Universidad de Alcalá

Departamento de Ciencias de la Computación

DOCTORAL THESIS

METADATA QUALITY ISSUES IN LEARNING REPOSITORIES

Author:

Nikos Palavitsinis

Supervisors:

Dr. Nikos Manouselis

Dr. Salvador Sánchez-Alonso

November 2013

Abstract

Metadata lies at the heart of every digital repository project in the sense that it defines and drives the description of digital content stored in the repositories. Metadata allows content to be successfully stored, managed and retrieved but also preserved in the long-term. Despite the enormous importance of metadata in digital repositories, one that is widely recognized, studies indicate that what is defined as metadata quality, is relatively low in most cases of digital repositories. Metadata quality is loosely defined as "fitness for purpose" meaning that low quality of metadata means that metadata cannot fulfill its purpose which is to allow for the successful storage, management and retrieval of resources. In practice, low metadata quality leads to ineffective searches for content, ones that recall the wrong resources or even worse, no resources which makes them invisible to the intended user, that is the "client" of each digital repository.

The present dissertation approaches this problem by proposing a comprehensive metadata quality assurance method, namely the Metadata Quality Assurance Certification Process (MQACP). The basic idea of this dissertation is to propose a set of methods that can be deployed throughout the lifecycle of a repository to ensure that metadata generated from content providers are of high quality. These methods have to be straightforward, simple to apply with measurable results. They also have to be adaptable with minimum effort so that they can be used in different contexts easily. This set of methods was described analytically, taking into account the actors needed to apply them, describing the tools needed and defining the anticipated outcomes.

In order to test our proposal, we applied it on a Learning Federation of repositories, from day 1 of its existence until it reached its maturity and regular operation. We supported the metadata creation process throughout the different phases of the repositories involved by setting up specific experiments using the methods and tools of the MQACP. Throughout each phase, we measured the resulting metadata quality to certify that the anticipated improvement in metadata quality actually took place. Lastly, through these different phases, the cost of the MQACP application was measured to provide a comparison basis for future applications.

Based on the success of this first application, we decided to validate the MQACP approach by applying it on another two cases of a Cultural and a Research Federation of repositories. This would allow us to prove the transferability of the approach to other cases the present some similarities with the initial one but mainly significant differences. The results showed that

the MQACP was successfully adapted to the new contexts, with minimum adaptations needed, with similar results produced and also with comparable costs. In addition, looking closer at the common experiments carried out in each phase of each use case, we were able to identify interesting patterns in the behavior of content providers that can be further researched.

The dissertation is completed with a set of future research directions that came out of the cases examined. These research directions can be explored in order to support the next version of the MQACP in terms of the methods deployed, the tools used to assess metadata quality as well as the cost analysis of the MQACP methods.

Keywords: Metadata, Quality, Repositories, Assessment, Tools, Methods

Acknowledgements

It seems really amazing how someone can go on and on, mumbling for almost 300 pages about metadata, but get completely stuck when called upon to write two pages of acknowledgements! I am trying really hard to summarize the journey of the past five years, into a couple of cohesive and accurate sentences. I cannot do it on my own; not without borrowing some words at least. Konstantinos P. Kavafis said it already in his poem "Ithaca": "*Ithaca has given you the beautiful journey. Without her you would not have set out on the road. Nothing more does she have to give you*". My PhD study, has been the perfect analogy to my personal Ithaca that I finally reached. An Ithaca not only related to the PhD study itself, but also an Ithaca of an endless list of personal ups and downs that made these five years a period of my life that I will never forget.

To begin with, during this journey I have had the priviledge to work with many inspiring and above all, generous colleagues and friends that did more than they had to, asking less than they deserved in return. **Hannes** (Ebner), you' ve been an enormous help in the first and most important experiment of my PhD with all the technical support. But above all, you have shared with me the agony for my PhD as I did with yours. Beers and laughing sure did the trick for me and I hope that you are also completing your thesis as we speak. **Kostas** (Makris), thank you for being there, despite the hectic schedule, to provide me with data from the Natural Europe project on demand, no matter the time or day of the week. **Vassilis** (Protonotarios), thank you for deploying the peer review experiments in a couple of cases, but most of all thanks for taking the time to understand and apply all the crazy ideas I came up within the projects you worked. **Miguel** (Refusta), thanks a lot for replying to all the requests for data from VOA3R from this crazy Greek (that's me), even if they went a bit outside your job description. **Charalampos** (Thanopoulos), thanks for allowing me to meddle with your workshops and sessions, to test my assumptions and theories. **Effie** (Tsiflidou), thanks for collecting data, presenting work and analyzing data for this final part of the PhD study. I hope that this has been a useful experience for you and also that I can return the favor helping out with your dissertation. **Miguel-Angel** (Sicilia) thank you for supporting my decision to start a PhD with the University of Alcala de Henares some time ago. Most of all, thank you for knowing when to intervene, silently and gracefully to keep me in track with my obligations. **Pythagoras** (Karampiperis), thanks for "tricking" me into publishing my first journal paper during a time that I was really over-researching instead of just writing! I always though of myself for smarter than that, but you managed to trick me, so hats-off to you! **Salvador** (Sanchez-Alonso), words cannot express the gratitude I feel for all the honest feedback, kind support but also hard

mentorship that got me through this experience in one piece! Thank you for keeping the balance for me, and I am sorry if sometimes I have been late with my obligations. **Nikos** (Manouselis), we have known each other since March 2005 and I know that the fact that I recall this, is a bit scary. If I had to complete a LOM metadata record to describe you, under the element Lifecycle.Contribute.Role, the values I would assign would be: “family”, “friend”, “colleague”, “mentor” and “boss”, in this exact order. I know that supervising me, has been a pretty hard “exercise” on you, because of all these connections between us. Especially for that, and knowing how hard I can be to handle, I owe you a big thank you for the fantastic job you did. Through hard decisions, problems and revisions of PhD topics, you always offered your experience and guidance, but mainly instilled in me a way of working and dealing with problems that will follow me through the course of my professional and personal life from this day on. Thank you.

Apart from the “scientific” contributors to the PhD work, there are also the people that had to put up with me on a daily basis during the times of crisis and doubt, when I became less than bearable. **Mom** (Katerina Vayena), you have shown me how it is to find peace in chaos for as long as I can remember. You have proven to me that nothing is impossible and I know that you are the only person in my life that believes there’s nothing I cannot do, making me believe it myself. **Dad** (Giorgos Palavitsinis), I know you’re still trying to figure out what on earth I am studying! You have balanced Mom’s chaos with order and method and you should know that this is what keeps me in track, most of the times. **Alex** (Palavitsinis), music is what I owe to you. Your music, your performances and success through these years have given me equal amounts of joy as the completion of this PhD does, really. You have been my reality-check to what really matters. **Dimitris** (Gogas), thank you for listening to all the nagging and complaints and for changing the topic to something more pleasant, each time! **Giannis** (Antoniou), thank you for being the best teammate I ever had, both in the court, helping me to blow off the steam, but mainly outside the court, being among the few that really understood and managed to help me through this. **Yiouli** (Dr. Kritikou), I was lucky to “find” you, in a workshop, half way through my PhD adventure. I was fortunate to have you close, sharing the burden and your PhD experiences every step of the way. Words are too limited to describe the ways in which you helped me. You sacrificed more than you had to, and this made the PhD way easier for me to handle. Thank you for being, among other things, my personal doctor for all of this…

Pappou (Nikos Vayenas), I have inherited your name and a couple of characteristics and character traits to go with it! You could have never known that I started a PhD or that I finished it, let alone have anything to do with it. Yet you did, as you have always been there, in flesh and in spirit, from the day I was born, supporting and caring for me as you did for all of us.

Setting the scene for the remaining of this thesis, I would like to let you know that the biggest part of what you are about to read, was written mainly after midnight, in two places. In Yiouli's living room, sat in front of the big blue sofa and in Kimolia Art Café, sat in the corner table of the middle room, with the company of my mother's paintings.

Coming to the end of the "thank you's", I have to admit that holding the PhD printout in my hands was definitely worth the effort. I am not sure about the financial or professional implications of having it but when doubts like this come to mind, I turn to Kavafis again for the answer: "*And if you find her poor, Ithaca has not deceived you. Wise as you have become, with so much experience, you must already have understood what Ithacas mean*".

Table of Contents

List of Acronyms

AP ..... Application Profile

CHO ..... Cultural Heritage Object

DC ..... Dublin Core

ESE..... Europeana Semantic Elements

IEEE..... Institute of Electrical and Electronics Engineers

ISO ..... International Organization for Standardization

LO..... Learning Object

LOM ..... Learning Object Metadata

LOR..... Learning Object Repository

LORI..... Learning Object Review Instrument

MMAT ..... MultiMedia Authoring Tool

MQACP ..... Metadata Quality Assurance Certification Process

MUS ..... Metadata Understanding Session

QA ..... Quality Assurance

W3C..... World Wide Web Consortium

List of Figures

List of Tables

1. Introduction

In the introductory part of the present thesis, the generic problem that it attempts to tackle is presented. To this direction, a short description of the research areas that are directly connected with this work is provided. To continue with, the generic research problem is analyzed into research questions that are explored within the thesis. Specific indicators connected to the research questions are presented and the specific contribution of this work is defined. Closing this chapter, the structure of the thesis is presented.

1.1. Introduction to the Problem

On the notion of quality, John Ruskin (English Writer and Critic of art, architecture, and society, 1819-1900) stated that, "Quality is never an accident; it is always the result of intelligent effort". In this sense, achieving high quality in any product, service or situation in general, requires practice, determination and intelligence. Defining quality in related literature, the concept "fitness for use" is adopted, taking into account the consumer viewpoint of quality as it is the consumer that will judge whether or not a product is fit for use (Juran & Gryna, 1980; Juran, 1989; Dobyns & Crawford-Mason, 1991). Therefore, quality in the case of data can be defined as "data that are fit for use by data consumers" (Wang & Strong, 1996). The need for the existence of processes that ensure data quality has been realized in the case of digital repositories by related studies (Barton et al. 2003; Stvillia et al. 2004; Margaritopoulos et al. 2008; Tate & Hoshek, 2009; Cechinel et al. 2011 and Clements & Pawlowski, 2012) that in many cases proposed well-defined metadata quality assurance frameworks to address this issue. A thorny issue regarding quality lies in the fact that there are not always absolute measures of quality which may be achieved when a specific rating passes a threshold (Harvey & Green 1993).

Looking at some indicative cases of digital collections, libraries and repositories, it is evident that there is a significant problem with the quality of metadata describing the resources that these infrastructures store. Stvilia et al. (2004) assessed the quality of metadata records within 16 collections of 155.000 metadata records coming from academic and public libraries, museums and historical societies. A manual metadata inspection on a small sample showed: lack of completeness, redundant metadata, lack of clarity, incorrect use of Dublin Core (DC) schema elements or semantic inconsistency, structural inconsistency and inaccurate representation. All of the examined records were incomplete as none of them used all 15 DC elements. 94% of the records contained elements with duplicate metadata whereas most of the date elements were ambiguous. Incorrect use of the DC elements was also common. Almost a half of the sample had consistency problems. Shreeves et

al. (2005) looked at a case of a large academic library that contained 14.425 records, as well as, a small academic library and public library collaboration that contained 1.599 resources. Taking a basic set of eight (8) DC elements, the authors measured completeness and found that the first collection contained 71% of incomplete records whereas the second contained 43% of incomplete records, meaning that they lacked at least one of these eight DC elements. Stvilia et al. (2007) examined 27.444 records from an aggregated digital collection with resources coming from museums, libraries and schools. The findings of this study reported that 56% of the records contained contradicting values of the same elements (termed as ambiguity), 54% of the records repeated elements containing the same values (redundancy) and 82% of the records contained inconsistent formatting or representation of the same elements (inconsistency).

Similar studies from other domains have identified such problems. Yen and Park (2006) focused on the National Science Digital Library (NSDL - http://nsdl.org/) Metadata Repository to find out that although many of the main metadata elements were well populated (such as Description, Title, Type and Subject), other elements (such as Relation, Rights, Language and Coverage) were quite poorly completed. Najjar et al. (2003), studied a small portion of the ARIADNE federation (http://ariadne.cs.kuleuven.be/finder/ariadne/) records (3.700), for their completeness to find that only out of eleven educational metadata elements used in the ARIADNE Network, only one educational data element is consistently used (Granularity) among all repositories in the federation whereas four other elements are used in maximum 50% of the cases. Completeness for the rest of the elements was reported even lower. Friesen (2004) took samples summing up to 250 metadata records from five major learning repositories to find that although some elements were used in most cases, elements such as Keyword, Educational Context, Typical Age Range and Rights are not used often. Finally, Ochoa et al. (2011) examined the GLOBE federation (http://globe-info.org/) repositories looking at a set of 630.317 metadata records in terms of their completeness. The main finding of this study is that only a small fraction of the metadata element of the standard adopted in this network is frequently used to describe the learning objects. Not more than 20 out of the 50 data elements are used more than 60% of the records.

As we see in the examined cases, the quality of the produced metadata records is not as high as someone would expect and there is a clear need for putting in place appropriate quality assurance (QA) procedures, methods and tools. Apart from the automated mechanisms of measuring information quality that can easily be scaled up and maintained, there is always the involvement of the human factor that adds value to the process but also includes high costs and is not scalable (Ballou & Tayi, 1999). Duval et al. (2002), Marshall & Shipman

(2003) and Treloar & Wilkinson (2008) point out the high costs of populating databases with metadata, and they all agree on the existence of strong economic incentives to create metadata with sufficient detail to meet the functional requirements of an application or domain.

Overall, from looking at the aforementioned cases, it is evident that problematic metadata quality is an issue faced by multiple cross-domain repositories. The process of creating high-quality metadata heavily involves human annotators that are inherently limited and do not scale up when the digital resources hosted in these repositories increase. Automated methods can only provide partial solutions to this problem and for specific elements. Elements that are domain specific and require examination of the digital content before being provided, have to be examined and treated using quality assurance processes such as the one suggested by our work.

1.2. Basic Concepts/Background

1.2.1. Digital Objects

A generic definition of a digital object is this of *an instance of an abstract data type that has two components, data and key-metadata* (Kahn & Wilensky, 2006). Essentially data of the digital objects refer to bit strings that form files such as photos, documents, videos etc. used in different contexts such as education, research or academia. The part of a digital object that is of great importance for the purpose of this thesis, is metadata.

The introduction of digital objects has its roots in the computer science field. Creating separate components that can be reused and remixed has been the premise of programming in order to save time when modifications were to be made in complex and lengthy code. Digital objects are usually distinguished based on the purpose they fulfill. In this sense, digital objects that are used for enabling and facilitating the use of educational content online are termed as *learning objects* (McGreal, 2004). Starting from this definition by McGreal (2004), digital objects that are used for enabling and facilitating the use of cultural content online are termed as *cultural objects*. In literature, digital objects are comprised from (a) the data that make up the object itself and (b) the metadata that describe the objects and its envisaged use. Drawing from the various purposes that a digital object may fulfill, there is a need for metadata that describe each object in context. This means that different metadata will be used to describe digital objects that are used in different contexts.

