capabilities become easier to understand.

Aalbersberg, IJsbrand Jan, et al.(2012), explained on topic “Adding article context” that Implementing a content enrichment strategy is important to existing and growing in this new world. Adding article context satisfies a need that is frequently observed from the behaviour

of researchers. After or even while reading an article, researchers search for related information or associated data sets. For instance, when reading an article on protein or gene sciences, the relevant protein or genomic information is fetched from the Protein Data Bank or GenBank for inspection. Similarly, in chemistry there is a need for basic information on the chemical compounds being mentioned in an article, plus references and links to the relevant databases in which these compounds are covered. The Article context can be added with this behaviour by presenting the related key information next to the article, saving the researcher the additional effort of searching for this information.

2.1.1 The Enrichment Approach

The approach for enrichment of metadata based on the contents of Institutional repositories. Metadata that is enriched is contextually and semantically linked and openly accessible; enriching solely deals with metadata content, filtering addresses all issues associated with its interface. Content Enrichment make deeper audience engagement and adds value, making your content more effective and compelling. The enriched contents with proper taxonomy, metadata and indexing will make easily discoverable and high ranking for today's searches. Also tagging approach like Semantic tagging enriches our content with very precise metadata. These precise tags denote elements specific to university or organization needs such as the function of the document, tags relevant to the universities and organization such as topics, dates, companies, geography and much more. Properly enriched content is easily searched, retrieved, and reused between people, programs, and computers via web, digital media or print etc.

2.1.2 Metadata and Enrichment.

Metadata is data about data; it is all the information about the asset. It can include any attribute that helps define or describe the asset. Metadata is a critical part of Content Enrichment because it is the key to helping your audience find the files they are looking for. So where does it come from? Fortunately a great deal of metadata is already provided. For example your computer operating system already tracks filenames, file size, date created and date modified which most digital access systems can be automatically catalog. Your devices automatically record metadata as well. For example digital camera records the data taken and, via GPS, the location where the image was taken. Even your suppliers provide data. For example purchased photos and purchased music come with embedded metadata. Your

vendors are already tracking information like part numbers, their description, size, etc. Content Enrichment will increase discoverability, enable reuse, and increase the marketing potential.(SMS Marketing, 2015).

2.2 Institutional Repositories

Institutional repositories play a key role in universities and research centers for the preservation and dissemination of the knowledge generated or collected by agents. According to Basefsk (2009), "The Institutional Repository (IR) concept was born out of competition for who was going to be responsible for dissemination of an institution's intellectual product via the internet." Furthermore, McClendon (as cited in McCormick, 2006) said the Institutional Repository (IR) concept has gained momentum as universities begin to question the logic of buying back their research. Lynch (2003) further specified "Institutional repositories will succeed precisely because they are responsive to the needs of campus communities, and advance the interests of campus communities and of scholarship broadly." Instead, Jone (2007) described that "The undergoing period of change on Information Environment, from the delivery mechanism of materials to the assumptions of the users of information service; institutional repositories are a feedback to some of these changes" p. 48.

Institutional repositories are often referred to in strategy and policy documents as they can support key institutional aims and objectives. Typically content can include research outputs such as journal articles or research data, e-theses, learning and teaching materials, and administrative data. Some repositories only store particular items (such as theses or journal papers), whilst others seek to gather any credible scholarly work produced by the institution; limited only by each author's retained rights from publishers.(Lou McGill, 2010).

Lou McGill(2010) explains Digital repositories may include a wide range of content for a variety of purposes and users. The focus of each repository, and, therefore, what content it will store, will depend upon the policy decisions made by each institution or administrator. To give an idea of the type of content currently held in repositories worldwide, the **Directory of Open Access Repositories OpenDOAR** has identified the following as the most common content types in repositories:

- Journal articles

- Bibliographic references (metadata only)
- Books, sections and chapters
- Conference and workshop papers
- Theses and Dissertations
- Unpublished reports & working papers
- Datasets
- Content–packaged learning objects
- Multimedia and audio–visual materials
- Software
- Patents
- Other special items

Moreover, Institutional repository grow an crucial infrastructure for scholar activities in universities on the world. This is evidenced by the development of thousand of IRs listed in DOAR (Directory of Open Access Repositories). Lynch (2003) defines IRs as: “a set of services that a university offers to the members of its community for the management and circulation of digital materials created by the institution and its community members” (p.1).

Heery and Anderson (2005) developed a typology that provides a helpful framework for exploring IRs, as presented in Figure 1 below:

This framework presents four main focus of IRs including content, coverage, users and functionality.

Content	Coverage	User Group	Functionality
Raw data Derived data Full text pre-print Full text peer-reviewed E-theses Learning objects Corporate records	Personal Journal (single or multiple) Departmental Institutional Inter-institutional National International	Learners Teachers Researchers	Enhancing access Preservation Dissemination Asset management Resource sharing

Table 2.2: Typology of IRs (Heery and Anderson, 2005, p.1)

2.3 Metadata practices in Institutional Repositories

Metadata, meaning “data about data” is a term that originated in the mid 1970s, incorporating concepts drawn from libraries, archives, records management, scientific data management, text markup, computer science, and related fields. By now, the term is so widely used as to be meaningless without qualification by context of use(Lynch et al. 1995).

Dempsey and Heery (1998, p.149), consider metadata is “data associated with objects which relieves their potential users of having to have full advance knowledge of their existence or characteristics. It supports a variety of operations. A user could be either a program or a person.” Miller (2004) defines metadata is “...the “extra baggage” associated with a resource that aids a user in finding that resource (find); discover where, and by whom it was created (identify); decide whether the resource is of value to the user (select); and conclude whether there is feasible access to the resource (obtain).” He explains that *metadata* should be aligned with FRBR’s (Functional Requirements for Bibliographic Records) user tasks.

Literally, “data about data,” metadata includes data associated with either an information system or an information object for purposes of description, administration, legal requirements, technical functionality, use and usage, and preservation.

Metadata is an essential building block in facilitating effective resource discovery, access, and sharing across ever-growing distributed digital collections. (Park and Tosaka 2010), Quality metadata is becoming critical in a networked world in which metadata interoperability is among the top challenges faced by digital libraries. However, there is no common data model that cataloging and metadata professionals can readily reference as a mediation mechanism during the processes of descriptive metadata creation and controlled vocabulary schemata application for subject description.

A popular system that operates as a digital repository for educational purposes is DSpace(D-Lib, 2010). DSpace uses the qualified Dublin Core (DC) element set as its base metadata schema. However, because the DC schema is sometimes proven to be inadequate for the efficient characterization of educational material, and researchers attempts development of an application profile, extended with the LOM(Learning Object Metadata) metadata standard, and tailor it to the needs of an educational repository..

In general, there are four groups of metadata creators: librarians, authors, users and machines (Mathes, 2004). Metadata created by librarians usually follows detailed sets of rules and workflows. To this effect, librarians receive formal trainings in cataloguing/metadata and the use and application of standards and principles, such as AACR, DDC, MARC, LCSH, MeSH, RDA and FRBR, to mention but a few (Lagoze, 2010; Lavoie & Gartner, 2005; Zeng & Chan, 2006). Consequently, the metadata created by librarians is considered to be of relatively higher quality, particularly with regard to accuracy, completeness and consistency (Bruce & Hillmann 2004; Haynes, 2004; Park & Tosaka, 2010). Furthermore, as Mathes (2004) points out, yet another main limitation of both librarian and author-created metadata is that “the intended and unintended eventual users of information are disconnected from the process”. Moreover, Doctorow (2001) contends that authors may want to see their works appear at the top of search engines listings, and, hence, may fail to accurately represent their contents, thus indicating the limitations of author-created metadata.

2.3.1 The Importance and characteristics of metadata

Metadata has been an essential component of the digital projects. That is because metadata is crucial for information retrieval especially in search accuracy, assisting evaluation, and the harvesting of digital resources. Particularly for non-textual resources, metadata is essential (Rettig, Liu, Hunter, & Level, 2008). Consequently, comprehensive and detailed metadata can influence the long-term discovery of resources (Hughes, 2004, p.206).

Gilliland (2008, p.6) details the roles of metadata in environments where users can access information without help from intermediaries, as follows:

- Certifies the authenticity and degree of completeness of the content;
- Establishes and documents the context of the content;
- Identifies and exploits the structural relationships that exist within and between information objects;
- Provides a range of intellectual access points for an increasingly diverse range of users; and
- Provides some of the information that an information professional might have provided in a traditional, in- person reference or research setting.

With carefully structured descriptive information, metadata can enhance a remote user's ability to discover resources and search effectively. In addition, metadata provides the context of an information object and maintains the linkage between the object and a digital surrogate. Besides, metadata supports managing digital objects and ensures that they will be accessible in the future by keeping technical data on producing, storing, and maintaining those objects. These data enhance the ability of museums, archives, and libraries to track the lineage of digital objects. Additionally, metadata also allows institutions to track rights, licensing and reproduction information. In term of interoperability, metadata allows diverse institutions to exchange and search for information across systems. Therefore, it expands the usage of collections in the digital age and reaches various users' needs regardless of geographical constraints and diverse institution types (Gilliland, 2008, p.15-17; Lagoze & Payette, 2000, p. 99; NISO, 2004, p.1- 2).

Once the metadata is collected, it must be reliably stored and is used to manage the resource (for example, it can be used to review the strength of a repositories collection in a particular domain or to check the currency of the file formats). Certain technical protocols or standards, such as **OAI-PMH** can be used to enable external services to access metadata records and facilitate resource discovery by a much broader audience.

It is important to consider these two contexts together. Strong communication between repository administrators and technical teams is needed to ensure the requested metadata input matches the requirements for the local management of the repository's items and enables the repository to be harvested for external use.

2.3.2 Digital and Digitized Documents

According to (name of book) digital documents and digital information, the terms refers to two classes of items: "digital"-those that are created originally in digital form such as electronic text or video, and "digitized"-those that are copied or translated into digital form from some other medium such as paper or film. Digital documents are sometimes referred as "born digital" to indicate that they originated in electronic form. An increasing portion of online materials are in this category, as text is written in word processors, images are captured by digital cameras, music is electronically synthesized, and content for many media are edited, manipulated, and composed in electronic forms. Many digital documents will exists

only in that form, such as electronic publications and electronic mail. Materials are digitized for a number of purposes, one of which is to improve access. Instead of one physical copy of a document, digital surrogate, or image of the document, can be distributed online, to multiple users. The Library of Congress is finding that more people visit their site online than the library building on any given day (Smith 1999). Once digitized, documents can be manipulated easily. Documents can be edited, merged, or otherwise changed in ways not possible with physical forms.

2.3.3 Metadata Schema and Factors Affecting the Choice of a Metadata Scheme

“A schema is a logical plan showing the relationships between metadata elements, normally through establishing rules for the use and management of metadata specifically as regards the semantics, the syntax and the optionality (obligation level) of values.” Also referred to as an element set. Like database records, digital library metadata should have a schema that represent what are the individual needs and decisions of their providers on describing their objects. Digital libraries should be able to communicate and share objects. There are some established schemas that cover many usual needs in digital libraries, but the specific needs of every library make some of them to not to use the same (or one of the well established schemas) or to modify them in a local way. Caplan (2003, p.5) indicates that “the term *scheme* and *schema* are used interchangeably with this general definition. *Schema*, however, has another meaning in relation to computer database technology as the formal organization or structure of a database, and another specialized meaning in relation to XML.” For this study, the term *scheme* is used as Caplan has suggested.

Due to the proliferation of metadata schemes from numerous communities, digitizing projects need to consider and evaluate many points before implementing a system. This is a crucial step influencing the effectiveness of resource discovery and the usability of information resources (Baca, 2003, p.48). In general, the best consideration for choosing a scheme is that scheme most closely fits identified requirements and has the widest acceptance within the community (Ma, 2006, p.8). In addition, a metadata scheme which is appropriate to the holdings and the potential end-users must be selected (Baca, 2003, p.54).

According to the article, “Choosing a Metadata Standard for Resource Discovery” (Kelly, 2006), it is recommended to consider several following issues before implementing a standard: 1) Granularity – At this point it is worth considering whether your metadata should be created at which material type we deal with and level we describe? 2) Interoperability – It is important, wherever possible, to choose one of the leading standards from within your subject community or domain in order to enable sharing information. 3) Support – Using a standard that is well supported by a leading institution can also bring cost benefits. 4) Growth – a standard may or may not be further developed. Are there working groups and workshops? 5) Extensibility – It may be necessary to be extensible and allow combinations with metadata elements from more than one standard. 6) Reputation-Resource bodies will be familiar with established, international standards allowing digitization. 7)Ease of use – a simple standard does not require much expertise or training to create metadata. 8) Existing experience – former experience with metadata schemes would probably reduce the implementation time.

While there is a variety of metadata schemata currently in use for organizing digital collections, only a few of them are widely used in digital repositories. In her ARL survey, Jin Ma(2007) reports that the MARC format is the most widely used metadata schema (91 percent), followed by Encoded Archival Description (EAD) (84 percent), Unqualified Dublin Core (DC) (78 percent), and Qualified DC (67 percent). Similarly, a 2007 member survey by OCLC Research Libraries Group (RLG) programs gathered information from eighteen major research libraries and cultural heritage institutions and also found that MARC is the most widely used scheme (65 percent), followed by EAD (43 percent), Unqualified DC (30 percent), and Qualified DC (29 percent). The different levels of use reported by these studies are probably due to different sample sizes and compositions, but results nonetheless suggest that metadata use at research institutions tends to rely on a small number of major schemata. (Karen Smith, 2007).

Dublin Core

The Dublin Core Element Set (DC) was developed in 1995 as a means in which to improve indexing of search engines by embedding metadata elements into web pages or encoding through the use of XML (Huddleston, 2008). This metadata standard was created to increase interoperability of metadata records, by bridging the differences of the existing objects

descriptions. it is common denominator of existing metadata standards, it has only 15 optional and repetitive elements that are very generic and clear in context, and they represent semantic crosswalks among metadata standards in different disciplines. Dublin Core (Dublin Core Metadata Initiative, 2011a) is an easy to learn and use schema that is a basic default metadata template in many digital content management systems. The goals of Dublin Core are simplicity and ease of use, commonly understood semantics, international scope, and extensibility (Intner, Lazinger, & Weihs, 2006). It was created to be intentionally “generic,” allowing user communities to define content standards and the use of controlled vocabularies that fit specific needs. The interoperability of Dublin Core metadata fields makes it easy to share data and create discovery opportunities. Unqualified, or Simple, Dublin Core is a requirement of the Open Archives Initiative (OAI) for its OAI Protocol for Metadata Harvesting (OAI-PMH). Allowing harvesting of metadata creates a wider discovery base, leading researchers back to the richer metadata at the original site. Dublin Core may also be expressed and exchanged using XML or Resource Description Framework (RDF), however, the nonspecialist metadata creator does not need to know how to use these standards if entering data into an easy-to-use content management system such as CONTENTdm, DSpace, or LUNA Insight.

Extensibility is especially important in creating the richer metadata needed for unique user communities because the core 15 elements may be extended by adding additional elements. This allows repositories to develop fields for enriching and cataloging practices. Extensibility creates more opportunities for researchers to discover content enrichment. Any element can be refined or qualified. For example, the *Date* element can be refined as date created, date valid, date available, date issued or date modified. The key principle in the element refinement is “dumping down”; the qualifier can be safely removed and the element value interpreted as a simple element. so the date created, for instance, can be safely interpreted as a Date. This principle allows qualified and simple Dublin core to co-exist easily inside digital libraries(Witten, Bainbridge & Nichols, 2010).

Encoded Archival Description (EAD)

Enables standardized exchange of descriptive data contained in specific types of archival finding aids known either as archival inventories or manuscript registers. It provides tools for a detailed, multilevel description, structured display, navigation, and searching. Primarily,

EAD is intended to accommodate descriptions of archival holdings in various media (Intner; Lazinger; & Weihs, 2006, p.90).

Categories for the Description of Works of Art (CDWA)

Provide access to information of cultural heritage resources, art databases for describing and collecting information about works of art, architecture, other material culture, groups and collections of works, and related images. This framework provides 532 categories in which to describe works of art.

MODS

Metadata Object Description Schema(MODS) was developed in 2002 by interested experts from Library of Congress' Network Development and MARC Standards Office, for a bibliographic element set that may be used for variety of purposes, and particularly for library applications.

2.4 Challenges of Metadata Practices

The diverse aims and practices of resource developers and their essential user groups are revealed through variation in metadata standards. MARC and Dublin Core have been widely adopted in library profession where digital metadata has been used constantly to both digital and non-digital works (Besser, 2002). Metadata is an increasingly well-accepted approach to organize digital collections in order to accommodate information organization, information retrieval, long-term preservation and interoperability. It poses challenges to existing cataloging practices, despite the fact that metadata provides numerous opportunities for libraries, archives, and museums to organize information in the digital environment, (Chen, Chen, & Lin, 2003, p.1; Ma, 2009). Current practice of metadata creation is problematic due to the lack of a mechanism for integrating various types of metadata schemata, content standards, and controlled vocabularies in ways that promote an optimal level of interoperability across digital collections and repositories,(Park and Tosaka 2010).