The research area around digital objects is quite rich, as various authors deal with many different aspects of them. On one hand, a strand of research (Govindasamy, 2001; Downes, 2001; Boyle, 2003; Brandley & Boyle, 2004) focuses on the practical aspect of designing digital objects in various contexts.

Such papers provide steps and guidelines for design, supporting experts to produce effective digital objects of learning. On the other hand, standardized processes are proposed for the quality assurance of digital objects, by introducing criteria that can be used to assess them (Kumar et al, 2005; Krauss & Ally, 2005; Leacock & Nesbit, 2007; Kay & Knaack, 2009). These authors provide the tools that can be used to assess the success and actual effectiveness of the digital objects produced. Also, a number of authors present tools that can be used to author digital objects in various contexts (Brusilovsky, 2003; Britain, 2004; Hoermann et al., 2005; Liu et al., 2005; Griffiths et al., 2005). These tools are based on learning theories and software design principles in order to facilitate the process of creating learning objects. Finally, a strand of research examines the intricate mechanisms that allow a digital object to fulfill educational or other purposes by looking into the theory of learning design for digital objects (Dalziel, 2003; Koper, 2005; Towle & Halm, 2005; Koper, R., & Tattersall, 2005).

For the purpose of this thesis, digital objects are dealt with, in relation to their purpose in various communities of users. We looked at research on digital objects that focused on the metadata that is used to describe them. Therefore, in the literature review chapter, only the papers that are related to metadata in digital objects are concerned, avoiding an exhaustive coverage of other research papers on digital objects.

1.2.2. Digital Repositories

An increasing number of research papers during the past years refer to the deposited digital content, using the term “digital repositories”. The question here should be how exactly is a repository differentiated from a traditional directory or database. As Heery & Anderson (2005) claim, for a digital repository to be different from any other online collection, the following have to apply:

- Content is deposited in a repository, whether by the content creator, owner or third party,
- The repository architecture manages content as well as metadata,
- The repository offers a minimum set of basic services e.g. put, get, search, access control,
- The repository must be sustainable and trusted, well-supported and well-managed

According to the same work, repositories, no matter the field of focus, (education, culture, science, etc.) share some common objectives that guide their operation and existence:

- The offer enhanced access to resources
- They introduce new modes of publication and peer review

- They offer corporate information management (records management and content management systems)
- They empower data sharing (re-use of research data, re-use of learning objects)
- The aid to the preservation of digital resources

Kahn & Wilensky (2006) define a repository as a "*network-accessible storage system in which digital objects may be stored for possible subsequent access or retrieval. The repository has mechanisms for adding new digital objects to its collection (depositing) and for making them available (accessing), using, at a minimum, the repository access protocol. The repository may contain other related information, services, and management systems*".

As in the case of digital objects, the research area around digital repositories is quite extensive. A number of authors provide guidelines on how to build a repository (Crow, 2002; Higgins & Inglis, 2003; Yang & Tsai, 2003; Palmer et al., 2008; Horstmann, 2009), focusing on the practical aspects that have to be considered in the repository design. Following the design of a repository, a significant consideration is the system that will be used to deploy it which is also heavily researched (Smith et al., 2003; Tansley et al., 2003; Barton & Waters 2004; Prudlo, 2005). On the other hand, a strand of research focuses on the lifecycle of the repositories and different phases of their evolution (Rosemann, 2000; Sicilia et al., 2005; Higgins, 2008; Greenberg, 2009) pointing out the stages that repositories go through. Looking at the performance of repositories, a number of authors suggest metrics and frameworks that guide the evaluation of repositories (Wei et al., 2002; Venturi & Bessis, 2006; Dobratz et al., 2007; Thomas & McDonald, 2007). In these studies, specific metrics, tools and methods are discussed that allow repository managers to assess the success of the deployed repositories. Finally, repositories are, in general, deployed in a number of contexts showing their increasing importance as infrastructures for knowledge management and preservation. Cases of repositories have been identified in fields such as medicine (Higginbotham, 2001, Brindis et al., 2001), public administration (Tzikopoulos et al., 2007; Charalampidis & Askounis, 2008), higher education (McGreal, 2002; Lynch, 2009), art and culture (Patra, 2008; Kounoudis et al., 2010), agriculture (Cebeci et al., 2008; Manouselis et al., 2009), architecture (Stefaner et al., 2007; Kruchten, 2006) and many more.

For the scope of this thesis, repositories are defined and specific case studies that describe their development are examined. In most of the cases, papers that were reviewed, dealt with metadata in one way or another, thus incorporating metadata closely into the digital repository development process. Studies that deal with repositories in a broader or more theoretical way, were considered as irrelevant and where not studied in the context of this thesis.

1.2.3. Metadata

Metadata has been defined as information about information (or data about data), and can be usually described as a data record that contains structured information about some resource. The structure of the metadata records aims to facilitate the management, discovery and retrieval of the resource they describe (Al-Khalifa & Davis, 2006). In addition, the existence of metadata about resources allows potential users to find out more about the resource without having to first examine it (Haase, 2004).

Metadata is made up of data items that are associated to the resource, the so-called metadata elements. Metadata schemas (or metadata models) are sets of metadata elements designed for a specific purpose, such as describing a particular type of resource. When they reach a stable implementation and adoption phase, and are endorsed by some particular community or organization, they often evolve to metadata specifications: well-defined and widely agreed metadata schemas that are expected to be adopted by the majority of implementers in a particular domain or industry. Then, when a specification is widely recognized and adopted by some standardization organization (such as W3C or ISO), it then may become a metadata standard.

Despite the existence of numerous metadata standards, there is no one all-encompassing one to be used in every application. Rather, there are various metadata standards or specifications that can be adapted or "profiled" to meet community context-specific needs (Kraan, 2003). This conclusion has led to the emergence of the metadata application profile concept: application profiles (APs) take one or more base standards or specifications as their starting point. By imposing some restrictions and modifying vocabularies, definitions or elements of the original (base) standard, they tailor the standard to fit the needs of a specific application (Duval et al., 2002). Through this process the aim is to increase interoperability of the AP well beyond the level of the base standard and, at the same time, not to break interoperability with the existing applications that are not aware of the profile (Duval et al., 2006).

Metadata have been deployed in numerous cases in literature as digital repositories have. A non-exhaustive list of cases where metadata were used to describe digital objects are oceanographic data (Han, 2001), statistical data and reports (Yamada, 2004), geospatial data (Devillerset al., 2002; Wayne, 2004; IVOA, 2004), geographical and marine data (Beard, 1996; NDN, 2004) and medical resources (Shon and Musen, 1999; Supekaret al., 2004; Supekar, 2005)

A major issue about metadata is related to who provides metadata. Working with metadata in any digital repository project is a complex and demanding task. A number of stakeholders with different backgrounds are involved: for instance, metadata experts may be concerned with the way metadata will be

presented and used in an information system. As Weinheimer (2000) and Greenberg et al. (2001) suggest, domain experts produce high quality metadata when it comes to their domain area. In the case of digital objects created for learning, educational experts may be concerned with the educational properties of the resources that will be reflected. In the case of digital objects that are meant to be reproduced with the use of advanced visualization devices, technical experts may be concerned providing the metadata to ensure the compatibility of the content with the devices it is meant for. These requirements call for metadata AP development processes that will involve and respect the needs of the various stakeholders, by appropriate and relevant evaluation procedures.

A significant strand of research is focusing on the social metadata that is metadata in the form of tags and annotations that are provided from communities that use the actual digital content. Efforts have been made already to incorporate user-contributed metadata into library catalogues through social features such as tagging, list-making, annotation, ratings and reviews. Overall, it seems unlikely that simple users with no domain expertise, will ever contribute the necessary quality and quantity of metadata that is required for complete digital collections. Another strand of researchers deals with metadata that are created automatically, by means of special software programs. Although this concept holds great promise, it seems that it can lead to errors in the metadata and to failed searches as it is accepted that human intervention is always needed in metadata generation. Overall, it seems that a combination of the aforementioned actors and automated methods will be most appropriate for digital repository projects.

When it comes to working with metadata, creating metadata guides in the form of guiding documents has proven to be really important. Providing annotators with the support in their metadata creation tasks is of paramount importance to support the generation of high quality metadata. In addition to that, it is also needed to involve the communities that will either contribute metadata or use them through the services deployed, in the design of the metadata application profile that will be used. Taking into account the experience and requirements of these communities has been documented in related literature and has generally speaking, led to bigger commitment to the digital repositories deployed.

Finally, an important aspect of metadata lies in the fact that their provision is not only a cumbersome process but also an expensive one. Taking into account the time needed to create a metadata record for a digital resource, in relation to the expertise needed on the metadata annotator side, it becomes clear that any repository project is faced with a significant cost. Taking into account related literature, we see that for one metadata record to be created, approximately 1 hour is needed as a minimum. In the case of a repository of

10.000 resources, this would sum up to almost 72 man months which is an enormous cost on its own, without discussing about even bigger collections of resources.

1.2.4.Quality

Quality can mean different things to different people in different contexts. In this way, if we want to really understand quality, we cannot study it in a vacuum but rather as a part of a given community of practice and a specific product. In quality literature, the concept "*fitness for use*" is adopted, implying the need to look at the consumers of the products we examine (Juran & Gryna, 1980; Juran, 1989). If we look at the generic case of digital repositories and the data they host, quality can be defined as "*data that are fit for use by data consumers"* (Strong & Wang, 1997). Again, if we discuss the meaning of the word "*use*", the definitions that come out of it for different contexts of use might be hundreds if not more. The need for the existence of data quality has been realized in the case of digital repositories already (Barton et al. 2003; Stvillia et al. 2004 and Margaritopoulos *et al.* 2008). But even if we know what quality means in a given context of use for a digital object, still different users of the same object may value the same quality attributes of data in different ways (Strong *et al.* 1997, Kelly et al., 2005). Robertson, (2005) elaborated on this, stating that different settings and purposes require different types of metadata quality. For example, the museum and the archive communities take a different approach to what represents quality in metadata. Museums record extensive detail about the provenance of an object as a necessary part of their purpose. Archives record extensive information but often only at the collection level, rather than object level, level due to the volume of materials they manage. These different purposes have existed side by side within the traditional knowledge management domain with little transference between (Robertson, 2005). The metadata record for the same book will look very different in each setting and no single option is objectively better for each expected use.

Many relevant studies have discussed the shortcomings that problematic metadata annotation can cause, as well as, their overall importance for the success of such systems (Heery and Anderson, 2005; Guy et al., 2004; Robertson, 2005). Greenberg & Robertson (2002) point out that accurate, consistent, sufficient, and thus reliable metadata is a powerful tool that enables the user to discover and retrieve relevant materials quickly and easily and to assess whether they may be suitable for reuse. Poor quality metadata can simply mean that a resource is practically invisible within a collection of resources. In turn, if high quality resources are being hidden because of low quality metadata, this would undermine the efficiency, performance and ultimately necessity of digital repositories.

Quality is interpreted through quality dimensions which are the aspects of quality we would like to measure in metadata. These dimensions are expressed through quality metrics that describe specifically how each quality dimension is linked to them. Finally, quality indicators are the statistical measures and thresholds that express the degree or level of quality metrics. Choosing the appropriate set of quality dimensions, metrics and indicator to be used in each case of a digital repository that seeks to assess metadata quality is of paramount importance. Choosing all possible dimensions and metrics might not be financially feasible. Also, each repository manager has to be aware of what each metric measures and therefore the claims that can be made regarding metadata quality looking at specific indicators. For example, if the percentage (indicator) of completeness (metric) is above 80% for all records in a repository, this does not imply that the records will also be grammatically correct but rather just filled out. As Bruce & Hillman (2004) also make the point that quality cannot be discussed in a vacuum: economic, political and technical constraints are a part of every decision affecting quality and perception of quality.

Attempting to ensure high quality in metadata, or generally in digital repositories, a number of frameworks have been proposed in literature containing predefined steps, methods and tools (Stvilia et al., 2004; Vinagre et al., 2011; Zschocke & Beniest, 2011). Some of the main shortcomings that were documented for quality assurance frameworks were the absence of tangible results of their application or evidence of their applicability in different contexts with similar results. Finally, as it was discussed in the case of metadata annotation, a crucial factor for all the aspiring metadata quality frameworks is the cost. A detailed documentation of the associated costs of a metadata quality framework is really important and apart from cost assessments of specific metadata quality assurance processes, no comprehensive estimation of cost was retrieved in the related literature.

1.3. Research Problem

From the literature review that was carried out, it became obvious that there was little evidence of a tested metadata quality assurance process that proved its effectiveness in one or more repositories. In addition, there we no evidence of any metadata quality assurance process that was both effective and transferable to other contexts, that is, other types of repositories.

1.3.1. Research Questions

To address the research problem that was identified, the present thesis discusses issues related to the quality of metadata records of three different case studies of digital repositories, an educational federation, a cultural federation and a research federation. Trying to address the research problem, we started by attempting to structure a process that would lead us to better

understand metadata quality. Therefore, the first question that this thesis will address is:

1. *Can we develop a modular process that is adaptable and its results are measurable, to support metadata quality assurance in Learning Object Repositories?*

To prove that the quality assurance process that we aimed at producing is indeed useful, we would need to demonstrate some kind of improvement in the metadata record quality of the repositories involved. But how much is enough when it comes to quality of metadata records? Completeness is a measure that ensures that the records will contain data but this does not ensure that this data will be correct. This question brings us to the next questions this thesis will try to address:

2. *Which are the metrics that can be used to effectively assess the metadata record quality?*

3. *At what levels of the metadata record quality metrics is a repository considered to have a satisfying metadata record quality?*

In the literature examined so far, all the quality assurance methods that were identified were targeted towards a specific type of repositories, being customized and contextualized to serve specific requirements. No evidence was found of a quality assurance process that was transferred to any other context with a minimum set of adaptations. This finding led us to the next questions that this thesis will try to address:

4. *Can we transfer a quality assurance process by adapting it for use in a different repository context?*

5. *What are the specific adjustments that have to be made to apply a quality assurance process in other contexts?*

Despite the fact that a quality assurance process may be transferred to other application contexts, this fact alone does not guarantee the success of the process. Parameters such as the size of the repository in terms of resources or the number of individuals involved in the application of the process can change significantly the requirements but also the outcomes of any given quality assurance process. This fact brings us to the next question this thesis will try to address:

6. *Are the results of the application of the same metadata quality assurance process in different repositories comparable in terms of the resulting metadata quality?*

When applying metadata quality assurance processes, a significant number of people are involved, investing time on metadata annotation and enrichment

but also on monitoring of the resulting metadata quality. Through this involvement, a significant cost of time comes up for any repository manager that wants to apply a quality assurance process. This cost has to be documented and analyzed, a fact that brings us to the next questions this thesis will address:

7. *Is the cost involved in the application of a metadata quality assurance process comparable in terms of magnitude for different repositories?*

8. *Is the improvement in metadata record quality comparable with the cost of the quality assurance method?*

In the next paragraph, we examine how the aforementioned questions were translated into goals for the thesis and how these goals are addressed through the contributions of the thesis

1.3.2. Goals and Contribution

To address the research questions of the thesis, a series of relevant goals are set forth and presented in this chapter. To achieve these goals a set of contributions was compiled in the context of our work which helped address the research questions.