- ***Metadata Standards***

Metadata standards,” reflects how some domain specific metadata schemata tend to dictate the use of particular controlled vocabularies. The rapid increase in numbers of emerging metadata schemes and controlled vocabulary standards provides alternatives for digital project implementation. Nevertheless, it causes difficulties because it requires the implementing community to choose the most appropriate standard for their particular contexts. According to Zeng, Lee, and Hayes’s (2009) research on major concerns regarding metadata and controlled vocabularies conducted by distributing a web-based questionnaire to the International Federation of Library Associations and Institutions (IFLA), the most remarkable concerns concerning the determination about element set standards are “to choose which metadata standard to use” and “to learn how to use different metadata schemes together” (62.40 % and 59.40% respectively). For decisions about authority files and controlled vocabularies, the major concern is “to decide whether to use existing controlled vocabularies or authority files (e.g. LCSH, ULAN [The Union List of Artist Names], LC Authorities)” (64.60%).

- ***Time and Cost***

Some professionals feels that a more structured metadata, one that is managed through controlled cataloging rules is the better option. what a more structured metadata can provide is increased standardization, precision and interoperability, but at what cost? No only would the structure form be more complex in nature, therefore requiring greater staff training, but the standardization process would be both time consuming and costly. Additionally, metadata production requires easy-to-use and standardized tools which are expensive (Zeng, Lee, & Hayes, 2009). In supplement, creating uniform metadata regardless a diverse of materials and repositories is costly and difficult (Ma, 2009).

- ***Uniformity***

Each institution has its own metadata guidelines. This guidelines and standard differ from project to project. This therefore affects the consistency of metadata creation within a collection and across collaborating repositories (Park, 2009, p.221). Associated with the metadata development. Detailed metadata can also lead to issues with consistent and standard metadata – as a rule it is typically easier to standardized simple metadata. Besides ease in

consistency, simple metadata is both less costly and provides a greater probability of interoperability. However, simple metadata is not without its flaws, the greatest being an increased chance of false results during information retrieval, due to less detailed and specific search parameters. Whether simple or detailed, in the end, “the richness of metadata descriptions will be determined by policies and best practices designated by the agency creating the metadata” (Duval et al., 2002). Park (2009, p.224) suggests that simple metadata guidelines embedded in Web form or a template provide benefits for the creation of quality metadata.

- ***Interoperability***

Interoperability literally denotes the ability of a system to work with or use parts of other systems. Metadata interoperability across distributed digital repositories and collections is fast becoming a major issue. The variety of metadata standards used for describing digital objects and providing subject access among communities causes difficulties for information sharing. This requires metadata crosswalk and mappings to accommodate metadata interoperability (Ma, 2009).

- ***Knowledge and skills***

Continuing education and training for metadata professionals enables them to potentially work in new digital circumstances, and it influences the effectiveness of metadata creation (Park, 2009, p.225). Although museum professionals realize the importance of metadata more and more, they feel it requires specialize skills to manage digital information, interpret it for remotely end-users, and preserve it for the next generation (Spinazzè, 2004, p.47)

Chapter : 3 Research Methodology

The research is based on qualitative methodology because it focuses on investigating the point of views from Hioa librarians, outside experts and authors as well as to analyzing the semantic of metadata elements being used in current digital publication at Hioa database. Since Enriching the contents for digital collection by making richer metadata in ODA at HiOA is a specific situation, the research methods used is case study. Pickard (2007, p.86) addressed that the purpose of a case study is to “provide a holistic account of the case and in-depth knowledge of the specific through rich descriptions situated in context”. She further stated that “using case studies is the most appropriate method when the purpose of the research requires holistic, in-depth investigation of a phenomenon or a situation from the perspective of all stakeholders involved” (p.93).

In this study, ODA at HIOA is chosen as a case for investigation. I performed an informal survey about the needs of e-journals when it comes to content enrichment. I have taken a look at some proposed journals. Authors and publishers were asked with some set of questions using google forms by email to ask them the necessity while enriching e-journals. Data collected from questionnaires are qualitative data because all questions were designed to get the opinions and experiences of informants about many kinds of research issues.

Chapter 3 comprises of five section. The first section presents the research objective and research questions. The second section describes whom the research population for this research is. In third section there is explanation of data collection instrument and methods of data collection as well as the rationale for selecting those methods. The fourth section describes how the collected data is analyzed. And the final section discuss the limitation of study.

3.1. Research Objective and Research Questions

This research attempts to investigate the current state of metadata practices for digital repositories collections in Dspace of HiOA Library. To accomplish the objective, the aims of the project can be divided into two main research questions with sub- questions as follows:

RQ1: What is the existing Dspace facility in HiOA and metadata that are in use?

Sub-questions 1.1 What are the existing Dspace facility in HiOA library.?

Sub-questions 1.2 what metadata standards/schema are used currently for digital Collections/Journals in HiOA Institutional Repositories.

RQ2: *Articles in Dspace that publishers and Librarians want to enrich.*

Sub-questions 2.1 what type of materials the faculty would like to keep in the repository?

Sub-questions 2.2 what contents of metadata in articles/journals do publishers and Librarians want to enrich?

RQ3: *What kind of metadata elements they need while enriching contents and issues concerns?*

Sub-question 3.1 what metadata elements are needed while enriching contents?

Sub-question 3.2 To identify the required qualified Dublin Core metadata for Dspace while enriching contents and issues on doing this.

Some supplementary questions have also been asked during the conversation based on the flow of the interview or sometimes to get more clarification from the respondents. The respondents involved in the study had freedom to suggest anything they considered relevant to the study. All the interviews were recorded with the permission of the respondents.

To achieve these research questions, descriptive survey was considered as an appropriate approach for this study. Leedy and Ormrod (2010, p.187) describe survey research as involving “acquiring information about one or more groups of people – perhaps about their characteristics, opinions, attitudes, or previous experiences – by asking them questions and tabulating their answers. The ultimate goal is to learn about a large population by surveying a sample of that population; thus, we might call this approach a *descriptive survey* or *normative survey*.”

3.2 Research Population and Sampling

In this study, a semi-structured face-to-face interview method has been used for the collection of data. Moreover, the study gains insight into respondents' opinions, feelings, emotions and experiences, which are more achievable through qualitative approach. To achieve the research objective, the researcher made an effort to collect data from the entire willing populations from the learning center of HiOA's Institution Repositories which have digital journal collections. The researcher requested that staff in charge of these institutions participate in this research.

A number of faculty members and publishers from HiOA have been contacted through e-mail to be a respondent. Based on their reply and availability 3 personnel have been interviewed and 5 of them replied through email and some data from outside experts. The respondents represented the faculty of Social Science (two persons); Faculty of Technology, Art and Design (one person); Publishers from Some proposed journals related to ODA (five persons). Due to some circumstances, it is not feasible to collect the total size of the population and do random sampling, so convenience sampling, defined as a sample upon selection which is appropriate to the convenience of the researcher and is readily available (Denscombe, 2007), was consequently applied for this research.

3.3 Data Collection Instrument

The instrument selected to collect data is online questionnaire. The questionnaire is designed to collect ideas and comments about research finding and issues from respondents at HiOA Library as well as from outside experts. Accordingly, the questionnaire was decided to employ as a data collection tool because it allows to collect data from wide range of institution types spread across wide area relatively inexpensively.

For this study the questionnaire has both open-ended and close-ended questions in English. However open-ended questions allow the respondents to answer in Norwegian to gather more detailed responses. The data collecting tool is divided into three sections (see appendix 1).

Section 1: The general data on respondents and their collections, such as institution types, cataloging databases, the objectives of content enrichment in institutional repositories, the existing Dspace facility. Current metadata schemas for digital collections and materials that are needed for archival preservation.

Section 2: The type of materials the faculty would like to put in institutional repository. This part includes several open-ended questions about the articles and collections that publishers and librarians really want to enrich. The respondents are asked specific questions about the use of metadata elements ODA database. The usage of Dublin Core elements and configuration of metadata registry in Dspace.

Section 3: This part investigates and identify the required qualified Dublin Core metadata for Dspace while enriching contents. The opinions on the problems while doing this. This part mostly provides questions as free space for sharing their opinions and experiences.

The final part in the questionnaire asks for respondent's profile such as name, position/role and email address. The information from respondent is declared to be kept secret and it is only used for further discussion about the study.

3.4 Online Questionnaire

The online questionnaire is increasingly recognized as an invaluable means of data collection (Doherty, 2012) due to such benefits as higher response speed (Adams and Deans, 2000); lower respondent error (Weible and Wallace, 1998); and online surveys provide several advantages to researchers. They can reach a wide range of potential respondents with Internet experience. (Van Selm and Jankowski, 2006)

The data collection method used to conduct this research was a partially structured questionnaire and purposive sampling method. The reason for using a semi-structured questionnaire was to provide an opportunity for the respondents to write their own thoughts in addition to the given alternatives if they have any. Another reason for using a questionnaire-based method for this research was to have more responses than by interviewing a few researchers. The decision to use purposive sampling in this study was taken because the study is a qualitative survey study where the purposive sampling method was thought to be appropriate. Interview by nature is time consuming for both the researchers and respondents. The chosen target groups for this research were also found to be difficult to interview because

of the nature of their work. Thus, taking the Social Science Department of the Oslo and Akershus college of Applied Sciences as the survey return point, online questionnaire using google forms were distributed among the publishers and faculty members from library. To develop the online survey free google forms has been used.

In addition, secondary data were also used to give more description and background information of the subject and about the institution such as online usage data of institutional repository in HiOA, data on number of contributed documents, type of submitted digital document to ODA and search hits from the ODA home page. Furthermore, review of related literature was also conducted.

In short, the process of developing the questionnaire includes the following steps:

- i. Identifying the main research issues.
- ii. Creating research questions
- iii. Reviewing previous studies related to research topic
- iv. Finding data required for research questions
- v. Manipulating structure of questions
- vi. Analyzing the questionnaire by pilot study
- vii. delivering the questionnaires to the informants

3.5 Data analysis Methods

The method for analyzing data is dissertation and narrative analysis. This approach helped to analyze qualitative data relies on the assumption that human experience is shapes, transformed and understood through linguistic representation. The data gathered from online questionnaire and interviews are also qualitative data because all the questions focus on finding out the perception and interpretation of respondents. constant comparative method is used for classifying data. The constant comparative method is a method for analyzing data in order to develop a grounded theory. Glaser and Strauss (1967) suggest that when used to generate theory, the comparative analytical method they describe can be applied to social units of any size.

All data gathered from respondents are online questionnaire and few from interviews has been arranged as precisely as possible to get the whole transparency of opinions during questionnaire and thus minimized the chances of the analysis being biased. After determining all opinions from respondents, the study has categorized of issues in the text in relation to objectives.

Other than these, meticulous literature has been reviewed to identify what practices are going on for the same kind of research in institutional repository. In addition based on the research objectives reviews are explored to see metadata practices in real life.

Finally, the results of the data analysis are illustrated in tables with explanations and a discussion in Chapter 4. Conclusions and recommendations are described in Chapter 5. Out of respect for the anonymity of respondents, the findings are presented without mentioning names or distinguishing characteristics of individuals or institutions.

3.6 Limitations of the research

Some limitations of the research are addressed below:

Firstly, the opinions from informants may be not sufficient for clarifying research issues because the enriching contents in ODA with actual metadata, is hard for informants to interpret many things which are on progression.

Another Limitation is online data collecting technique. The respondents might have become impatient with the online English questionnaire (with 15 questions), because completing the questionnaire required approximately 15-20 minutes. Moreover, distributing questionnaires via email could have affected the response rate. The recipients might have ignored the survey invitation email from an unknown person.

secondly, language is another limitation. All the questions and answers are written in English so that the informant may feel uncomfortable to express the ideas. Furthermore, some technical terms can be difficult for informant to understand.

Finally, some informants are so busy with the work that they might not take enough time to answer the questions or they will refuse to participate in the study.

CHAPTER 4: DATA ANALYSIS AND DISCUSSION

This chapter presents findings collected from the questionnaire. The data gathered from questionnaire are mainly qualitative data because all the questions focus on finding out the perception and interpretation of respondents. A method called constant comparative analysis is used for coding and categorizing data. Constant comparative analysis is “one method that can be used to identify broad themes and patterns, or categories that emerge from qualitative research studies” (Hewitt-Taylor, 2001, p.42).

The first section is the analysis of data collected from the online questionnaires with librarians, faculty members, publishers, library professions and outside experts about the current practices and facilities in HiOA institution repository. The second section of projects aims to explore metadata practices with in ODA based on the analysis of documents describing structures and meanings of metadata elements in the systems. The final section will describe the finding of the research.

The table below gives brief description of informants’ profile. Their names are coded because the confidentiality was assured to them. All their original answers are put in quote. More additional explanations to clarify their words are placed in square brackets.

Informants	Role	Institution
#1	Digital Services Librarian	HiOA Library
#2	Associate Professor	Faculty of Education and International Studies(HiOA)
#3	PhD Candidate	University of Oslo.
#4	Advisor	Pro-Rector Education, HiOA
#5	Master Student	Universal Design of ICT, HiOA

Table 4.1 The profile of Informants

4.1 Analysis of data collected by online questionnaires

The online questionnaire have been delivered to 20 informants who are involved in different roles of ODA projects. Altogether 11 responses were collected from online questionnaires via the online survey taken using google forms. Then, the raw data from google forms was exported in order to correct the data and enable further analysis of the data. The responses from the respondents are grouped into various categories or issues and sub issues based on objectives of this report. The results are ordered as tables with explanations and can be divided into three sections:

The table below represents description of replied informants profile. informants name were kept confidential and answers are in quote.

All the responses are are kept in PDF file. Each questions from the questionnaire is designate with an idea which is illustrative for different answers from informants. And similar ideas are grouped for finding answers on research issues.

4.1.1 Present practices in HiOA' institution repositories ODA.

The questionnaire were asked to present their opinions on important and present metadata practices in IR of HiOA. The materials in the ODA is of open access policy.

#1 *"All materials in ODA is open and freely available in full text of all."*

furthermore

#1 *"main purpose of ODA is to provide to the greatest feasible information of scientific production, the benefit of further research, education, business media and for the general interested citizen."*

Therefore ODA presents open access materials and contributes possible dissemination of research.

The institution repository ODA is operated by Dspace software for education purposes. It is commonly used because of it's easy customization and interoperable capability. Precisely many institution in Norway use the DSpace software for their institutional repository it is easier to share code, no longer necessary to develop own software. DSpace is highly customizable open source software.

#5 "We found that libraries are preserving both born digital and digitized materials including, book chapters, thesis, scientific reports and some other materials without research value".

The libraries give priority to born digital materials than digitally converted objects, and their main concern is to journals, e-books, and thesis databases. The university college libraries emphasize on digitizing thesis, reports, and scientific documents. In all these cases, there is an attitude to provide access to the resources freely (no direct cost) available in the internet. some libraries chose metadata schema based on the default setup of the software.

Respondents felt that they need Institution Repository were they can keep their materials like presentations, lecture slides, streamed/taped lectures, etc. as well as non-reviewed articles, records, images and so on.

#3 "It is easy to access from everywhere and everyone can read it. If we can put all the materials together in an organized way."

I have taken responses of library and publisher personnel about the awareness of Institution Repositories and ODA as well as to pick out their concern to contribute to the IR of University College(ODA). From their response it is seen that it has been found low level of IR awareness, high level of ODA awareness and interest to contribute to ODA from all personnel.

During the interviews with repository personnels were asked to describe the current contents in the repository, and some proposed and current facilities of ODA.

#1 "All peer-reviewed articles prepared by researchers at HiOA are made available in HiOAs IR, ODA, as soon as possible after publishing provided that publishers allows self- archiving and parallel publishing of scientific work. Everyone must self archive their scientific journal articles in Cristin".

#1 "ODA contains peer-reviewed journals articles, doctoral and masters theses of HiOA's students and other scientific documents. Hioa also contains repository that includes academic work and student papers produced by HiOA named Fagarkivet".

All peer-reviewed journals articles written by staff at HiOA should me made available in ODA as quickly as possible after publication, provided that the journals publisher allow self- archiving of the scientific work.(R & D committee proposal is 19/2000).

#1“*Digital services has overall system responsibility for both cristin & the IR ODA, and can provide technical solutions between the systems. This makes it possible to work closely together in terms of solutions, workflow etc.*”

It seems that HiOA’s open institution repository, ODA include peer-reviewed or editorially evaluated scientific publications. All peer-reviewed journals articles prepared by researchers at HiOA will be available in HiOAs IR ODA, after publishing, provided that the publisher allows self-archiving and parallel publishing of the scientific work. Everyone must self-archive their scientific journal articles in Cristin.