1st Goal

The primary goal in order to tackle the questions set forth, was to really understand the basics of knowledge organization. To this end, an initial study was carried out, to examine the knowledge organization systems used in a specific field, that of agriculture. This selection was based on the fact that the first experiment where we would apply the quality assurance process was on the field of agriculture. To this direction we started working on the environmental education domain, researching the way that information related to the environment and in specific agriculture is organized and stored. We also worked on issues related to application profiling of standards as well as issues related to the lifecycle of a digital resource and its metadata within a repository. This work provided us with the practical experience of working with a community of users on metadata concepts and also served as the first stage for deploying the proposed Metadata Quality Assurance Certification Process.

- Palavitsinis, N. and Manouselis, N. (2013). "Agricultural Knowledge Organisation Systems: An Analysis of an Indicative Sample" in Sicilia M.-A. (Ed.), Handbook of Metadata, Semantics and Ontologies, World Scientific Publishing Co.

- Palavitsinis, N., Manouselis, N. (2009). "A Survey of Knowledge Organization Systems in Environmental Sciences", in I.N.

Athanasiadis, P.A. Mitkas, A.E. Rizzoli & J. Marx-Gómez (eds.), Information Technologies in Environmental Engineering, Proceedings of the 4th International ICSC Symposium, Springer Berlin Heidelberg.

- Palavitsinis, N., Kastrantas K. and Manouselis, N. (2009a). "Interoperable metadata for a federation of learning repositories on organic agriculture and agroecology", in Proc. of the *Joint International Agricultural Conference 2009 (JIAC 2009)*, Wageningen, The Netherlands, July 2009

- Palavitsinis, N., Manouselis, N. and Sanchez, S. (2010). "Preliminary Discussion on a Digital Curation Framework for Learning Repositories", in Massart D. & Shulman E. (Eds.), Proc. of *Workshop on Search and Exchange of e-le@rning Materials (SE@M'10)*, Barcelona, Spain, CEUR 681, September 2010

2nd Goal

After defining the context of the study, the next step was to work towards addressing the first research questions (1-3). Having a comprehensive image of the domain, as well as the generic principles of repository development, we introduced parts of the Metadata Quality Assurance Certification Process (MQACP) through focused experiments that used specific metrics to assess quality of metadata records. The first full version of the MQACP and some initial results were presented after completing the majority of the experiments, also defining the thresholds that each quality metric had to surpass to be considered as satisfying. All this effort, from the initial research on the environmental education domain, to the first experiments and the presentation of the findings of applying the MQACP, was presented in the following work:

- Palavitsinis, N., Ebner, H., Manouselis, N., Sanchez S. and Naeve, A., (2009b). "Evaluating Metadata Application Profiles Based on Usage Data: The Case of a Metadata Application Profile for Agricultural Learning Resources", in Proc. of the *International Conference on Digital Libraries and the Semantic Web (ICSD 2009)*, Trento, Italy, September 2009

- Palavitsinis, N., Manouselis, N. and Sanchez, S., (2009c). "Evaluation of a Metadata Application Profile for Learning Resources on Organic Agriculture", in Proc. of *3rd International Conference on Metadata and Semantics Research (MTSR09)*, Milan, Italy, October 2009

- Palavitsinis, N., Manouselis, N. and Sanchez, S., (2011). "Metadata quality in learning repositories: Issues and considerations", in Proc. of

the *World Conference on Educational Multimedia, Hypermedia & Telecommunications (ED-MEDIA 2011)*, Lisbon, Portugal

- Palavitsinis, N., Manouselis, N. & Sanchez, S., (*in press*). Metadata Quality in Learning Object Repositories: A Case Study. The Electronic Library.

3rd Goal

Addressing the next set of questions (4-6) we applied the MQACP in the case of a federation of repositories with cultural content as well as the case of a federation of repositories with research content. Looking at the parameters of each case we see that the MQACP was easily transferable to different contexts with a minimum set of adaptations on the process itself, mainly coming from small differences in the metadata application profiles which had to be factored in when preparing the forms and questionnaires used within MQACP.

An initial analysis of the corresponding results from the two new repository cases showed that similar results for the first phases of the MQACP were generated from domain experts creating metadata records or peer-reviewing them. Since the analysis of the two repositories was recently completed, there is no published work that references the full set of results from the application of MQACP in all three cases. The analysis of the full set of results is only presented within this thesis.

- Palavitsinis, N., Manouselis, N. & Sanchez, S., (in press-b). Metadata Quality in Digital Repositories: Empirical Results from the Cross-Domain Transfer of a Quality Assurance Process. Journal of the American Society for Information Science and Technology.

4th Goal

As regards the last two research questions (7-8), the cost of applying MQACP is also discussed within the work that is referenced below. In this paper we attempted to provide a first discussion on the cost, breaking down the MQACP in the corresponding experiments organized and assessing the cost of each one in terms of time. We tried to assess the magnitude of the cost and how it may be connected to the size of the repositories examined. Extensive discussion and detailed analysis of the cost parameters is also provided within the thesis.

- Palavitsinis, N., Manouselis, N. & Sanchez, S., (in press-b). Metadata Quality in Digital Repositories: Empirical Results from the Cross-Domain Transfer of a Quality Assurance Process. Journal of the American Society for Information Science and Technology.

In total, the results that are presented in this thesis were published in six (6) international conferences, one (1) book chapter and two (2) journals.

1.4. Structure of Thesis

The research that was carried out within this thesis is presented in detail in the next chapters:

In **Chapter 2**, a literature review is carried out focusing on the main research areas that the thesis covers. It provides the main definitions and findings that guided the development of the MQACP and also analyzes the literature that was studied using quantitative and qualitative criteria to provide a better understanding on the domain of digital repositories, metadata and metadata quality altogether. In the end of the chapter, a short overview of the literature focusing on quality assurance processes for metadata is presented.

In **Chapter 3**, the Metadata Quality Assurance Certification Process is presented in detail, elaborating on the tools, methods, material and actors involved in the process. Each stage of the process is described in detail in terms of the method it deploys, the tools it uses and the actors that are involved in it.

In **Chapter 4**, the application of the MQACP on the case of a federation of learning repositories is presented. The experiments carried out in each specific phase are described and their results are presented, explaining the main implications for the quality of metadata records. In the end of the chapter, a short discussion on the overall results is carried out.

In **Chapter 5**, the application of the MQACP on the case of a federation of cultural repositories is presented. We elaborate on the experiments carried out in each phase and we present the results that were collected through the questionnaires and automated methods used. In the end of the chapter, a short discussion is carried out on the overall results.

In **Chapter 6**, the case of the research repositories is presented through the experiments carried out. The results of these experiments are provided and explained, for each given phase of the repository lifecycle. Finalizing this chapter, the main results of the case of the research repositories are presented followed by a short discussion.

In **Chapter 7**, we present the general conclusions and the future directions of this thesis. More specifically, we present the main contributions of this thesis and we re-examine the research questions discussed in Chapter 1. For each question, we discuss the degree on which the question was researched and we describe future directions of research for each one of the research questions that this thesis deals with.

In **Appendix A**, the tools that were used in the literature review are presented.

In **Appendices B and C**, the statistical analyses and tables from the literature review are presented, as well as the distributions of research papers per journal, according to the classification parameters used.

In **Appendix C**, we provide the documentation and instruments that MQACP utilizes in its application.

Finally, in **Appendices D to F**, we provide the data that came out of each experiment and served as the basis of the results presented in the corresponding chapters of this thesis.

2. Literature Review

The aim of this chapter is to provide an overview of the areas of research around the main topics that this thesis deals with. To this end, we focus on digital objects, repositories/digital libraries, metadata and quality of metadata. More specifically, the objectives of this chapter are:

- First, to serve as an introduction to the concept of metadata, digital objects, digital repositories and metadata quality,
- Second, to provide an overview and a classification of existing research literature on the above,
- Third, to extract research directions for the research problem that was described in the previous chapter, that is the focal point of this thesis

The chapter focuses on the quality aspects of metadata and examines quality approaches from existing literature. Nevertheless, it also addresses some generic theoretical aspects related with the digital resources that are hosted in the digital repositories but also with the digital repositories themselves. This chapter contains a review of existing literature that reflects the current status of research on the wider field of metadata and quality of metadata for digital resources and repositories and identifies research trends and issues.

2.1 Methodology

This report is based upon an extensive literature study. Papers related to metadata but also involving digital resources and digital repositories/libraries, from journals that publish related research have been reviewed. More specifically, thirty-nine (39) scientific journals related to Information Science, Libraries and Information Technology have been examined. The review covered mostly publications of the last 15 years (1997-2012) but in several cases, influential publications prior to this date were considered. Almost 1.000 papers have been initially located in this way. We briefly examined each one of these papers in order to filter out ones those who were irrelevant. This process narrowed down the number of papers to 605. We carefully studied each one of these papers and carried out an initial classification based on a short description of each paper and its content. From this set of papers, access to their full content was possible for a set of 506 papers that were contained in 32 out of the initial 39 journals.

As a next step, a set of metadata related to the papers examined was selected, based on which the papers were annotated with metadata containing information such as the date of publication, number of pages, authors, etc. that would then help us categorize and analyze the research carried out in the wider field that is relevant to this thesis. In addition, to provide a meaningful

classification of the papers, apart from the standard metadata that were referenced previously, we adapted a framework (Groenlund, 2004) for the classification and assessment of the literature.

Groenlund (2004) presented an assessment framework of research characteristics, which he applied to assess the state-of-the-art in e-government research. He classified 170 papers from three e-government related conferences in 2003. Groenlund characterizes the proposed framework as an intuitive one, whose purpose is not to compare the status of e-government research with other fields, but to stimulate discussion about this field and measure its maturity. Since we found such an assessment required in the case of learning resources, repositories, metadata and their quality issues, we adapted the proposed framework for the needs of this study. The framework assesses the status and maturity of research publications around a specific topic, by classifying them according to two major dimensions:

- Rigor: depending on the maturity of the field, the balance among methods used is expected to change over time from simple arguments, philosophical discussions and case stories to more methodologically sound examination of relevant issues (e.g. through experiments or ethnological observation). This dimension mainly examines the following aspects: Research Type, Research Method, and Claims.

- Relevance: identifying the fields which this research involves and estimates its current benefits for the practitioners and the society, depends on aspects such as whom this research concerns, whom it is addressing, and from whom it is conducted. This dimension mainly examines the following aspects: Focus unit, Target audience, and Origin (Institution & Discipline).

The extended definition of all the dimensions used to categorize the papers identified in the literature review can be found in “Appendix A: Literature Review Dimensions”. In Figure 2.1, an example of how the research papers retrieved were described with metadata is presented.

A	B	C	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
			RIGOR						RELEVANCE						
No	Paper Title	Author (s)	Research Type	Method	Claim	Number of Pages	Number of Refs	Refs per Pages	Focus Unit	Target Audienc	Institution	Discipline	Number of Auth	Country	Continent
			Descriptive	Argument	Normative	14	26	1,86	Society	Researchers	University	Information Systems	2	USA	
			Theory Generation	Survey	Lessons	6	10	1,67	Organisation	Researchers	Government	Management	1	Belgium	
1	A semantic web architecture for integrating competence management and learning paths	Fotis Draganidis, Paraskevi Chamopoulou, Gregoris Mentzas	Applied (Implementation)	Product Description	Ongoing Research	16	16	1	Method	Researchers	University	Computer Science	3	Greece	Europe
2	Development of personalized learning objects for training adult educators of special groups	Maria Pavlis Korres, Elena García-Barriocanal	Applied (Implementation)	Case Story	Descriptive	13	46	3,54	Group	Practitioners/Implementers	University	Computer Science	2	Spain	Europe
3	Developing a digital preservation strategy at Edinburgh University Library	Najla Semple	Applied (Implementation)	Case Story	Lessons	5	9	1,8	Method	Practitioners/Implementers	University	Information Science	1	UK	Europe
4	Digitizing library resources for new modes of information use in Uganda	Helen M. Byamugisha	Theoretical	Case Story	Normative	16	21	1,31	Method	Practitioners/Implementers	University	Information Science	1	Uganda	Africa
5	How library practitioners view knowledge management in libraries: A qualitative study	Md Roknuzzaman, Katsuhiro Umemoto	Theory Testing	Survey	Descriptive	14	47	3,36	Group	Practitioners/Implementers	University	Information Science	2	Japan	Asia
6	A new look at the university libraries in context: European Research Area	Diana Pietruch-Reizes	Theoretical	Case Story	Descriptive	6	12	2	Sector	Researchers	Research Institute	Information Science	1	Poland	Europe
7	Towards a national print repository for Australia: where from and where to?	Paul Genoni	Descriptive	Argument	Lessons	13	31	2,38	Organization	Managers	University	Information Science	1	Australia	Australia
8	Managing the life cycles of the document and library collections in Finnish academic libraries: Two case studies: Aalto and UEF	Ari Muhonen, Jarmo Saarti, Pentti Vattulainen	Theory Generation	Interpretative	Descriptive	9	9	1	Organization	Managers	University	Information Science	3	Finland	Europe
9	Educating the academic librarian as a blended professional: a review and case study	Sheila Corrall	Theoretical	Literature Study	Normative	27	101	3,74	Individual	Practitioners/Implementers	University	Information Science	1	UK	Europe
10	The economics of repository libraries in the context of the future conventional libraries	Steve O'Connor	Philosophical	Argument	Descriptive	8	3	0,38	Organization	Managers	Research Institute	Economics	1	Australia	Australia
11	OAI compliant institutional repositories and the role of library staff	Lynne Horwood, Shirley Sullivan, Eve Young, Jane Garner	Theoretical	Case Story	Descriptive	7	18	2,57	Organization	Researchers	University	Information Science	4	Australia	Australia
12	Leveraging quality web-based library user services in the digital age	Lili Li	Theory Generation	Argument	Descriptive	11	5	0,45	Sector	Practitioners/Implementers	University	Information Science	1	USA	America

Figure 2.1: Example of annotation of research papers

2.2 Literature Review Outcomes

This chapter provides an introduction to the main concepts of this literature review, focusing on their definition, as well as other parameters that characterize them.

2.2.1 Data, Digital Objects and Types

This paragraph deals with the existing definitions the different types of digital content that were identified during the literature review. Starting from the concept of data in general, we define data as a raw sequence of symbols (Stvilia et al., 2007) or in a more general sense as objects in digital form (Külcü & Külcü, 2010). Stemming from this definition we tried to define what an object in digital form is, that is, a digital object. Looking at the two most prominent definitions we see that a digital object/resource can be termed as:

- A compound object that must have these components: the material, descriptive metadata, technical metadata, an activity/event log, representation information, and a unique identifier (Harvey & Thompson, 2010).

- A resource that is generated through some electronic medium and made available to a wide range of viewers both on-site and off-site via some electronic transferring machine or internet. Saye (2001).

This object-oriented paradigm has its roots in the computer science field. Object-orientation highly values the creation of components (called “objects”) that can be reused in multiple contexts (Dahl & Nygaard, 1966; Cohen et al., 2011). One of the prominent contexts or applications that are researched in literature is learning. The idea of learning objects had already emerged in the early nineties (Persico, Sarti, & Viarengo, 1992). The concept was concerned with storing the learning materials into databases. From then on, the issue that emerged was the reuse of learning materials (Olimpo, Chioccariello, Tavella, & Trentin, 1990; Rada, 1995b; Sarti & Marcke, 1995).

One expression of a digital object that will be researched in this thesis is a learning object in the sense that a learning object is a digital object that is used in order to achieve the desired learning outcomes or educational objectives (Nash, 2005). In the case studies that this thesis covers, there are two main types of content hosted in the repositories examined. On one hand educational content, that is digital objects that are used in learning situations and on the other hand, research/scientific objects that are used by practitioners and researchers in specific scientific fields. Going back to the context-driven definition of a learning object, we perceive research/scientific objects as digital objects that are used in research of various scientific disciplines. The next step is to define in a more accurate way learning objects and

research/scientific objects. Table 2.4, presents the main definitions that were retrieved for learning objects.

Authors	Definitions of Learning Objects
L'Allier, 1997	The smallest independent structural experience that contains an objective, a learning activity and an assessment.
Merrill, 2000	Combined knowledge object and a strategic object representing a mental model to be developed by a learner through incremental elaboration
IEEE LTSC, 2002; Hodgins, 2008	Any entity, digital or non-digital, which can be used, reused or referenced during technology supported learning
Gallenseon et al, 2002	A unit of instructionally sound content centered on a learning objective or outcome intended to teach a focused concept.
Wiley, 2000; Wiley, 2003	Any digital resource that can be reused to support learning
IMS Global, 2003	Atomic or aggregate learning resources
Sosteric & Hesemeier, 2004	A digital file (image, movie, etc.) intended to be used for pedagogical purposes, which includes, either internally or via association, suggestions on the appropriate context within which to use the object
Hummel et al., 2004	To be a learning object a digital resource must be reproducible, addressable (ie. connected with a URL and has metadata), used to perform learning or support activities, and made available for others to use
Cebeci & Erdogan, 2005	Digital units of educational content designed and created for many purposes such as maximizing the number of learning situations, reducing development costs, and exchanging of learning materials between different platforms.
Cochrane, 2005	Interactive digital resource illustrating one or more concepts
Nash, 2005	A digital object that is used in order to achieve the desired learning outcomes or educational objectives.
Harman & Koohang, 2005	Digital resources of any kind that can be similarly combined, shared and repurposed in different educational contexts
Cohen and Nycz, 2006	Knowledge based object that is self-contained and reusable
Varlamis & Apostolakis, 2006; Ouyang & Zhu, 2008	A chunk of digital learning resources or data used for learning and instruction. They are authored, stored, indexed, assembled, delivered, and evaluated
Kay & Knaack, 2008	Interactive web-based tools that support the learning of specific concepts by enhancing, amplifying, and/or guiding the cognitive processes of learners
Mogharreban & Guggenheim, 2008	Context-free digital elements, whether text, audio, video, animation, etc., that become learning objects only when applied within a context of learning

Authors	Definitions of Learning Objects
Alberta Learning, 2002-2008, Online Glossary section	One or more digital assets combined and sequenced to create or support a learning experience addressing a curricular outcome(s) for an identified audience(s)

Table 2.1: Various definitions of learning objects

Looking at the amount of different definitions recovered, we see that the topic of learning objects has been discussed extensively in literature. Filtering the definitions provided above, we came up with a set of statements related to learning objects:

- Learning objects are digital objects that serve an educational purpose,
- Learning objects can be combined amongst them,
- Learning objects can be repurposed to fit different needs,
- Learning objects can be atomic or aggregations of smaller objects,
- Learning objects are designed based on learning objectives and outcomes

Interestingly enough, many of the definitions refer to learning objects as knowledge objects or knowledge based objects which also points out the meticulous educational design that is required for a learning object to be created. Learning objects are not simple digital files or streams of bytes attached to a "manual of use". They are carefully designed and aligned with learning objectives and sometimes specific curriculums to lead the learned through a specific learning path. As Downes (2003) argues, for a digital resource to be classed as a learning object, it must have some instruction inherent in it:

A mere picture is not a learning object because there is no instruction inherent in the picture…The presentation of a picture, therefore, must be accompanied with some context. The context would describe what is to be learned from the picture.

Taking all the above into account, we propose the following definition of a learning object that will be used in the context of this thesis:

A learning object is a digital object (or an aggregation of digital objects) that is designed based on learning objectives and contains information that supports its use within a specific environment. Learning objects can be combined and repurposed for different learning needs in different contexts.

Having articulated the working definition of a learning object, we see that each learning object has to contain specific information that facilitate its use in an educational setting but also enable its repurposing and combination with

other learning objects. According to (Varlamis & Apostolakis, 2006) learning content must be labelled in a consistent way to support the indexing, storage, discovery (search), and retrieval of learning objects. This is achived by describing the learning objects with metadata (Yordanova, 2007). Metadata allow learning objects to be well described and easily searched, assembled in desired learning context, and delivered according to the learners' preferences and needs of education. As a result, users have access to the learning content most appropriate for them, and receive only the necessary and needed information, instead of being overloaded with learning materials that they already know or are not related and appropriate to their expertise, professional background and educational needs (Yordanova, 2007; Cohen et al., 2011).

Attempting to elicit a definition for a research/scientific object, we start from the working definition of a learning object to state that:

A research/scientific object is a digital object (or an aggregation of digital objects) that is the result of research/scientific work and contains information that support its use within a specific environment. Research/scientific objects can be combined and repurposed to address different research purposes in different research domains.

2.2.2 Digital Libraries and Digital Repositories

This paragraph deals with the existing definitions on the technical infrastructures that are deployed to host digital objects. The terms that have been used to describe them vary greatly but the two major ones that were retrieved in relevant literature are "digital repositories" and "digital libraries". Digital repositories are also are also referred to as "learning repositories" describing digital repositories that host learning objects. During our search we have identified a significant conceptual overlap between digital repositories and digital libraries which was also really apparent through some of the definitions collected such as this one:

A digital repository/library is collection of digital objects. It is either a local, institutional, or central (e.g. subject- or discipline-based) archive for depositing and providing access to digital contents. (Patra, 2008)

Looking into more definitions we found that in some cases like in (Baker, 2006) digital libraries were thought to contain the term digital repositories "*An organizational entity that brings together a wide range of (academic) assets, including metadata, catalogues, primary source materials, learning objects, datasets, and digital repositories – in a structured and managed way*". Enhancing that notion, Tsakonas & Papatheodorou (2008) point out that "*the term digital library is vast, covers many and different applications and has been used interchangeably for systems, like digitized collections, e-*

journals platforms, network databases, library websites, etc." and (Cassella, 2010) argues that "*institutional repositories are some of the most innovative and creative components of digital libraries*". On the other hand, in other cases we saw that repositories were a more broad term, containing digital libraries (Zuccala et al., 2006) "*examples of repositories are digital libraries, speciality repositories, institutional repositories, research data repositories and e-learning repositories*".

In order to clearly identify if the term "digital library" is broader than the term "digital repository", or vice versa or if they are just used interchangeably, we have to look closer at the literature. To this end, in Tables 2.2 and 2.3 the main definitions for digital libraries and digital repositories are presented.

Authors	Definitions of Digital Libraries
Lynch, 1995	Electronic information access system that offers the user a coherent view of an organized, selected, and managedbody of information
Lesk, 1997	A collection of information that is both digitised and organised, and which offers capabilities beyond those of the traditional library.
Raitt, 1999 Walters, 1998	Organizations that provide the resources, including the specialized staff,to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of,and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities
Borgman, 2000	A collection of digital objects (such as digital text, images, and videos) and a set of associated techniques and services that help to collect, organise, retrieve, and preserve those digital objects for a community of users
William, 2001	A managed collection of information, with associated services, where the information is stored in digital formats and accessible over a network.
Xiao, 2003	A distributed system that has the capability to store and effectively utilize various electronic documents, which may be conveniently accessed by end users via network transmission
Wang, 2003	Digital information resources center containing multimedia information resources. A digital library exists by digitizing information, such as characters, videos, and audios, and providing users with quick and convenient information services via the internet, to deliver a digital information system in whichshare of resource is available
Bhattacharya, 2004	Managed collections of digital objects, created or acquired according to the principles of collection development, in which information is stored and distributed in digital form with the associated value-added services necessary to allow users to retrieve and exploit the resources just as in a traditional library

Authors	Definitions of Digital Libraries
Gonçalves, et al., 2004	Complex information systems built to address the information needs of specific target communities
Saeed, 2006	Not only a digitized collection with information management tools, it is also a series of activities that include collections, services, and people in support of the full life cycle of creation, dissemination, use, and preservation of data, information, and knowledge
Fuhr et al., 2007	A particular kind of information system and consists of a set of components, typically a collection (or collections), a computer system offering diverse services on the collection (a technical infrastructure), people, and the environment (or usage), for which the system is built
Shiri, 2008	Multifaceted and complex information structures that offer a wide range and variety of information bearing objects. They vary in their content, subject matter, cultural characteristics, language etc.

Table 2.2: Various definitions of digital libraries

Trying to sort out the definitions, we chose to present a list with the main characteristics of a digital library. So, according to the existing definitions, a digital library:

- …contains digital objects grouped in collections,
- …is organized and managed through tools,
- …offers services on the content it hosts like collection management, organization, searching and preservation,
- …provides specialized support from experts,
- …serves communities of users,
- …offers curatorial services on the digital content,
- …may as well be distributed among different systems,
- …hosts digital content from different domains, languages and types

From this initial analysis, it seems that digital libraries are not that different from what is defined as a repository in related literature, apart from the fact that libraries are also concerned with managing the full life cycle of a digital object whereas something like this is not apparent in the case of digital repositories. To confirm or reject this notion, we also analysed the existing definitions on digital repositories.

Authors	Definitions of Digital Repositories
Crow, 2002	Institutional repositories are "digital collections that capture and preserve the intellectual output of a single or multi-university community"
Chang, 2003	A method for capturing, collecting, managing, disseminating, and preserving scholarly works created in digital form by the constituent members of an institution

Authors	Definitions of Digital Repositories
Lynch, 2003	An organizational commitment to the stewardship of . . . digital materials, including long term preservation where appropriate, as well as organization and access or distribution
Duncan, 2003	The purpose of a digital repository is not only to store, catalog, and distribute learning objects, but also to allow their sharing and reuse
Downes, 2004; López 2005; Namuth et al., 2005	A system that manages the access to reusable learning content
Benjelloun, 2005	A digital archive. It is used to amalgamate and diffuse the scholarly publishing produced by faculties, and institution's research staff, in order to make it accessible to users within and outside the institution.
The JORUM Team, 2006	A system that "enables the storage, discovery and retrieval of metadata and/or electronic objects stored at a local or distributed level"
Zuccala et al., 2006	A store where electronic data, databases or digital files have been deposited, usually with the intention of enabling their access or distribution over a network.
Shreeves & Cragin, 2008	A set of services and technologies that provide the means to collect, manage, provide access to, disseminate, and preserve digital materials produced at an institution
Aschenbrenner et al., 2010	The persistent management of networks of files and their associated metadata is achieved by systems called digital repositories
Yen et al., 2010	And an open database that provides data storage, searching, delivery, and exchange functions.

Table 2.3: Various definitions of digital repositories

Overall, from looking at the existing definitions, we found that a digital repository:

- …supports the storage, discovery and retrieval of digital objects,
- …fosters the sharing and reuse of the retrieved digital objects,
- …supports the knowledge management needs of given communities,
- …offers metadata management tools,
- …supports long term preservation of digital objects,
- …offers a set of services on the digital content hosted

Looking at the main characteristics of both digital libraries and digital repositories as these were extracted from literature we see that no significant difference was located between the two terms. Both digital libraries and digital repositories offer added value services on the content, support communities of users and preserve the content hosted over its entire lifecycle. Two differences

that are pointed out, is the fact that repository definitions discussed metadata management explicitly whereas the definition of digital libraries often contained the notion of collections of digital objects but also the library staff as a focal part of the digital library.

Before we propose the working definition of digital libraries and repositories, we also examined existing work on the attributes and characteristics of repositories and libraries to compare and contrast them. First of all, Rachel Heery and Sheila Anderson (2005) have shown that a digital repository is differentiated from other digital collections by the following characteristics:

- Content is deposited in a repository, by the content creator, owner or a third party,
- The repository architecture manages content as well as metadata,
- The repository offers a minimum set of basic services (e.g. put, get, search, access control) and
- The repository must be sustainable and trusted, well-supported and well-managed.

On the other hand, Zhou (2005) argues that the unique characteristics of digital libraries include:

- Mass storage of information resources,
- Information resources in diversified media,
- Network transmission of information resources,
- Distributed information resources management,
- Highly shared information resources,
- Intelligent retrieval technologies; and
- Information services without space and time limitations.

Examining these two opinions and comparing them with our findings, we see that two differences exist between the definitions or characteristics of a "digital repository" and those of a "digital library". The first difference is the fact that digital libraries are being researched and developed mainly from librarians and related professions, whereas digital repositories are developed and researched mainly from computer and information scientists. The second difference is that digital repositories seem to be more specific in the nature of the content they host, to serve specific communities on research fields or topics of interest, a fact that was also pointed out by Koenig and Mikeal (2010).

To sum up, the present thesis will use the term digital repository as a common term in the remaining chapters that also reflects digital libraries. In addition, lots of the lessons and concepts that are introduced in the field of digital libraries are being transferred and discussed as possible solutions in the field

of digital repositories, based on the similarity of characteristics that was defined above. Our working definition of a digital repository/library is:

A digital repository/library is the ecosystem of processes, metadata, people, services and tools that facilitate the storage, discovery, retrieval and preservation of digital objects for a given community of users.

Metadata are an intricate part of the definition, purpose and function of any digital repository. Confirming our argument, Groenewald and Breytenbach (2011) state that any digital object does not have any meaning to a human being unless the content is described with descriptive, structural and technical (or administrative) metadata.

2.2.3 Metadata and Related Issues

This paragraph deals with the existing definitions of metadata that were identified during the literature review. The purpose of this paragraph is to provide the working definition of metadata used in this thesis and identify the main issues that were brought up by the literature review. Connecting with the previous paragraph, in this section we will examine how digital objects can be described with metadata to facilitate their use in different contexts by looking at related literature.