4.1.2 Contents for Metadata enrichment for collection in ODA

However, a metadata schema does not specify how to determine the value or description for the defined elements. In Dublin core, for example element “title” is defined as a name given to the resources(DCMI, 2006). It does not say where the name should come from , where to look for the name, & how to determine a ‘title’? The current metadata schema used in open access IR of HiOA is Dublin Core Metadata Sets.(DCMS). The current metadata schema used in ODA is Dublin core Metadata format. HiOA a union catalog of digital resources gets its collection by harvesting from many data providers. These data providers are required to expose their metadata in Simple Dublin Core format. While Dublin Core is a widely-adopted standard, the interpretation and population of the fifteen Dublin Core elements is ultimately up to the providers creating the metadata. Some institutions have resources to ensure high quality metadata.

#5 “*OAI records harvested from the Library of Congress repository have, not surprisingly, highly uniform Library of Congress Subject Headings in the Dublin Core Subject element*”.

However, many institutions incorrectly or inconsistently use the Dublin Core fields. and some informants were asked about the current metadata schema for digital journals in IR and its advantage of using. some respondents replied with some advantages of using Dublin Core.Respondents focus on first three options :

#4“*DC elements are less rigorous content rules, it is widely used and simple to use too.*” This used schema we found are with *flexible and extensible*,

are easier to train and implement. Also allows the Open Archives Initiative(OAI) harvesting of metadata and is supported by digital library products like ContentDM, Encompass etc.

A large and diverse collection of metadata records contains a varying amount and quality of subject information (and sometimes none). In our context for metadata enrichment experiments, it has been found from the DC elements while using some elements like Title, Subject and Description. It can be determined **that** these fields contained the volume of text relevant to determining the Title, subject and Description of a record. Words from the three fields were considered to be equally important because there was no way of knowing (in advance) from which field useful descriptive text might come. Using the combined text from three Dublin Core elements reduces the problem of inconsistent use of individual elements.

Considering one example of article in ODA 2. here record is described by means of sets of metadata and this one runs on DSpace and uses a schema based qualified dublin core. I have taken an example as a reference that represents enriching the contents by cataloguing items(this example represents cultural cataloging) using dublin core in other institution repositories 3.

In this metadata records each record is described by means of a set of metadata, which are clear, precise and complete as possible.

There it seems, multilingualism is recognized and the descriptions can be enriched enough for enabling some semantic use of metadata i.e **DC tags**(in some cases such as on the right part of “**dc.rights**” and “**dc.type**”, due to the use of concepts collected in namespaces),

For instance, there are some elements of the description that can be used to mark applicable dates such as the moment the data was generated (**dc.coverage.temporal**), when the project finished (**dc.date.issued**) or when it was published (**dc.date.available**).

2 https://oda.hio.no/jspui/handle/10642/2373?mode=full&submit_simple>Show+full+item+recordfull+item+record

3 https://addi.ehu.es/handle/10810/7096?mode=full&submit_simple>Show+full+item+record

With reference to the spatial location, we can use several example of the field **dc.coverage.spatial**; it would be favorable to link this term to an mandate record, it would be possible to relate it with more pieces of information about the same enriching item that are located in other repositories.

and **dc.coverage.spatiallink** metadata is applicable to link Google's map viewer; by doing so, as well as creating the link to a map we avoid the need to explicitly state the reference system.

With the implementation of these metadata elements in contents of Institution Repositories we could be able to enrich some contents from digital collections.

4.2 The usage of metadata elements while enriching contents ODA

The usage of metadata elements can be categorized on the basis of ways of customization of metadata elements and usage of qualified dublin core for enriching.

4.2.1 The ways of customization of metadata elements in enriched ODA

I have taken some of the customizing metadata elements ways from the response of informants. some ways that i have found is creating new qualifiers for Dublin Core Metadata Schemata and then developing and adding some schema in ODA.

#3 "This will certainly increase the existing Dublin Core schema in order to preserve similar metadata support on what ODA are using now."

I found some other ways of customizing metadata elements in in ODA .

#3 suggests that " use of additional metadata schema in order is useful to keep valuable data in ODA".

This implies that from these ways of responses it seems that one way is creating new qualifiers for default Dublin Core metadata set, only default Dublin Core metadata set in ODA is used as standard metadata schema for the enriching and new qualifiers might be added to enable Dublin Core element to fit with data elements taken from same ODA. and another way

is using additional metadata schemata and then developing a custom schema in enriched ODA. Elements from different schemata might be combined to create a custom metadata schema in ODA which can be mapped as closely as possible from data elements of original records in ODA.

4.2.2 Usage of qualified Dublin Core in ODA while enriching

Qualified Dublin Core Metadata Set has 15 original elements plus 6 additional refining elements and many qualifiers for each element (Dublin Core Usage Board, 2005). Here few elements of them are used regularly and are count as essential elements while the other elements are rarely used as elective elements. Therefore, the informants were asked in three levels: use, possibly use and Not use, to give their opinions about the usage of elements of qualified Dublin Core in Dspace. The results are presented in the following chart:

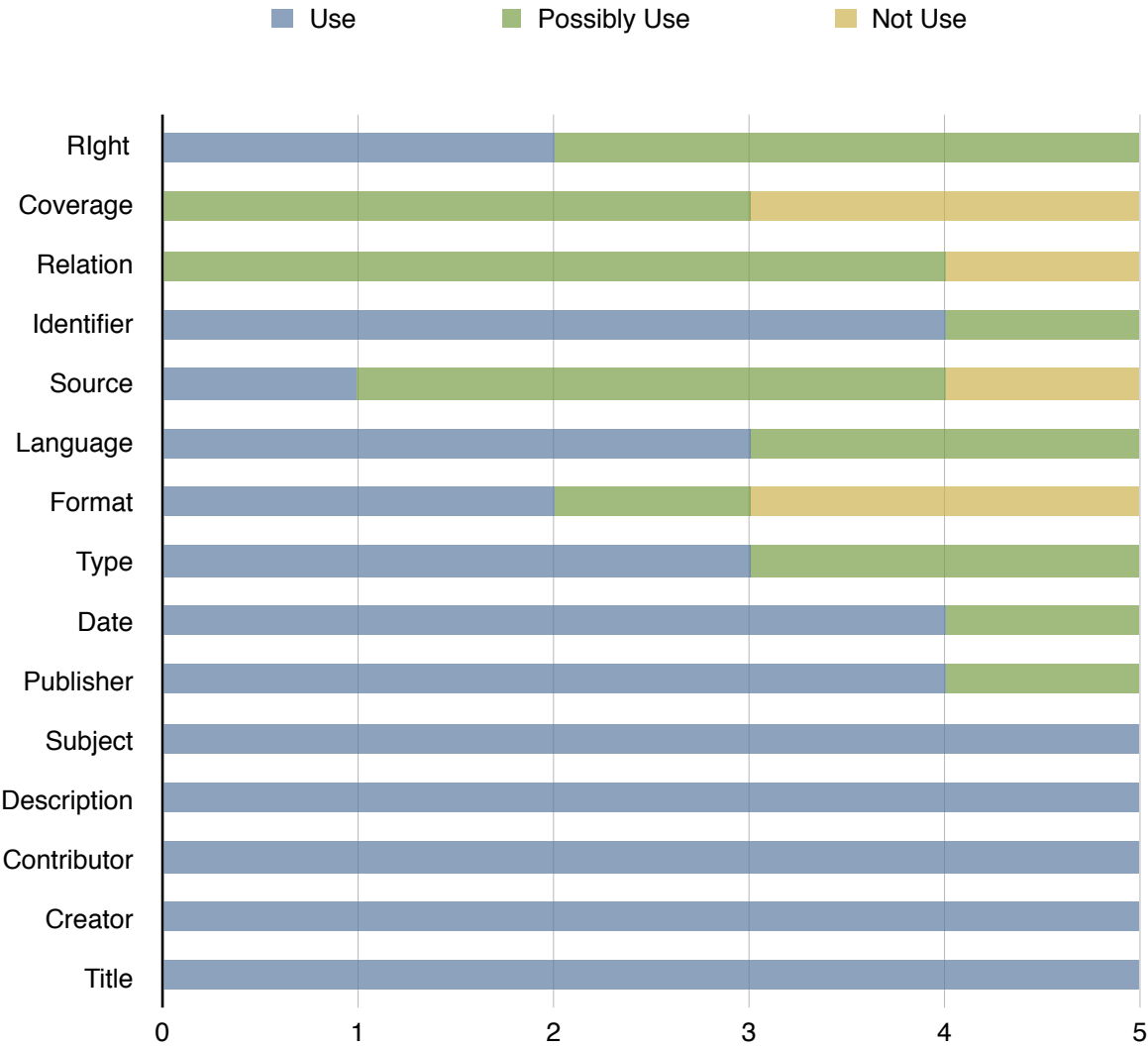


Fig 4.2 Usage of qualified Dublin Core in ODA

From figure 4, most of the elements in qualified Dublin Core receive high “use” support for the data elements transferred from ODA. However, some elements of Dublin Core such as Coverage and relation are much less supported. The data helps to work for selection of metadata elements of Dublin Core for creating a custom schema in ODA.