Authors	Definitions of Metadata
Weibel 1995, 1997	Structured data about data. Metadata are simply data about data or the contents of a surrogate record that characterize an object
Dempsey et al., 1997	Data which describe attributes of resources
Burnett, Ng, & Park, 1999	Data that characterizes source data, describes their relationships, and supports the discovery and effective use of source data
Day, 2001	Data about data, information about information, and are used to describe document contents and structure, and to provide information about accessibility, organization of data, relations among data items, and the properties of the corresponding data domains
Greenberg, 2001	Structured data about data that supports the discovery, use, authentication, an administration of information objects.
Guinchard, 2002	Data that records information about a resource in the library community
Miller, 1996; Steinacker et al., 2001; Taylor, 2003; NISO, 2004; Sen, 2004	Data about data or information about information"

Authors	Definitions of Metadata
Taylor, 2004	Structured information that describes the attributes of information packages for the purposes of identification, discovery, and sometimes management
Malaxa and Douglas, 2005	Structured data about data that is helpful to the efficient discovery and reuse of digital assets.
Zhang & Dimitroff, 2005	Metadata provides an effective mechanism for describing and locating data that is relevant to a particular user.
Nash, 2005	Searchable information stored about an object to identify or explain it.
Patra, 2008	Information required to describe, organize and index digital material in ways that allow users to locate and browse it in useful ways.
Cohen et al., 2011	Descriptive information about the resource that describes the learning object and is critical for its sharing and reuse.
Kovasevic et al., 2011	Information that describes the characteristics of objects that are stored in digital repositories.

Table 2.4: Various definitions of metadata

Choosing a working definition for metadata is a straightforward task, as most authors agree on metadata being termed as "data about data". Apart from that, authors seem to also discuss other aspects of metadata in their definitions, such as the structure behind them but also the fact that they characterize the digital object in a way that would allow someone to understand the content of the object itself just by looking at its metadata. In addition, definitions of metadata seem to take into account the functionalities that metadata support that is the description, organization, discovery, use and reuse of a digital object. In this sense, a working definition that will be used in the context of our work is the following:

Metadata is structured data about data that reflect the latter faithfully. Metadata contribute to the description and organization of the data which ultimately leads to their efficient discovery, use and reuse.

Apart from the definitions presented, the literature review brought up some important issues about metadata that are discussed below, enhancing the background of this thesis and providing some useful insights that guided the experiments planned.

2.2.3.1Metadata in Context

As Lagoze (2000) pointed out, the notion of "metadata" is not absolute but relative: it is only really meaningful in a context that makes clear what the data itself is. In the wider context of digital repositories, the use of metadata has been recognized as a crucial component for discovering and sharing

information resources in a networked information environment (Beall, 2005). For example, given a collection of bibliographic information, metadata might comprise information about each bibliographic item, such as who compiled it and when.

In addition, the significance of metadata for digital resources is acknowledged in a number of different applications. More specifically, in the educational context, a number of authors (Besser, 1999; Singh, 2000; Yordanova, 2007; Cohen et al., 2011) argue that metadata of learning resources is crucial. The use of metadata in an educational context is extensively researched (Sutton, 1999; GESTALT, 1999; Recker and Wiley, 2001; ETB, 2002; Dron et al., 2002; EQO, 2004). Educational metadata are deployed to describe learning objects used in various domains such as Ceramics (Patra, 2008), Architecture (Vuorikari et al., 2010) and Medicine (Nikolaidou et al., 2005).

On the other hand, focusing on research and science, we see that digital repositories introduced in the scientific domain provide publications and information collected, but also sets or collections of primary data (Lyon, 2006). These traditional library competencies will play an important role in successful e-Science-based research (Osswald, 2008). Some cases where digital repositories containing digital objects for science and research described with metadata are the following:

- Oceanographic data (Han, 2001).
- Statistical data and reports (Yamada, 2004)
- Geospatial data (Devillerset al., 2002; Wayne, 2004; IVOA, 2004; INSPIRE, 2005)
- Geographical and marine data sets (Beard, 1996; NDN, 2004)
- Medical resources and ontologies (Shon and Musen, 1999; MedCIRCLE Workshop, 2002; Supekaret al., 2004; Supekar, 2005)

Apart from educational and research/scientific applications, description of cultural heritage materials using metadata is also widely researched (Attig et al., 2004; Research Libraries Group, 1999; Visual Resources Association, 2003; Weibel and Miller, 1997; Yoon & Kim, 2000). Numerous digital archives programs in museums are conducted worldwide to preserve and sustain mankind's cultural heritage. Such programs aim to preserve cultural heritage and collections; popularize fine cultural landmarks; and encourage information/knowledge sharing. They are considered as the prerequisite and foundation for developing digital museums from which a museum's conventional functions can be extended and developed through information technologies (MacDonald, 2000; Alonzo, 2001; Chen, 2003; Shindoet al., 2003; Hemminger, 2004; Mei, 2004, Fan et al., 2008; Hsu et al., 2006).

Examining the different applications of metadata retrieved in literature, we identified three broader cases on which metadata can be applied, that is education, research/science and culture. To this direction, this thesis will examine metadata quality in the cases of three different collections of digital objects, namely an educational collection, a research/scientific collection and a cultural collection.

2.2.3.2 Issues for Metadata Creation

Who Provides Metadata

A significant issue that was identified in the literature review was the background and skills of individuals that provide metadata. In the origins of cataloguing practice, people that provided metadata were metadata or information experts that received specific training on the topic. Nevertheless, with the introduction of digital repositories that serve specific domains which demand specialized knowledge on a variety of issues, this paradigm has shifted. These days, professional metadata creators (metadata experts) and resource authors (or domain experts) represent two distinct groups of metadata creators. Metadata professionals, such as catalogers and indexers, are people who have had formal training and are proficient in the use of metadata standards. Although researchers have noted problems with inter-indexer consistency (Chan 1989), we see that professionals and experts of each domain, generally produce high quality metadata (Weinheimer, 2000; Greenberg et al., 2001).

Although in the early days of digital libraries, metadata creation was managed by trained professionals, due to the large number of objects that must be tagged with metadata, people who are not metadata experts are becoming involved in the metadata authoring process. These people can be apart from professional metadata creators or technical metadata creators, also domain experts, and/or community/subject enthusiasts (Zschocke & Beniest, 2011). This is based on the fact that creators have intimate knowledge of their creations, whereas indexers and catalogers can assist them with knowledge of metadata schemas and classification systems (Greenberg, 2002). In many cases, domain experts are the individuals responsible for the creation or metadata annotation of the digital resources hosted in digital repositories. Researchers regularly produce abstracts, keywords and other types of metadata for their scientific and scholarly publications. Visual artists, another class of authors, generally sign and date their works (Greenberg et al., 2001).

This notion is enhanced by Barker and Ryan (2005) that found that several metadata elements caused some difficulties to metadata experts given the catalogers were not the creators of the resources being tagged. In particular, the technical requirements and domain specific elements such as the educational description required specialist knowledge and care in maintaining

consistency between entries from different catalogers. As (Zhang & Dimitroff, 2005) points out, metadata development requires substantial cooperation and collaboration across different potential user groups. In addition, Trigget et al. (1999) and Crystal & Greenberg (2005) indicate that having metadata experts create metadata and domain experts use them may create a significant disparity between ideal metadata elements for a specific domain and actually recorded metadata (Lubas et al., 2004). As Parsons et al., (2011) point out, collaboration between data managers and data creators increases the ecosystem transactions and balance of communication and has broad, positive repercussions. At one level, this sort of collaboration can improve data and metadata completeness. In short, both theory and practice suggest that metadata is best generated through collaboration between the subject-knowledge holders (termed here as domain experts) and information specialists (termed here as metadata experts). This makes the data more useful within and without specialist communities. Drawing from this, our experiments that are described in the main part of this thesis include both metadata experts and domain experts in the processes set up to ensure metadata quality.

Metadata generation from domain experts is not the only case retrieved in literature. A significant strand of research is focusing on the social metadata, that is metadata in the form of tags and annotations that are provided from communities that use the actual digital content (Monge et al., 2008; Goh et al., 2009; Ding et al., 2009). Focusing on this strand of research is not within the scope of this thesis although it would be an interesting future direction. Already, efforts have been made to incorporate user-contributed metadata into library catalogues (Chua & Goh, 2010). For example, several libraries have started making use of metadata from LibraryThing (www.librarything.com) as well as other social features (including tagging, list-making, annotation, ratings and reviews) and having them available using interesting and dynamic interfaces (e.g. tag cloud) (LibraryThing, 2010; National Archives, 2010). Another term for social metadata, implying a more structured approach to metadata creation from non-experts, comes with the term “Content Curation Communities”. Content curation communities are distributed communities of volunteers who work together to curate data from disparate resources into coherent, validated, and oftentimes freely available repositories. Content curation communities helped to develop resources on drug discovery (Li, Cheng, Wang, & Bryant, 2010), worm habitats or bird migration (Sullivan et al., 2009), astronomic shifts (Raddick et al., 2007), and language (Hughes, 2005).

Overall, (Lampert & Chung, 2011) statement covers our appreciation of social metadata, “*though it is unlikely that users will ever contribute the necessary quality and quantity required for complete digital collections metadata, experiments in user tagging and folksonomies have shown that this is a very*

compelling area for exploration". For the purpose of this thesis, only the domain experts are included in the metadata annotation of digital objects to allow for a focused study on the requirements of this user group on metadata related processes and the overall metadata quality.

Another strand of researchers deals with metadata that are created automatically, by means of special software programs. The position of this thesis is that although this concept holds great promise, this process can lead to errors in the metadata and to failed searches because metadata creation generally needs human intervention to be successful. Similarly to the case of social metadata, we feel that automated metadata is an interesting prospect to consider but always in connection to human annotation as a means to control metadata quality that come out of such systems. As Beall (2006) points out, even the most sophisticated computer program might not be able to differentiate among locks (hydraulic engineering) or locks (hardware) or locks of hair, air locks, etc., or among authors with similar names. In addition, as Malaxa & Douglas (2005) points out, discovery and reuse of digital assets, particularly non-textual assets such as images and videos dictate annotation through human-created metadata. Previous studies show that automatic generation of metadata provides acceptable performance (Liddyet al., 2002; Han et al., 2003; Peng and McCallum, 2004; Takasu, 2003) but in spite of that, researchers generally conclude that the best results are achieved by integrating automated and manual methods (Schwartz, 2001).

An example of the combination of the aforementioned approaches to metadata annotation is presented in Stefaner et al., (2007) that presents the MACE repository that contains learning resources for architecture. More specifically, the MACE system relies on three ways to enrich learning resource metadata and to create relations among the resources. Each learning resource is classified through a number of architecture experts (domain expert metadata). Using automatic entity recognition, names of buildings and architects are identified in learning resource descriptions (automated metadata). Last but not least, user community tools are deployed (social metadata).

Metadata Guides

Although metadata experts have the skills and material needed to effectively annotate resources with metadata, the same thing does not apply to domain experts. Those individuals must also be provided with effective just-in-time support (Malaxa & Douglas, 2005; Chen et al., 2002). As Cassella (2010) points out, ongoing staff training is an essential prerequisite to building a successful repository and to assessing the repository's potential for growth and development. Zhang & Dimitroff (2005) argue that the effect of metadata cannot be maximized until people understand the domain clearly, interpret it

in the context of metadata correctly, and comprehend the behavior of metadata elements fully.

Researchers, serious collectors of information and even users of information should know what metadata guidelines to use for capture, management, storage and/or preservation of digital objects (Paradigm Project, 2005-2007). Programs that foster the development of digital repositories for archives, libraries and museums in the USA, provide training in digital imaging and metadata standards, access to scanning equipment, and software tools that streamline the creation of metadata records (Middleton, 2005). Adopting good practice at the outset of a document is advised in a number of studies as a way to increase the longevity of the digital content (Groenewald & Breytenbach, 2011; Külcü & Külcü, 2010; Lubas et al., 2004; Friesen, 2002). Drawing from these findings, the deployment of metadata guidelines in all three cases examined in this thesis is also undertaken to facilitate metadata annotators in their tasks.

2.2.3.3 Community Driven Metadata

There are numerous studies that have pointed out the need to take into account input coming from domain experts of the application domain when developing a metadata application profile (Chen et al., 2002). In this way, domain experts are not only becoming the authors of metadata as it was discussed before but also participate in shaping the metadata used within the deployed collections or digital objects. To this end, the task of using any metadata schema in a project or indexing task is a complex, resource-intensive undertaking. It requires elements to be chosen, interpreted, used, and then possibly reinterpreted by each group or individual collecting or developing resources. Varying implementations of this element set, moreover, threaten to create problems for the effective searching and exchange of metadata records between projects and jurisdictions (Friesen et al., 2002).

In most of the applications of metadata mentioned previously, literature describes the way in which metadata have to be customized to fit the purpose of a specific community of users/practitioners. Apart from realizing the need to use metadata to manage the massive scale of digital collections effectively, establishing a metadata model and application profile is a fundamental part of any digital repository project (Patra, 2008). In addition, involving the domain experts in this process is a key decision that can also ensure the usage of the digital repository developed. As Choo (2002) pointed out, knowledge creation is everyone's concern, and not the responsibility of a specialized few. Adding to this, Bainbridge et al. (2003) pointed out that digital library systems have the potential to empower users, not just librarians, to conceive, assemble, build and disseminate new information collections.

Already, there are cases where the domain experts are offering their input to the metadata application profile development. In the work carried out by Sreenivasulu (2000) initial meetings were held with the different departments of a university that would be served through a digital repository project. As it turned out, different disciplines had very different metadata requirements and it became very clear that the departments would have their own views of what and how they wanted to use the services provided. The authors reflected that part of the success of the repository has their ability to be responsive to these differences. An example of such "requirements" comes from Chu et al., (2010) where a digital repository for chemistry was developed. In this case, the following type of information (metadata) was deemed essential: Certain details of the crystal structure data space group, authors, institution, quality indicators, color of the crystal, keywords, what kind of compound it is, organic or inorganic, organometallic compound.

This part of the literature review, guided our decision to involve domain experts in the first stages of metadata application profile development, taking into account their requirements and structuring the application profile accordingly.

2.2.3.4 Metadata Application Profiles

Metadata schemas (or metadata models) are sets of metadata elements designed for a specific purpose, such as describing a particular type of resource (NISO, 2004). The definition or meaning of the elements themselves is known as the semantics of the schema. The values given to metadata elements are the content. Metadata schemas generally specify names of elements and their semantics. Optionally, they may specify content rules for how content must be formulated (for example, how to identify the title), representation rules for content (for example, capitalization rules), and allowable content rules (for example, terms must be used from a specified controlled vocabulary). There may be also syntax rules for how the elements and their content should be encoded.

In order to facilitate the reusability of metadata in different applications, as well as the interoperability between different systems that use metadata to describe the same type of resources, several metadata specifications have been developed. Specifications can be said to represent standards early in their development, prior to receiving approval from standards bodies, and they tend to be experimental, incomplete and more rapidly evolving (Farance, 1999).

When a specification is widely recognized and adopted by some standardization organization (such as ISO), it then becomes a metadata standard. Standards can be defined as documented agreements containing

technical specifications or other precise criteria to be used consistently as rules, guidelines, or definitions of characteristics, to ensure that materials, products, processes and services are fit for their purpose (Bryden, 2003). As Slaton & Abbate (2001) state, the adoption of standards may simplify some aspects of a system while creating a demand for more skilled labor elsewhere. This is true for metadata as well, since the deployment of standards has created a series of requirements for the usage of metadata in practice, as the creation of guiding documents and consequently the training of metadata annotators. This need was also identified from related literature as Lindner (2001) states that promotion of learning technology standards should take high priority and urgency, and the action should start with easy access to information on standards activities by repositories, explanatory documents, and translations (Cebeci & Erdogan, 2005).

But standards alone are not the answer in addressing the need for structured and well-defined metadata as there is no one all-encompassing metadata standard to be used in all applications. Rather, there are various metadata standards or specifications that can be adapted or "profiled" to meet community context-specific needs (Kraan, 2003). This conclusion has led to the emergence of the application profile concept. Roberts (2003) defines an application profile as a customization of a metadata standard or specification to meet the needs of a particular community of implementers with common needs and requirements. For Duval et al. (2002), the purpose of an application profile is to adapt or combine existing standards or specifications into a package that is tailored to the functional requirements of a particular application, while retaining interoperability with the original base schemas. An example is the work of Guinchard (2002), which provides input for the development of DC application profiles in libraries (Tambouris et al., 2007).

Although application profiling and contextualization of metadata seems the way to go, metadata experts are divided in two opposing movements related to metadata usage. On one side are the minimalists, who contend that metadata should be a very simple set of only a few elements so that it is equally useful across domains and resource types (Campbell, 2002; Lagoze, 2001). Proponents of this type of simple metadata argue that when metadata standards become more narrowly defined, the risk of lower element usage from users is higher. On the other side of this debate, those in favor of stricter standards and more complex element sets argue that in order for digital objects to be retrieved easier and used as they are supposed to, the metadata elements must be consistent (Chepesuik, 1999; Sokvitne, 2000; Tennant, 2004).

Standards and their application profiles have shown great promise and results during their deployment, but still some of the problems initially identified related to their application, have not been resolved. More specifically, as a

number of authors stated, a key problem with metadata is with interpretation of the words used (Hiddink, 2001; Rada, 1995a; Rada, 2001; Tate & Hoshek, 2009). Different developers interpret words differently and, therefore, when searching for the object it may not come under the same word. Through this argument, the need for guidelines and corresponding training for metadata annotators becomes even clearer.

2.2.3.5 Types of Metadata

Researchers on the field of metadata have identified different classes (or types) of metadata in order to provide a more concrete definition for this concept. Lagoze, Lynch, and Daniel (1996) provide one of the more extensive analyses in this area by identifying seven types of metadata:

1. Identification/description,
2. Terms and conditions,
3. Administrative data,
4. Content ratings,
5. Provenance,
6. Linkage/relationship data,
7. Structural data

Gilliland-Swetland (2000) offers a taxonomy of five types of metadata:

1. Administrative metadata,
2. Descriptive metadata,
3. Preservation metadata,
4. Technical metadata,
5. Use metadata

In her discussion of metadata for cultural resources, and Besser (1999) notes six types of image metadata which are comparable to Gilliland-Swetland's groupings:

1. Descriptive metadata,
2. Discovery and retrieval metadata,
3. Structural metadata,
4. Administrative metadata,
5. Intellectual metadata,
6. Other metadata

The examples presented here and other metadata classifications reveal a number of similarities in labels used. The distinction among different types of metadata is, however, not absolute (Making of America, 1998), because a single metadata element can support functions that fall into more than one class. For example, location metadata, such as a URL or a shelf location number for a physical image, is a "multifunctional element" because it

facilitates object discovery (retrieval), object use (networked or physical access), and assists with administrative activities (record-keeping), (Greenberg, 2001).

The categorization of metadata elements in different classes is provided here to allow for a better understanding of the functions that each different metadata element fulfills within a metadata schema, metadata standard or application profile. Developing a concrete categorization of metadata elements is not within the scope of this thesis.

2.2.3.6 Beyond Cataloguing

Building upon the functionalities of metadata discussed in the previous paragraph, we focus on the importance of metadata in terms of the preservation of the digital objects they describe. As Day (1998) pointed out, there is a growing awareness that metadata has an important role in digital resource management, including preservation. To that end the effort or cost of preservation metadata today can be considered negligible compared with cost associated with a catastrophic loss of digital material in the future that might have been mitigated had preservation metadata been available. According to Lee (2001) cataloguing is one of the most crucial aspects of any digital repository project. Without cataloguing and metadata, digitizers will have lost an invaluable opportunity to record technical information that could prove essential in the future. Enhancing this notion, (Anido et al., 2003; Cebeci & Erdogan, 2005; Anido et al., 2002) argue that the lack of information about the properties, location or availability of a resource could make it unusable. This situation is even more crucial in an open, unstructured environment like the Internet. To achieve long term preservation of digital objects, complex metadata is needed to ensure that the provenance of each set of data is fully understood. Adding to this notion, Namuth et al., (2005) point out that in order for repositories to be sustainable, the individual learning objects of which they are comprised should carry certain characteristics which facilitate longevity. Overall, it becomes clear that there is a need for continuous curation of metadata throughout its lifecycle (Brophy, 2008). To this end, this thesis will analyze the quality assurance mechanisms that can be put in place to ensure that metadata are curated and enhanced throughout the lifecycle of a digital repository.

2.2.3.7 Metadata Cost

Gathering from the discussion carried out so far, we can say with certainty that metadata are an intricate part of the digital object discussion. Within this discussion, the involvement of domain experts in the metadata definition and creation process is advised. The effort vested in the definition of metadata application profiles but more importantly in the creation of metadata records, always comes with a cost in time and therefore in money. It is evident that the

more metadata an object has the easier it will be to locate it through searching. On the other hand, the more metadata mean bigger effort from the human annotators and therefore higher costs. The creation of metadata is a task requiring major labor and financial cost and, most important, the involvement of knowledgeable and experienced people (Liddy et al., 2002; Barton, Currier, & Hey, 2003).

A number of studies demonstrate that entering complex metadata efficiently, accurately, and consistently can be confusing, costly, time consuming and error prone (South and Monson, 2000; Greer, 2002; Kunze, Brase, & Nejdl, 2002; Marshall, 1998). Despite that, we have showcased that only through complex and domain specific metadata will the digital repositories deployed offer added-value services to their users, fulfilling their *Raison d'être*, which is to offer access and discoverability to digital objects that are highly specialized and unique.

Pointing out the cost and time needed to annotate resources, Crystal and Land (2003), point out that it would take about 60 employee-years to create metadata for one million documents (Kovacevic et al., 2011). According to a study by Reerink (2003), the cost of metadata creation accounts for 30% of the total cost of a digitization process, indicating its significance in our discussion as well. One of the main challenges for human-created metadata noted in the literature is the potentially high cost of production in terms of human effort, time, money, and the errors and inconsistency that occur (Geisler, McArthur, & McClelland, 2002; Marshall, 1998). In this sense the overhead involved in creating and managing metadata is a potential barrier to the successful use of metadata to facilitate reuse and sharing (Malaxa & Douglas, 2005).

Looking at relevant literature we were able to locate only a handful of a systematic documentation of the effort vested in creating metadata for digital objects. Some examples were found in the field of digital libraries but their generalization was not easy mainly because of the specific nature of the case studies examined in this thesis. Despite that, discussion on these is essential to begin to understand the costs involved in metadata annotation of digital objects.

In Willer et al., (2008), the costs associated for populating the National and University Library of Croatia with 385 digital objects was discussed. The main findings of this paper are summarized below:

- Cataloguing of the digital objects, which accounted for adding metadata to them, took up 4.334 minutes (72 h) which was 33% of the entire time that was vested in all four processes (Identification, Selection, Cataloguing & Archiving),

- The cataloguing process for one digital object demanded 55 minutes per record,
- The cataloguing process for one non-digital object, demanded 45 minutes per record,
- Updating an existing metadata record, demanded 21 minutes per record

Phillips (2005) carried out a similar study covering 937 objects that were archived in the National Library of Australia. In his study, the average time needed to catalog one object was 81 minutes amounting to 18% of the total time needed for the corresponding processes that Willer et al., (2008) also discussed. This aspect of our literature review offered really useful insights in the costs associated with creating metadata for a set of digital objects. To this end, we will also incorporate the cost discussion in the experiment carried out within this thesis and we will discuss alternatives to lowering metadata costs, like automated metadata mechanisms and social metadata or content curation communities.

2.2.4 Quality of Metadata

We have already defined the concept of a digital object and the concept of digital repositories that host them. We also defined metadata and discussed a series of aspects related to their development in various contexts. This paragraph will focus the discussion on the quality of metadata as an absolute measure of how well digital repositories perform in one of their fundamental functions, which is the discoverability of digital objects.

Before defining quality of metadata, it is essential to clarify the title of this paragraph. As Subirats et al., (2008) pointed out, two quality-related topics can be identified in the field of metadata: "quality of metadata" and "quality in metadata." The first topic concerns finding ways to evaluate and ensure the quality of the metadata itself which is the objective of this thesis as well (Duval et al., 2002; Currier et al., 2004; Hillman et al., 2004; Sicilia et al., 2005; Robertson, 2005). The second topic refers to the usage of metadata to represent information, such as the quality assurance process for a digital resource or results from its evaluation which is a topic heavily researched that will be discussed briefly but overall it is considered to be outside the aim of this thesis.

To narrow on the definition of quality of metadata, it is needed to start defining quality from a more abstract level, focusing on the specific domain that is examined in this thesis. According to Merriam-Webster's Online Dictionary (2009), the meaning of quality is defined as "*a degree of excellence*" (Babalhavaeji et al., 2010). Donaldson (1994) states that "*quality does not lend itself to easy or precise definition, but there is general*

agreement that the quality of any activity should be assessed in relation to its purpose" (Babalhavaeji et al., 2010). This statement is aligned with Juran's (1989) standpoint that quality is defined as "*fitness for purpose*". Other definitions like "*meeting or exceeding customer expectations*" or "*satisfying the needs and preferences of its users*" (McClave & Benson, 1992; Evans & Lindsay, 2005) put more emphasis on user needs as drivers of quality. Pawlowski (2007) offers a definition of quality that involves the internal processes that have to take place within an organization, arguing that "*quality is appropriately meeting the stakeholders' objectives and needs which is the result of a transparent, participatory negotiation process within an organization*".

The definition of quality at a generic level that will be considered in this thesis is the ISO 9000 definition for quality as "*The totality of features and characteristics of a product, process or service that bears on its ability to satisfy stated or implicit needs*" (ISO, 2005).

To pin point the exact definition of quality of metadata, we will look into a series of definitions that were retrieved from literature, discussing data quality, information quality and metadata quality as we have found that these terms are in some cases used interchangeably to talk about an overarching concept which is the quality of data or information provided by individuals in a given situation. Table 2.5 provides the main definitions that were collected.

Authors	Definitions of Data, Information & Metadata Quality
Marschak, 1971	How accurately information represents a particular event
Stronget al., 1997; Wang and Strong, 1996	Information or data quality can be defined as data that are fit for use by end-users.
Wand and Wang, 1996	Quality of mapping between a real world state and an information system state
Taylor, 1998	The value or worth the information has in relation to the purposes at hand
Eppler, 2003	The degree to which the information at hand either meets the requirements of the particular activity in which the user is engaged or the degree to which the information meets the expectations of the user
ISO 11620, 2006	Totality of features and characteristics of a product or service that bear on the library's ability to satisfy stated or implied needs.
Pawlowski, 2007	Appropriately meeting the stakeholders' objectives and needs which is the result of a transparent, participatory negotiation process within an organization'

Authors	Definitions of Data, Information & Metadata Quality
Hilligoss and Rieh, 2008	Subjective judgment of goodness and usefulness of information

Table 2.5: Various definitions of metadata quality

We see that data or metadata quality is a notion that is not adequately defined within the existing literature. Overall, it seems that the notion of quality in general, is closely connected to the satisfaction of users or consumers of data through a set of services offered. It seems that in order to get a clearer and more precise definition of metadata quality, we have to focus more, taking into account the domains of application. Nevertheless, for a generic definition of quality of data, we will build upon the definition provided by ISO (2005) to argue that metadata quality is:

The totality of features and characteristics of metadata that bears on its ability to satisfy stated or implicit needs.

The need to look quality in context is confirmed by the statements from Vlasceanu *et al.* (2007) that argued that quality means different things to different people but also from Cullen & Chawner (2010) that argued that quality has no single accepted definition because it has multiple perspectives and has been defined differently under different conditions. As stated in a number of studies, the usefulness and quality of metadata, is contextual and dynamic (Taylor, 1986; Jörgensen, 1995b; Strong et al., 1997; Greenberg, 2001b; Stvilia & Gasser, 2008; Stvilia et al., 2007). With changes in domain culture, activity systems, knowledge and technology, and user expectations, the quality of these systems can quickly become outdated and require regular intensive maintenance and upkeep. Keeping this in mind, we consider that generating a definition of metadata quality would not be of any use as opposed to looking at cases of quality in different fields that can guide the development of metadata quality assurance processes.

2.2.4.1Digital Objects and Metadata Quality

As mentioned before, assessing quality of digital objects is a difficult and complex task that often revolve around multiple and different aspects that must be observed in each context. For instance, in the context of digital libraries, Custard and Sumner (2005) claim that concerns about quality are mainly related to issues of:

1) Accuracy of content,
2) Appropriateness to intended audience,
3) Effective design, and
4) Completeness of metadata documentation

In the specific field of learning multimedia resources, the so far most recognized instrument for quantitatively measuring quality, the Learning Object Review Instrument (LORI) (Nesbit et al., 2003), approaches quality from nine different dimensions:

1) Content quality,
2) Learning goal alignment,
3) Feedback and adaptation
4) Motivation,
5) Presentation design,
6) Interaction usability,
7) Accessibility,
8) Reusability, and
9) Standards compliance.

We notice that in the case of the most widely used quality instrument (LORI) in the field of digital resources for education, there is no direct reference to metadata. The closest reference to metadata is the ninth dimension that is the "Standards compliance" which we feel is not adequate to capture the essence of metadata quality.

Metadata are not considered a focal point of the aforementioned approaches for automatically measuring quality as it also happened more than once in the relevant literature (Krauss & Ally, 2005; Buzzetto-More & Pinhey, 2006; Maceviciute & Wilson, 2008; Sanz-Rodríguez et al., 2010). Nevertheless we see that specific dimensions that are measured and characterize quality ("Appropriateness to intended audience", "Presentation Design", "Accessibility" and "Reusability") depend on the existence and availability of metadata attached to the objects, or on measures of popularity about the objects that are obtained only when the resource is publicly available after a certain period of time (Cechinel et al., 2011).