4.3 The ways how repositories suited into the Scholarly communication

“Scholarly communication refers to the formal and informal processes by which the research and scholarship of faculty, researchers, and independent scholars are created, evaluated, edited, formatted, distributed, organized, made accessible, archived, used, and transformed” (ARL Libraries, 2000). It is convenient to the university community to have the ability to search and use a database that contains an institution's easily accessible and peer reviewed research, such as post-prints, reports, conference proceedings, and monographs. Professional materials that could include preprints, images, audio, video, and respective supporting data can also be valuable to researchers. Faculty and other scholars are able to view and expand on what projects are retrospectively and currently underway, opening opportunities for collaboration and further peer review. It also has the ability to showcase faculty work to prospective higher education students or current students looking for an expert in his or her field of interest to study with. Repositories will form a indefinite and crucial significant part of the scholarly communication process. The primary role of repository is to provide the Open Access publications and literature. Also, we can add services on repositories to equip with added functionality. let’s consider one example, a usage-reporting service provides authors, publishers and the institution by giving information on the content of repository about how they are being used. A service that organizes content in specific ways may help authors, for example, to download a list of articles into their CV, or aid institutions in assessing the institution’s research program or for reporting data to governments or for other statutory requirements. A few scholarly society publishers encourage authors to notify them when a paper has been deposited in a repository and is ready to be peer reviewed and published. IRs also encourage the reform of the scholarly publishing system by supporting the open access movement, which advocates free online access to scholarly materials with minimal restrictions on their use. By providing free access to digital scholarly works at universities, IRs help to realize the goals of the open access movement.

CHAPTER 5: CONCLUSION AND FUTURE RESEARCH

This final chapter of the thesis presents conclusions about the findings of this research. It summarizes the key findings drawn from the interviews and document analysis. It focuses on the main issues learnt from the study. This has been done by answering the research questions in a summarized form as well as pointing the implications of this research and possible future research ideas.

This study adopts a qualitative approach and uses the case study strategy. The literature discussed in chapter 2 revealed that there are gaps in the implementation of enriching metadata standards from theory to practice and as a result has its own challenges from the very aim of enriching metadata in repositories. This study examined the extent of implementing metadata standards and preservation metadata into practice at institutions repository. Identifying the extent to which international metadata standards have been adopted for the enrichment process will allow to analyze the extent of which metadata is used to support the institution repository processes. Therefore, the intent of this study was to add the case study researches that show about the application of preservation metadata standards in to practice along with the problems and challenges in the process and to provide some potential ideas for future research.

5.1 Treatment of the research questions

The major aim of this study has been discussion of the appropriate choices for metadata contents enrichment in HiOAs institution repository ODA To achieve this purpose, three research questions have been formulated at the beginning of the study:

RQ1: What is the existing Dspace facility in HiOA and metadata that are in use?

Sub-questions 1.1 What are the existing Dspace facility in HiOA library.?

HiOA has a department for digital services (the digital library). The department has 4-5 employees and is responsible for the learning center website, the digital publications, the management of shared electronic information resources, and the research documentation for HIOA (Rahman, 2011). The HIOA have an institutional repository named ODA (Open Digital Archives) that includes peer-reviewed journal articles and other scientific documents,

approved theses of master and PhD students from HIOA's own research programs. Faculty members should upload their scholarly journal articles here. This applies to documents published after 01 January, 2010. In those cases where journals do not allow open publishing, or where the co-author does not approve the publication, the publications are stored in a closed archive. IR often contains presentations, historical research conducted at the university that has been converted into digital form, working papers, technical reports, electronic theses and dissertations, and datasets. The Learning center of HIOA also has IR named 'Fagarkivet' for archiving non-peer reviewed materials and other teaching and learning materials that the faculty members have and do not fall with the submission policy of ODA. 'Fagarkivet' to be started by non-peer reviewed materials of HIOA with the DSpace open source software.

Sub-questions 1.2 what metadata standards/schema are used currently for digital Collections/Journals in HiOA Institutional Repository?

The metadata schema used currently is qualified Dublin core it includes elements refinements and encoding schemas. The institutional repository should offer a good number of metadata field as there are different types of need and description addressed by various departments. The minimum required Qualified Dublin core metadata fields have been mentioned in appendix 2. This metadata standard was created to increase interoperability of metadata records, by bridging the differences of the existing objects descriptions. This Dublin core has extensible architecture which is especially important in creating the richer metadata needed for unique user communities because the core 15 elements may be extended by adding additional elements. This helps repositories to develop fields for enriching and cataloging practices. Any element can be refined or qualified. For example, the *Date* element can be refined as date created, date valid, date available, date issued or date modified.

RQ2: Type of materials in Dspace the faculty, publishers and librarians would like to keep in the repository and they want to enrich.

The faculty would like to have scholarly journal articles in repository materials like conference presentations, lecture slides, streamed/ taped lectures, bachelor theses, non-reviewed articles, chronicles, images, and so on. The faculty members of different faculties have lots of materials and they are preserving it by themselves in a scattered way. We found

that libraries are preserving both born digital and digitized materials including, book, book chapters, thesis, journal published by the organizations, organizational newsletters, manuscripts, images, newspaper articles, preprint and post print of articles, annual report, organizational magazine, old question papers, protocols, scientific reports, bibliographies, abstracts, letters, conference proceedings, monographs, audio-visual materials, reports and some other materials without research value. The libraries give priority to born digital materials than digitally converted objects, and their main concern is to journals, e-books, and thesis databases. The college libraries emphasize on digitizing reports, and newspaper articles.

RQ3: To identify the required qualified Dublin Core metadata elements for Dspace while enriching contents and issues on doing this.

‘cataloguing’ or bibliographic information gets a new identity as ‘metadata creation’ in the digital age. In a sense, metadata is the life of the documents, and obligatory part of any Digital Resource Management system (Rahman et al., 2011). It can be found that the institution repositories contains descriptive metadata (for example, title, subject, etc.), followed by administrative metadata (access privileges, rights, ownership of material). Only a few of them considered technical metadata (information describing the production process or digital attributes of the work) too, while structural metadata (for purposes of linking different parts or units of data) is largely ignored. In ODA of HiOA currently Qualified Dublin Core is used, as it is specified by the OAI-PMH as a ‘lowest common denominator’ format well suited to support harvesting into a commonly structured repository, and therefore, supports discovery interoperability (Jones, 2006). Some misunderstanding has been observed as some libraries indicate about using simple Dublin Core, while they are actually using qualified Dublin Core. For example, ‘Date’ is one of the 15 basic Dublin Core metadata elements. The ‘Date’ is then further specified to identify it as a particular kind of date, for example, ‘dc.date.created’, ‘dc.date.valid’, ‘dc.date.available’, ‘dc.date.issued’, ‘dc.date.modified’ etc.

5.2. Implications of the Research

The implication of this study is that the results can be used for librarians, Learning center/ University or people, or for any one that are interested in developing open access metadata standard, software and tools to notice the application of them at institution repository. The

implementation of theoretical standards to practice is imperfect. From the study one can get the information about the metadata enrichment that could likely happen in the process of implementation of theory into practice at institution repository. This research helps expand our understanding of metadata practices in Institution Repositories in the context of digital collection, library. Enriching the Metadata contents can be benefit to initiatives aimed at a re-conceptualisation of contemporary metadata principles. Designing a new approach for improving the content of metadata was considered as a part of framework developed to address the main challenges identified in the research.

5.3 Future Research Ideas

This survey research can present a summary of opinions at a certain time, it may work out the present situation when the time passes. Additionally, due to limitations regarding the research population and sampling, it is recommended that the same topic should be surveyed again. As a result, the findings would more accurately represent the current state of metadata practices in the Institution Repository, ODA. Further, follow-up interviews should be administered for more in- depth information from the respondents.

It is also recommended that a further study on the needs of standardized subject heading lists be conducted. The study should investigate opinions from various relevant communities such as practitioners, scholars in library and information science and in other related areas.