In addition, most quality approaches (targeting either digital objects or metadata or both) face problems of scalability since that, as the number of objects or records increases, it becomes impossible to provide evaluative metadata for every single object. Such situation leaves many objects of the repositories without any measure of quality at all. A recent study (Cechinel & Sánchez-Alonso, 2011) has shown that in MERLOT, from the total amount of digital objects available, approximately 12% were rated by users or peer reviewers, and only 3% presented at least one peer-review and one user rating at the same time. As pointed out in the case of costs associated with digital object cataloguing, costs for digital object or metadata quality are not scalable as well. To this direction, efforts have to be made to automate the process in the extent possible so that the biggest possible number of digital objects and metadata records are quality certified.

2.2.4.2 Why do we need quality of metadata?

S.R. Ranganathan is a well-known systematic thinker that made immense contributions to librarianship with his "five laws of library science" (Ranganathan, 1931). The fourth law, "*save the time of the reader*," gets at the core of library operations and points out the need for mediating tools like indexes, taxonomies, classifications, library catalogs, and metadata. The importance of metadata quality cannot be overstated. According to Guy, Powell, and Day (2004) there is an increasing realization that the metadata creation process is key to the establishment of a successful digital repository. Robertson (2005) enhanced the previous argument by stating that the need to support the development of quality metadata is perhaps one of the most important roles for professionals working in the field of digital libraries.

Thousands of digital resources are published online every day, and their quality control, assurance and evaluation are of paramount importance for potential users. Furthermore, since the properties of digital resources are reflected in their metadata, quality should also be of primary importance in the agenda of metadata research (Glover et al., 2001). As stated by Subirats et al., (2008), content description through quality metadata creation and use of standard terminologies is the basis of efficient content management as well as the development of value-added services on top of digital repositories.

A series of arguments that were retrieved through the literature review to support the need for quality of metadata are the following:

- Poor quality metadata can mean that a resource is essentially invisible within a repository and remains unused (Barton et al., 2003),

- Since the properties of digital resources are reflected in their metadata, quality should also be of primary importance in the agenda of metadata research (Subirats et al., 2008),

- Poor data quality can be harmful to an information system, having an adverse effect on decision making processes of end-users who depend on these systems and well-managed quality metadata (Shankaranarayanan et al., 2003),

- There is a concern that the abundance of information presented will lead to time wasted in search of information and difficulty in the absence of appropriate tools for evaluating the quality and reliability of the information and its management (Taibi et al., 2005),

- High quality metadata ensure accurate and complete access to digital resources and enable end-users to easily find and retrieve the resources they require (Shankaranarayanan & Even, 2006).

- The level of description in learning objects annotation is often deficient: most metadata elements are either never or rarely used by annotators (Sánchez-Alonso, 2009),

- For consumers of data it is critical that they can make informed judgments about the quality of data for a particular application in a specific context (Zschocke & Beniest, 2011),

- Both the benefit and the acceptance of information systems depend heavily on the quality of data provided by these systems (Ballou & Tayi, 1999; Wang & Wang, 2009)

Adding to the last argument, a series of studies have indicated the need for control and quality assurance of metadata that describe various digital resources, when sharing knowledge (Stephens, 2004; Sturdy, 2001), conducting business on-line (Manouselis and Costopoulou, 2006), making available scientific research outputs (Schweik et al., 2005; Thomson et al., 2003), offering learning resources in education (Currier et al., 2004) or using the semantic web in general (Greenberg et al., 2003; Sicilia, 2006). The scope of the different applications that are affected by low quality of metadata covers also the cases that we examine within this thesis (education, research and culture).

Taking into account all of the above, digital repositories that maintain institutional or disciplinary digital objects, have to decide whether they will apply quality control to metadata. Weighting their decisions, they will have to balance the benefits of having authors create metadata and the difficulties of requiring or encouraging authors to adhere to standard practices or metadata creation tools that impose restraints on data input (Ranganathan, 1931).

2.2.4.3 Quality of Digital Repositories

The field of research that focuses on the assessment of digital repositories presents a number of metrics used to assess digital repositories performance and success. Westell (2006) suggests eight factors of success selected for the evaluation of Canadian digital repositories. Six measures are internal factors: mandate, integration with planning, funding model, measurement, promotion, and preservation strategy, while two measures are external: relationship with digitization centers and interoperability. Thibodeau (2007) proposes a more general framework for the evaluation of digital repositories including five dimensions: service, orientation, coverage, collaboration and state. Within each dimension Thibodeau poses questions to help managers of digital repositories to assess success. Swan (2007) recommends a quality framework based on four domains: content recruitment, user awareness and involvement, workflow practices, and financial discipline. Xia and Sun (Xia & Sun, 2007) suggest following

measures: number of deposits, availability of full text, cost per deposit, usage assessment. Xia and Sun also stress the importance for digital repositories' evaluation of factors such as authors' attitude, information on depositor, usage assessment and interoperability.

These studies are only a sample of the ones identified that focus on pinpointing quality for digital repositories. Nevertheless, despite the quantity of studies, quite few of them take into account metadata quality as a decisive factor on the overall quality of the digital repositories. Usually, metadata quality is examined in isolation, not within the context of a generic quality framework as the following studies have indicated. In the context of this thesis we will present data that reinforce our proposal that metadata quality is an integral part of any digital repository project and a critical success factor.

2.2.4.4 Dimensions, Metrics and Indicators of Quality

In a time of financial crisis the need to assess digital repositories' success by adopting valid, appropriate and relevant performance indicators has become stringent (Cassella, 2010). As Armstrong (1995) pointed out, in an industry that is essentially a service, and whose primary currency is intangible information, "*quality is not only difficult to define; it is difficult to quantify*".

To quantify quality we need to discuss about indicators, dimensions and metrics of quality. The first step is to clearly define those to avoid confusion. A data *Quality Dimension* is an aspect or feature of information and a way to classify information and data quality needs (McGilvray, 2008). A *Quality Metric* describes what is being measured and how it will be measured during the quality assurance processes deployed (Heldman & Mangano, 2009). *Quality Indicators* are statistical measures that give an indication of output quality. However, some quality indicators can also give an indication of process quality (ESS Quality Glossary, 2010). This shows that quality indicators are the quantitative way to express quality metrics. Overall, we consider the term "*Quality Dimension*" as a series of aspects related to data quality that we want to examine and "*Quality Metrics*" as the specific measures that express each quality dimension. Finally, "*Quality Indicators"* are defined as the statistical measures and thresholds that express the degree or level of quality metrics.

Having defined quality and quality in metadata, the need to look at exact dimensions, metrics and indicators to measure it, is more than clear. In respect to the definition of the quality dimensions we will start the discussion by defining a set of quality dimensions that will guide our selection of metrics and after this, indicators. To this end, we adopt Lee et al.'s (2002) categorization of derived dimensions of information quality, adapting it to fit the metadata context. In their study (based on earlier work by Wang & Strong,

1996), four high-level dimensions of information quality were defined, providing comprehensive coverage of the multi-dimensionality of quality:

- **Intrinsic Metadata Quality**: Represents dimensions that recognize that metadata may have innate correctness regardless of the context in which it is being used. For example, metadata for a digital object may be more or less 'accurate' or 'unbiased' in its own right,
- **Contextual Metadata Quality**: Recognizes that perceived quality may vary according to the particular task at hand, and that quality must be relevant, timely, complete, and appropriate in terms of amount, so as to add value to the purpose for which the information will be used,
- **Representational Metadata Quality**: Addresses the degree to which the metadata being assessed is easy to understand and is presented in a clear manner that is concise and consistent,
- **Accessibility Metadata Quality**: References the ease with which the metadata is obtained, including the availability of the metadata and timeliness of its receipt.

For our study, we will focus mainly on the Intrinsic, Contextual and Representational dimensions of metadata quality. These dimensions will also guide the selection of the metadata metrics. To choose the metadata metrics that is needed for the Metadata Quality Assurance Certification Process, we studied the relevant literature to see which specific metrics are introduced or adopted by other authors. Looking at Table 2.6 we see clearly that there is no consensus among researchers on a concrete set of metadata quality metrics but rather some commonly used metrics.

Metric	Gonçalves et al., 2007	Kim & Kim, 2008	Margaritopoulos et al., 2008	Moen et al., 1997; Greenberg et al, 2001	Tozer, 1999	Rothenberg, 1996	Alkhattabi et al., 2010	Wang & Strong, 1996	Bruce & Hillman, 2004
Accessibility							Yes		Yes
Accuracy	Yes	Yes		Yes	Yes			Yes	Yes
Appropriateness						Yes	Yes		
Completeness	Yes	Yes	Yes	Yes	Yes		Yes	Yes	Yes
Conformance	Yes								Yes
Consistency		Yes		Yes	Yes		Yes		Yes
Correctness			Yes			Yes			
Currency									
Intelligibility					Yes				
Objectiveness							Yes	Yes	
Presentation								Yes	
Provenance									Yes
Relevancy			Yes				Yes		
Timeliness					Yes		Yes		Yes

Table 2.6: Metadata Quality Metrics proposed in literature (sorted alphabetically)

Accuracy, Completeness and Consistency are the prevalent metadata quality metrics that have to be chosen in order to be compatible with the majority of the related literature. To decide on the remaining metrics to be used, we revisited the adapted model for quality assurance dimensions and we decided to include at least two metrics per dimension of quality, adding to the initial three. Overall, the quality dimensions that will be examined in this thesis are the following ones:

- **Intrinsic Metadata Quality**: Accuracy, Correctness,
- **Contextual Metadata Quality**: Completeness, Appropriateness,
- **Representational Metadata Quality**: Consistency, Objectiveness,

The specific definitions that are provided in the literature for these metrics are not relevant in this part of the discussion. In the full description of the Metadata Quality Assurance Certification Process, a comprehensive definition is provided for each one of the metrics.

Having selected the metadata quality metrics, we turn to literature to search for a quality assurance metrics for the metrics themselves. Looking at Heinrich & Klier (2011) we adopted and adapted a set of requirements that each metadata quality metric has to follow. These requirements guided the development and definition of the selected quality metrics as well as the supportive documentation needed for human annotators to be able to utilize them in peer-reviewing metadata records. More specifically, the requirements as collected from related literature are the following:

1) **Normalization**: An adequate normalization is necessary to ensure that the values of the metrics are comparable (to compare different levels of metadata quality over time),
2) **Interval scale**: To support both the monitoring of an improved metadata quality level over time and the economic evaluation of quality measures, the metric has to be interval scaled. This means that the difference of 0.2 between the values 0.7 and 0.9 and the values 0.4 and 0.6 of the metric represents the same extent of improvement of data quality,
3) **Interpretability**: The values of the metric have to be comprehensible and easy to interpret by users. For instance, considering a metric for completeness, it could be interpretable as the percentage of attribute values which are stored (i.e. they semantically differ from NULL) in the digital repository at the instant of assessment,
4) **Adaptivity**: To assess data quality in a goal-oriented way, the metric has to be adaptable to the context of a particular application (for

example, if a metric is not relevant in the context of digital repository, the metric has to be adapted in order to assign a weight of 0 to this attribute),

5) **Feasibility**: To ensure applicability, the metric shall be based on input parameters that are determinable. When defining the metric, methods to determine its input parameters have to be defined. From an economic point of view, it is also required that the assessment of metadata quality can be conducted at a high level of automation.

Having defined the specific quality dimensions and metrics to cover each one of them, the next step is to define the indicators that are used to measure them. Completeness is acknowledged by a significant number of studies as a fundamental quality characteristic that is also the easiest to be quantified and measure automatically without human intervention (Margaritopoulos et al., 2012). Indeed, it can be considered a prerequisite to assess completeness, since incomplete records are in any case not of quality due to lack of essential information (Sicilia et al., 2005). Completeness is indicated by the number of complete metadata elements compared to the total number of available elements in a metadata instance. It's expressed using a percentage (i.e. if 5 of the 10 elements of an instance are completed, the completeness of this record is 50%).

Apart from this absolute definition of completeness, Ochoa & Duval (2009) argue on a more abstract notion of completeness say that it represents the degree to which the metadata instance contains all the information needed to have a comprehensive representation of the described resource. This definition means that even if one or two not important elements are missing from a metadata instance, still the record can be considered as completed. For this reason, in the methods deployed in the Metadata Quality Assurance Certification Process, we also collect evaluations of completeness through peer-reviews of domain experts that rate completeness from a scale from 1 to 5, meaning the perceived completeness of an instance, in contrast to the absolute one (calculated automatically). In addition, to allow for a simpler and more straight forward process for the reviewers of metadata records, we also measure the remaining quality metrics through a scale from 1 to 5, 1 being the lowest, drawing from similar studies in the existing literature.

2.2.4.5 Metadata Completeness Problems

In the introduction of this thesis, we carried out a short discussion on studies that identified metadata quality problems. In the following paragraph, a set of similar studies that were retrieved during the literature review is presented. In general, it is obvious that problems with the cataloguing of resources are not new. Reports on poor functionality and difficult use of catalogues can be traced continuously almost 30 years (Markey, 1984; Borgman, 1996a;

University of California Libraries, 2005). As Borgman (1996b) and Tennatt (2005) point out, through time, the process and tools of cataloguing changed, but most of the modifications were made on the surface and not in the core functionality that would truly affect the user experience.

Sánchez-Alonso (2009) points out some of the main shortcomings that have been identified in metadata records pointing out that the level of description in digital object annotation is often deficient: most metadata elements are either never or rarely used by annotators. In addition, Pages et al., (2003) argues that metadata records in existing repositories are often fragmentary and unstructured (Pages et al., 2003). Quantitative surveys carried out in the field of digital repositories for education (Najjar and Duval, 2006; Ochoa and Duval, 2009b; Sicilia et al., 2005) on the actual use of LOM and the completeness of metadata records show that only a small amount of the available elements is actually used. Relevant surveys by Guinchard (2002), Najjar, Ternier, and Duval (2003) and Friesen (2004) have shown that indexers tend to fill out only particular metadata elements that could be considered "popular," while they ignore other less popular elements. The issue of incomplete metadata records is also problematic in collections resulting from harvesting from metadata databases (Dushay & Hillmann, 2003) or from automatically generated metadata (Greenberg, Spurgin, & Crystal, 2005; Ochoa, Cardinaels, Meire, & Duval, 2005; Margaritopoulos M., Margaritopoulos, T., Kotini, & Manitsaris, 2008). The research community considers completeness of metadata a fundamental quality characteristic (Margaritopoulos et al., 2008).

Efron (2007) carried out a study on the use of metadata in institutional repositories that participate in the Open Archives Initiative (OAI). He sampled 86.522 records from 19 repositories and identified problems in the completeness of almost half of the elements of Dublin Core that were used (Description, Subject, Publisher, Contributor, Rights, Coverage and Source). The aforementioned elements presented an average completeness of 35.3%. Overall, the average completeness of the set of 15 DC elements was 65.7%.

Mark (2006) harvested 5.445 records from 9 institutional repositories in Canada that also used Dublin Core as a metadata standard. He found out that the element with the highest completeness was Date with a mere 14.6% followed by Subject (13.7%) and Format (13.4%). Overall, the average completeness of the set of 15 DC elements was as low as 6.6%.