In addition, it was shown that the indirect and direct metadata enrichment models proposed by metadata enrichment approach could connect the end users to spatial metadata creation and improvement process, and hence enrich the content of metadata through the end users' interaction.

It is conceived as an open access initiative and supported by the institutional repository on which we rely in order to ensure the document retrieval, the visibility and the interoperability of the solution. However, it was found that such topic, especially in connection to institution repository' current metadata practices and problems which is the focus of the present study,

has been inadequately acknowledged in the context of Norwegian University and University college.

Digital information is always at risk of disappearing due to its fragility. Institutional repositories are a powerful service provided by universities and research centers that might help in the difficult task of preservation; anyway there is no point in preserving if there are not any re-use of the data. The repositories are rather easy to use for the creators and they are maintained by the universities and research centers, for these reasons, they are a quick and accessible way of creating contents.

This study has also shown that Institution repository focus on preservation metadata standards/schema to record metadata elements as well as developed their own metadata specifications. It would also be interesting to study the comparison and harmonization of various metadata specifications as well as the cooperation between the many metadata initiatives that have an interest in enriching metadata contents.

REFERENCES

ARL Libraries (2000), "Create change", retrieved oct 30, 2015 available at: www.createchange.org/resources/CreateChange2003.pdf.

Antelman, K . (2004). Do Open-Access Articles Have a Greater Research Impact? retrieved nov 9, 2015 from http://www.lib.ncsu.edu/staff/kaantelm/do_open_access_CRL.pdf.

Arunachalam, S. (2008). Open Access to Scientific Knowledge, retrieved nov 9, 2015, from <http://eprints.rclis.org/13859/>

Arms, W. (2000). Digital Libraries. Chapter 10: Information Retrieval and Descriptive Metadata Arms, W. (2000).

Aalbersberg, IJsbrand Jan, et al. "Elsevier's Article of the Future enhancing the user experience and integrating data through applications." *Insights: the UKSG journal* 25.1 (2012).

Rahman, A. I. M. J., Khatun, M., & Mezbah-ul-Islam, M. (2011). Education for information professionals in Bangladesh: A case study of the University of Dhaka. Retrieved from <http://digitalcommons.unl.edu/libphilprac/678>

Adams, S. and Deans, K.R. (2000). Online Business in Australia and New Zealand: crossing a chasm, in Treavor, A. and Ellis, A. (Eds) *Proceedings of AUSWEB2K, The Sixth Australian World Wide Web Conference*, Southern Cross University, Lismore, 12-17th June, Cairns, pp. 19-34.

Baca, M. (2003). Practical issues in applying metadata schemas and controlled vocabularies to cultural heritage information. *Cataloging & Classification Quarterly*, 36(3-4), 47-55.

Basefsk, S. (2009). The End of Institutional Repositories & the Beginning of Social Academic Research Service: An Enhanced Role For Libraries, retrieved nov 9, 2015 from <http://www.llrx.com/features/socialacademicresearch.htm>.

Borgman, Christine L. From Gutenberg to the global information infrastructure: access to information in the networked world. Mit Press, 2000.

Chau, D. V. (2011). Challenges of metadata migration in digital repository: a case study of the migration of DUO to Dspace at the University of Oslo Library.

Dublin Core Metadata Initiative. (2005). DCMI glossary, retrieved nov 12, 2015 from <http://dublincore.org/documents/usageguide/glossary.shtml>.

Dublin Core Metadata Initiative. (2011a). *Dublin Core metadata terms*, retrieved nov 12, 2015 from <http://dublincore.org/documents/dcmi-terms/>

Dublin Core Metadata Initiative. (2011b). *Mission and principles*, retrieved nov 12, 2015 from <http://dublincore.org/beta/about-us/>

Duval, E., Hodgins, W., Sutton, S., & Weibel, S. (2002). Metadata Principles and Practicalities. *D-Lib Magazine*, 8(4).

Doherty, E. (2012). Broadband adoption and diffusion: A study of Irish SMEs, *PhD Thesis*, University of Ulster, Coleraine.

Ferran, N., Casadesús, J., Krakowska, M., & Minguillón, J. (2007). Enriching e-learning metadata through digital library usage analysis. *The Electronic Library*, 25(2), 148-165.

Glaser, BG. & Strauss, AL. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. New York: Aldine De Gruyter.

Genoni, P. (2004). Content in institutional repositories: a collection management issue. *Library Management*, Retrieved oct 2, 2015, 300-306.

Gilliland, A. J. (2008). Setting the stage. In M. Baca (Ed.), *Introduction to Metadata* (pp. 1-19). Los Angeles, Calif.: Getty Research Institute.

Heery, R., and Anderson S. (2005). Digital Repositories Review, retrieved oct 6, 2015 from <http://www.icbl.hw.ac.uk/perx/sourceslisting.htm#broadlandscape>.

Huddleston, R. (2008) *HTML, XHTML, and CSS*. Indianapolis, IN: Wiley Publishing, Inc.

Intner, S.S., Lazinger, S. S., & Weihs, J. (2006). *Metadata and its impact on libraries*. Westport, CT: Libraries Unlimited.

Iglezakis, I. (Ed.). (2010). *E-Publishing and Digital Libraries: Legal and Organizational Issues: Legal and Organizational Issues*. IGI Global.

Jin Ma, *Metadata*, SPEC Kit 298 (Washington, D.C.: Association of Research Libraries, 2007).

Jones, C. (2007). *Institutional Repositories: Content and culture in an open access environment*. Chandos Publishing. Oxford, England.

Klungthanaboon, W. (2010). *Metadata Practices for Digital Photographic Collections in Archives, Libraries, and Museums in Norway*.

Koutsomitropoulos, Dimitrios A.(2010). *D-Lib Magazine*, retrieved oct 12, 2015 from <http://www.dlib.org/dlib/january10/kout/01kout.html>

Karen Smith-Yoshimura, *RLG Programs Descriptive Metadata Practices Survey Results* (Dublin, Ohio: OCLC, 2007): 6–7, retrieved nov 17 from <http://www.oclc.org/programs/publications/reports/2007-03.pdf>.

Kelly, B. (2006). *Choosing a metadata standard for resource discovery*, retrieved oct 14, 2010 from <http://www.ukoln.ac.uk/qa-focus/documents/briefings/briefing-63/html/>.

Li, C., Han, M., Hong, C., Wang, Y., Xu, Y., & Cheng, C. (2011). *Building a Sustainable Institutional Repository*. *D-Lib Magazine*, Retrieved oct 2, 2015.

Lou McGill,(2010). *Digital Repositories: Jisc InfoNet*, Retrieved oct 5, 2015 from <http://tools.jiscinfonet.ac.uk/downloads/repositories/digital-repositories.pdf>.

Lawrence, S. (2001), *Online or Invisible?*, NEC Research Institute, Princeton, NJ, retrieved oct 30, 2015, available at: <http://citeseer.ist.psu.edu/online-nature01>.

Lynch, C. A. (2003). *Institutional Repositories: Essential Infrastructure For Scholarship In The Digital Age*. *Libraries and the Academy*, Vol. 3, No. 2 , pp. 327–336.

Lynch, Clifford. (2008) “ *Institutional: Essential infrastructure for Scholarship in the Digital Age*”. *ARL Bimonthly Report*, No 226, retrieved oct 9, 2015 from <http://www.cai.org>.

Lynch & Lippincott. (2005). Institutional repository deployment in the United States as of early 2005. D-Lib Magazine, retrieved oct 14 from <http://www.dlib.org/dlib/september05/lynch/09lynch.html>

Lagoze, C. and Payette, S. (2000). Metadata: Principles, practices, and challenges. In A. R. Kenney and O. Y. Rieger (Eds), *Moving theory into practice: digital imaging for libraries and archives* (pp. 84-100). Mountain View, CA: Research Libraries Group.

Lavoie, B. F. (2008). PREMIS With a Fresh Coat of Paint: Highlights from the Revision of the PREMIS Data Dictionary for Preservation Metadata. D-Lib Magazine, Retrieved oct 1, 2015.

Mathes, A. (2004). Folksonomies - Cooperative Classification and Communication Through Shared Metadata, retrieved nov 10 from <http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html>

McDowell, C. S. (2007). Evaluating institutional repository deployment in American academe since early 2005. D-Lib Magazine, Retrieved oct 2, 2015

McCormick, M. J. (2006). Filling Institutional Repositories by Serving the University's Needs. A Master's paper for the M.S. in L.S. degree, retrieved nov 13, 2015 from <http://etd.ils.unc.edu/dspace/bitstream/1901/327/1/monicamccormick.pdf>

Markos, J. *Director, Content Enrichment, Wiley* Content enrichment – the opportunity for societies

Miller, P. (2004). Metadata: What it means for memory institutions. In G.E. Gorman & D.G. Dorner (Eds), *Metadata applications and management* (pp.4-16). Lanham, MD: Scarecrow Press.

National Information Standards Organization. (2004). *Understanding metadata*. MD: NISO Press. retrieved nov 14, from <http://www.niso.org>

OCLC. (2012). Best practices for CONTENTdm and other OAI-PMH compliant repositories: Creating sharable metadata, Version 3.0, retrieved nov 14, 2015 from http://www.oclc.org/gateway/support/best_practices.pdf

- Park, J. R., & Tosaka, Y. (2010). Metadata creation practices in digital repositories and collections: Schemata, selection criteria, and interoperability. *Information Technology and Libraries*, 29(3), 104-116. retrieved nov 12, 2015 from <https://ejournals.bc.edu/ojs/index.php/ital/article/download/3136/2750>
- Rahman, M. A. (2011). Challenges for international students in using electronic resources in the Learning Centre: a case study of Oslo University College (Master thesis). retrieved on oct 9, 2015 from <https://oda.hio.no/jspui/handle/10642/989>
- Rahman, A. I. M. and Bayramova, Elnara(2001),Fagarkivet: a proposed Institutional Repository for non-peer reviewed materials in HIOA,
- Rettig, P.J., Shu Liu, Hunter, N. & Level, A.V. (2008). Developing a metadata best practices model: The experience of the Colorado State University Libraries. *Journal of Library Metadata*, 8(4): 315-339.
- Rodríguez Miranda, Á., Valle Melón, J. M., & Porcal Gonzalo, M. C. (2013). Enriching the content provided by cultural catalogues with data from institutional repositories.
- Shreeves, S. L. (2009). Cannot Predict Now: the role of repositories in the future of the journal. Chapter in the future of the academic journal edited by Bill Cope and Angus Phillips. Oxford, U.K.: Chandos Publishing.
- SMS Marketing: Content Enrichment to Enhance Value (2015). retrieved oct 1, 2015 from <http://www.superiormediasolutions.net/content-enrichment-enhance-value/>
- Van Selm, M. and Jankowski, N.W. (2006). Conducting online surveys, *Quality and Quantity*, 40, 435-456.
- Westell, M. (2006). Institutional repositories: proposed indicators of success, retrieved nov 14 2015 from <http://www.emeraldinsight.com/Insight/viewContentItem.do?contentType=Article&contentId=1558875>
- Weible, R. and Wallace, J. (1998). Cyber research: the impact of the Internet on data collection, *Market Research*, retrieved oct 3, 2015, 19-31.

Zeng, M. L., & Chan, L. M. (2006). Metadata Interoperability and Standardization – A Study of Methodology Part II: Achieving Interoperability at the Record and Repository Levels. D-Lib Magazine, Oct 6, 2015.

Zeng, M. L., & Qin, J. (2008). Metadata. London: Facet.

Zhang, A., & Gourley, D. (2014). Creating digital collections: a practical guide. Elsevier.

APPENDIX 1: DUBLIN CORE METADATA INITIATIVE-DUBLIN CORE QUALIFIERS

(Approved in 2007 by the Dublin Core Usage Board)

Retrieved on October 25th, 2015 from <http://dublincore.org/documents/usageguide/qualifiers.shtml>

DCMES Element	Element Refinement(s)	Element Encoding Scheme(s)
<u>Title</u>	<u>Alternative</u>	-
<u>Creator</u>	-	-
<u>Subject</u>	-	<u>LCSH</u> <u>MeSH</u> <u>DDC</u> <u>LCC</u> <u>UDC</u>
<u>Description</u>	<u>Table Of Contents</u> <u>Abstract</u>	-
<u>Publisher</u>	-	-
<u>Contributor</u>	-	-
<u>Date</u>	<u>Created</u> <u>Valid</u> <u>Available</u> <u>Issued</u> <u>Modified</u> <u>Date Accepted</u> <u>Date Copyrighted</u> <u>Date Submitted</u>	<u>DCMI Period</u> <u>W3C-DTF</u>
<u>Type</u>	-	<u>DCMI Type Vocabulary</u>
<u>Format</u>	-	<u>IMT</u>
	<u>Extent</u>	-
	<u>Medium</u>	-
<u>Identifier</u>	-	<u>URI</u>
	<u>Bibliographic Citation</u>	-
<u>Source</u>	-	<u>URI</u>
<u>Language</u>	-	<u>ISO 639-2RFC 3066</u>
<u>Relation</u>	<u>Is Version Of</u> <u>Has Version</u> <u>Is Replaced By</u> <u>Replaces</u>	<u>URI</u>

	<u>Is Required By</u> <u>Requires</u> <u>Is Part Of</u> <u>Has Part</u> <u>Is Referenced By</u> <u>References</u> <u>Is Format Of</u> <u>Has Format</u> <u>Conforms To</u>	
<u>Coverage</u>	<u>Spatial</u>	<u>DCMI Point</u> <u>ISO 3166</u> <u>DCMI Box</u> <u>TGN</u>
	<u>Temporal</u>	<u>DCMI Period</u> <u>W3C-DTF</u>
<u>Rights</u>	<u>Access Rights</u>	-
	<u>License</u>	<u>URI</u>
<u>Audience</u>	<u>Mediator</u> <u>Education Level</u>	-
<u>Provenance</u>	-	-
<u>Rights Holder</u>	-	-
<u>Instructional Method</u>	-	-
<u>Accrual Method</u>	-	-
<u>Accrual Periodicity</u>	-	-
<u>Accrual Policy</u>	-	-

APPENDIX 2: THE ONLINE QUESTIONNAIRE

1. Are you aware of Institution Repository(IR) ?

- yes
- No

2. What do Open Digital Archive(ODA), the Institutional Repository of HiOA Contains ? (Please select all that apply)

- Approved theses of students from HiOA's own master's degree programmes.
- Monographs & Books
- Approved doctoral theses of students from HiOA's own research programmes.
- All of above
- Scientific Documents
- Peer-reviewed journal articles and other scientific documents registered in Cristin.
- Other: _____

3. What are the existing Dspace facility in HiOA repository ?

Your answer _____

4. What metadata scheme are used currently for digital collections/Journals in HiOA Institution Repositories ?

- MARC
- Dublin Core Metadata Element Set (DCMEs)
- Encoded Archival Description (EAD)
- Other : _____

5. Why the above mention schema is chosen? (Select all that apply)

- It is flexible and extensible
- It is widely used
- it is simple and easy to use
- It is supported by leading organizations
- It supports information sharing
- Previous experience
- all of those
- Other: _____

6. What kind of materials do you like to keep and share in an institution repository?(select all that apply)

- non-peer reviewed articles
- class lectures
- notes
- presentations
- all of those
- Other: _____

7. How do repositories fit into the scholarly communication aspects?

Your answer _____

8. what kind of materials do you like to preserve for long time or for archival preservation?

Your answer _____

9. what kind of policy do you like to have in the repository?

- Open Access
- Close Access
- Other: _____

10. Please rate your awareness of the importance of content enrichment in institution repositories?

- Very Important
- Somewhat unimportant
- Very Unimportant
- Somewhat important

11. Metadata is a critical part of enriching articles. why? *

Your answer _____

12. What kind of articles that publishers and Librarians want to enrich?

Your answer _____

Which metadata elements of qualified Dublin Core in enriched ODA will the Library Use? *

	Definitely Use	Maybe Use	Won't use
Title	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Creator/Author	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Contributor/Co-author	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Description/Abstract	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Subjects/Keywords	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Publisher	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Date	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Type (image, sound, text...)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Format (physical/digital form of object)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Language	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Source (where content is derived)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Identifier (URL, ISBN, DOI,...)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Relation (part/version of)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Coverage (spatial/temporal topic in object)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Rights (license)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

14. Please specify other ideas or explain more your choice about metadata elements of qualified Dublin Core in Dspace the library can use ?

Your answer

15. What do you think the Challenges of Metadata Practices for Enriching articles in ODA ?(Select all that apply)

- To decide which metadata standard to use.
- Time and cost
- Consistency
- Interoperability
- Knowledge and Skills
- all of those

16. If you have more ideas and suggestions about my topic, please feel free to write here.

Your answer

17. Thank you very much for your kind contribution. Please kindly provide your contact information. (Your identification is kept secret). *

Name, Position, Email Address

Your answer