Shreeves et al., (2003) harvested 613.813 metadata records coming from 23 institutions, including museums, academic libraries and public libraries that followed the Dublin Core standard. They found out that 5 out of 15 elements of DC were completed in less than 50% of the records (Source – 11%, Contributor – 20%, Format – 32%, Relation – 39% and Language – 41%). Overall, the

average completeness of the set of 15 DC elements was 54.5%. Furthermore, Shreeves et al., (2003) analyzed completeness grouping the content providers per domain (museums and cultural societies, academic libraries and digital libraries) to see how metadata usage is affected by the context. The average completeness for museum and cultural societies was 69.1%, whereas for academic libraries was 39.3% and for digital libraries was 53.3%.

Metadata may be viewed as a quality control device in that it helps to filter non-relevant documents and produce retrieval results that have more precision than the results generally obtained via most of today's Web engines and indexes (Roszkowski and Lukas 1998). Although these research studies provide rich empirical account of metadata quality problem incidents, the proposed models are context specific. They lack an integrated approach and may not be readily generalizable and operationalizable (Stvilia et al., 2007).

2.2.4.6 Quality Approaches

The previous paragraph focused on really specific problems with metadata completeness. In general, metadata quality has been studied in many diverse contexts and communities for varying functions besides information retrieval. Such applications included digital preservation (Rothenberg, 1996), data used in simulations and models (Rothenberg & Rand, 1997), databases (Medawar, 1995), and museums (Marty and Twidale, 2000).

In this section, we are looking at existing research on metadata quality focusing on aspects other than completeness. A series of studies that mostly focus on the correctness and appropriateness of metadata have been identified showing the need to examine quality using a series of indicators (Barton, Currier, & Hey, 2003; Howarth, 2003; Moen, Stewart, & McClure, 1998; Park, 2006; Wilson, 2007). As Wilson, (2007) points out, little research has been done in addressing how well a given schema is constructed to meet contextual requirements of users. This lack of research will be tackled by incorporating users in the metadata application profile design.

Another set of studies focused on frameworks that have been proposed and used in metadata quality assessment. Moen and McClure (1997), and Moen, Stewart, and McClure (1997, 1998) conducted a series of studies to evaluate the utility of U.S. Government Information Locator Service (GILS), a metadata schema for improving public web access to government resources via the Internet. They proposed an evaluation framework with five dimensions: policy, users, technology, contents, and standards and rules. Based on a literature review, they identified a group of assessment criteria and they employed qualitative and quantitative techniques to assess selected metadata records. They introduced a pool of 50 measures via an internal examination on sample records under the metadata schema. These measures

were grouped under four criteria categories: completeness, profile, accuracy, and serviceability.

In a similar study, Dushay and Hillmann (2003) identified several criteria and used them to assess the performance of automatic metadata harvesting. Despite that, their batch judgments on sufficiency, accuracy, and completeness of metadata harvested were done internally by the developers without a consideration of its target users' preferences and needs. Park (2006) assessed the interoperable Canadian architecture collections, by conducting a survey of metadata fields from nine architecture databases and analysed the structure of these databases. The survey revealed some issues of metadata description, such as different fields with the same semantic meanings and different meanings for the same field. In this study, the same shortcoming is found. Although the findings are meaningful to developers, they ignore users' perception on metadata quality.

Involving users in the quality assessment, Liddy and colleagues (Liddy et al., 2002; Liddy, Allen, & Finneran, 2003) conducted metadata evaluations from harvest to use with users' participation. They compared manually and automatically generated metadata in terms of retrieval effectiveness and users' satisfaction. The findings indicated that only minimal differences existed between these two types of metadata. Their studies also showed that involving users in metadata evaluation is effective. Howarth (2003) assessed Namespace that comprises 17 element labels, including contact information, rights/restrictions on use, edition, roles, summary and description, identifiers, etc. A focus group with 19 participants was conducted to evaluate these element labels in terms of their clarity and potential usefulness. The study demonstrated some problems in terms of understandability. It also suggests the necessity of involving real users in metadata evaluation.

In Zhang and Li (2008) two studies were carried out on a metadata application profile to (a) examine the usefulness of each metadata field with respect to satisfying different user groups' needs, and (b) discover additional metadata fields that may be useful. With similar objectives a third study was conducted after a pool of 1,000 metadata records had been created. This last study was aimed at examining users' perceived usefulness of the metadata fields and their values when the users interact with metadata records for given topics and scenarios. The three evaluation studies were conducted during different stages of the metadata implementation which showed the need for continuous assessment of metadata quality. In the study, the usefulness of metadata was assessed through an online survey that targeted a mixed audience from thirty-three domain experts, metadata experts and the public. The participants were asked to indicate on a 5-point Likert scale from not useful (1) to very useful (5) the perceived usefulness for each metadata element. The findings of the survey provided an insight about the perceived

usefulness of the elements that allowed the better design of the system and metadata tools.

All the aforementioned studies yielded really interesting results for several aspects of metadata quality. The two major shortcomings that were identified were (a) the application of the studies' outcomes on the examined digital repositories as a feedback mechanism were limited and (b) none of the studies proposed a comprehensive approach to metadata quality assurance.

2.2.4.7 Comprehensive Quality Assurance Studies

In order to improve and assure the metadata quality in digital collections, libraries and repositories, a number of quality approaches (QAs) have been proposed in the literature. This shows the importance of QA approaches in all contexts. These QAs bring forward either comprehensive quality assurance methods or specific tools and techniques to improve some particular metadata quality dimension. In this section we review this work, also reflecting on the application context of any practical implementations.

In related work, Stvilia et al., (2004) presented a framework of metadata quality dimensions and discussed specific metadata quality metrics that can be applied to measure them. The study continued with a set of recommendations for ensuring high metadata quality in federated collection and it concluded with a case study of one collection. In this case, the proposed quality metrics were applied to assess the quality of metadata records within a collection of 150.000 metadata records but no concrete method of improving the metadata quality was suggested. Park, (2007) looked at architectural collections of Canadian universities, examining metadata quality of existing records in order to support their aggregation for building an interoperable metadata and search interface for Canadian architecture collections. Park found that there had been noticeable issues in metadata description, especially the categorizing of groups of tables and fields. This study showed the poor quality of metadata of heterogeneous collections and also indicated the need for a stage of metadata design to address contextual characteristics of the architectural domain.

Adopting ideas and concepts from service science, Vinagre et al. (2011) presented a Library Service Quality Model specifically designed to evaluate digital libraries. Questionnaires were filled from stakeholders as a part of this study that pointed out the need of efficient information retrieval and thus high quality metadata. The authors argued in favor of the need for continuous application of the Library Service Quality Model to regularly monitor digital libraries and implement continuum improvements. Waaijers and van der Graaf (2011) confirmed the previous argument by investigating the concept of quality in the various phases in the life cycle of research data. The 2.811 interviews that were collected, indicated that metadata play a vital role in the

retrievability of research data. The findings and contribution of this study is strengthened as they were tested via a national academic survey of three disciplinary domains as designated by the European Science Foundation Metadata.

Finally, in the educational domain, Zschocke and Beniest (2011) analyzed different quality metrics for metadata and proposed a quality assurance framework that can be applied on the metadata creation process in the case of an agricultural learning repository. The framework proposed although not tested using actual data is really similar to MQACP in the sense that it also uses peer review methods to assess the metadata coming from content providers as well as automated methods for the assessment of metadata records. In addition to the studies above, a number of studies that are quite relevant to our study were identified each one of them justifying the need of specific methods included the MQACP. More specifically, Sutton (1999) supported the need of metadata standard application profiling to fit the domain. Poll (2008) showcased the need for tangible metrics for the assessment of digital libraries that must be incorporated to their lifecycle. Lefoe et al. (2009) emphasized the need of peer-review methods in any quality assurance process that is established for a digital library, it pointed out the need for supportive documentation but also incentives for the peer-reviewers. Babalhavaeji et al. (2010) found a lack of authoritative guidelines or criteria to help library professionals define how the quality of an academic library can be properly measured and improved to serve end-users better. Complementing this, Kastens et al., (2005), presenting the work carried out in the Digital Library for Earth System Education (DLESE) project (http://www.dlese.org/) argued in favor of providing cataloging best practices & guidelines for the metadata annotators. He also supported the existence of a QA process to check the quality and consistency of the metadata provided by them. Finally, the improvement of the skills of the metadata annotators through the establishment of QA processes as a long-lasting effect with impact on future projects was pointed out by Saarti et al. (2010).

In addition, creating quality assurance mechanisms is not a process that is independent of the context in which they are applied. Fuhr et al. (2001) created a generalized schema for a digital library that takes into account the application domain, the technologies used, the users and the type of data a digital library contains, when discussing its evaluation. Overall, the same data may have different levels of quality in different contexts of use. Additionally, different users of data may value the same quality attributes of data in different ways (Strong et al. 1997, Kelly et al., 2005). Borgman et al. (1996) argued that digital libraries are constructed, collected and organized, by and for a community of users. Therefore, the functional capabilities of these libraries support the information needs and uses of that community. Poll & te Boekhorst (2007) argues that libraries use the quality indicators that best fit

their needs and that local circumstances may also affect the methods of data collection. From the above it is made obvious that using a quality assurance process in a new context does not simply mean taking it and applying it *ceteris paribus*. It takes effort and adaptations to each application domain so that it fits the users, the data and the technologies used.

2.2.4.8 Considerations for Quality

Completing this chapter, we feel that it is important to also examine a part of the literature that offers practical advice on setting up quality assurance mechanisms for metadata but also assessing their success. This input also served as the basis to design and implement the Metadata Quality Assurance Certification Process. Westbrook et al., (2012) pointed out that future projects that will carry out quality audits have to consider whether using human, automated, or combined evaluation is most efficient for determining the quality of the metadata. Decisions on the number of staff members and other individuals that are necessary for a good quality evaluation are important. Since the quantity of metadata records being assessed is usually large, deciding whether or not sampling of the metadata records is sufficient will also come into the picture. An assessment of how the communication of metadata creation guidelines can impact the quality of the product has to be carried out. Finally, an important consideration is the one of the overall cost of the metadata quality assurance process deployed. As pointed out by various studies (Ehlers et al., 2006; Even & Shankaranarayanan, 2007; Otto et al., 2009; Even & Kaiser, 2009), researchers have taken an economic viewpoint for data quality management and developed methodologies to assess data quality accordingly.

Apart from these considerations when designing a metadata quality assurance process, it is crucial to assess the process once it is deployed. To this direction, Clements and Pawlowski (2012) suggest that for each quality assurance process it is important to analyse and understand three aspects:

1. **The effect of the quality approach**: as quality approaches aim at different objectives and scopes, it has to be clear which effects can be achieved with a quality approach. These effects have to match the users' needs and requirements,

2. **The perception of the stakeholders**: one important aspect is how users perceive quality. Even though lots of efforts might be spent for quality assurance, the value and the awareness of the users about quality is a main concern. It has to be clear which stakeholders benefit,

3. **The cost of applying a quality approach is of crucial importance**: Most of the repositories are not commercial; thus, there

is no budget for quality assurance tasks. This means that solutions need to be simple and cost-effective.

Attempting to elaborate on the three aspects discussed by Clements and Pawlowski (2012), we examined the Quality Maturity Model (QMM) proposed by Wilson & Town (2006) that measures the quality of a library service on a five-step scale. Each point on the scale of QMM has a general description of the level of quality in the organization that is accomplished through with a list of specific attributes. Although the model was created for library services, we feel that adapting it in the case of metadata, would be a useful guide for metadata quality assurance processes to assess their maturity and guide their continuous development. The adapted model is termed as the Metadata Quality Maturity Model (MQMM) and the five levels, as they were adapted, are the following:

Level 1 (Initial): The metadata quality process is ad hoc, and occasionally even chaotic. Few methods are defined for ensuring metadata quality and their success depends on individual effort. Overall:

- Quality is achieved in an ad hoc way,
- Customer satisfaction is reactive and unpredictable,
- Quality depends on the capabilities of individuals, and varies with their innate skills, knowledge and motivations, and
- Training for metadata authoring is ad hoc and unstructured and depends on the motivation of each metadata or domain expert.

Level 2 (Repeatable): Basic metadata quality methods are established. A basic process is in place to repeat earlier quality levels:

- Quality methods, and tools to implement these methods, are established,
- There are effective methods to allow the digital repository to repeat earlier success in user satisfaction,
- Such methods are practiced, documented, enforced, trained, measured, and able to improve, and
- Training for metadata annotation is provided in a structured way and it is reactive to specific events such as low metadata quality for specific elements.

Level 3 (Defined): The metadata quality process is documented and standardized.

- There is a defined, documented metadata related strategy, from which all methods are derived; there is an understanding of the activities, roles, and responsibilities of each content provider of the repository, and how they fit into the overall strategy, and

- Training for metadata quality is a cycle of training needs assessment and programme provision.

Level 4 (Managed): Detailed measures of the application of the quality process are collected. The quality process is quantitatively understood and controlled:

- Quality measures/metrics are part of every work that is carried out on the metadata,
- These measurements form the basis for evaluating products and services of a digital repository,
- Changes are implemented to improve the quality of services and products,
- The organization sets quantitative goals for quality and user satisfaction, and
- Training for metadata annotation is a cycle of training needs assessment, programme provision, and measurement of the effectiveness of the programme.

Level 5 (Optimizing): Continuous quality improvement is enabled by quantitative feedback and from piloting innovative ideas:

- The entire digital repository is focused on continuous improvement in every service, product and the metadata quality assurance process itself,
- All staff are encouraged to continuously improve themselves and their work;
- The digital repository is able to identify weaknesses, and the means to strengthen the process, proactively with the goal of preventing problems;
- Innovations that exploit the best practices are identified and transferred throughout the content providers contributing metadata and
- Training for quality is focused on preparing staff for future product and service requirements.

We feel that the adapted Quality Maturity Model (QMM) introduced by Wilson & Town (2006) will help us assess the proposed Metadata Quality Assurance Certification Process but also identify the next steps to reach level 5 (Optimizing) of the whole process. This scale can also serve as a threshold for future research on metadata quality assurance processes proposed.

2.3 Literature Review Analysis

2.3.1 Introduction

This chapter provides an overview of the research literature in the area of metadata, digital resources, digital repositories/libraries and metadata quality. More specifically, it engages a conceptual framework for the categorization and assessment of research that is relevant to these terms that has been published in thirty-two (32) related journals. The results are expected to demonstrate the nature and the status of metadata research. In this review, we examined only the papers that included in their title, abstract, keywords or full text one or more of the following keywords: "resource", "repository", "library" and "quality" combined with the term "metadata".

Table B.1 (Appendix B: Literature Review Tables & Figures) shows all the journals that were searched for relevant publications, along with the ones that were retrieved and were relevant to our field of study, either we had access to them, or not. The total number of journals identified was 39 whereas access was possible for 32 of them. The table also presents the time period covered by our study. In most cases, we collected relevant papers published in these journals from late 1990s to 2012. Overall, apart from some minor problems in accessibility we feel that we considered a rather extensive list of as it included 39 of the most appropriate journals for metadata, metadata quality, and digital libraries and repositories research.

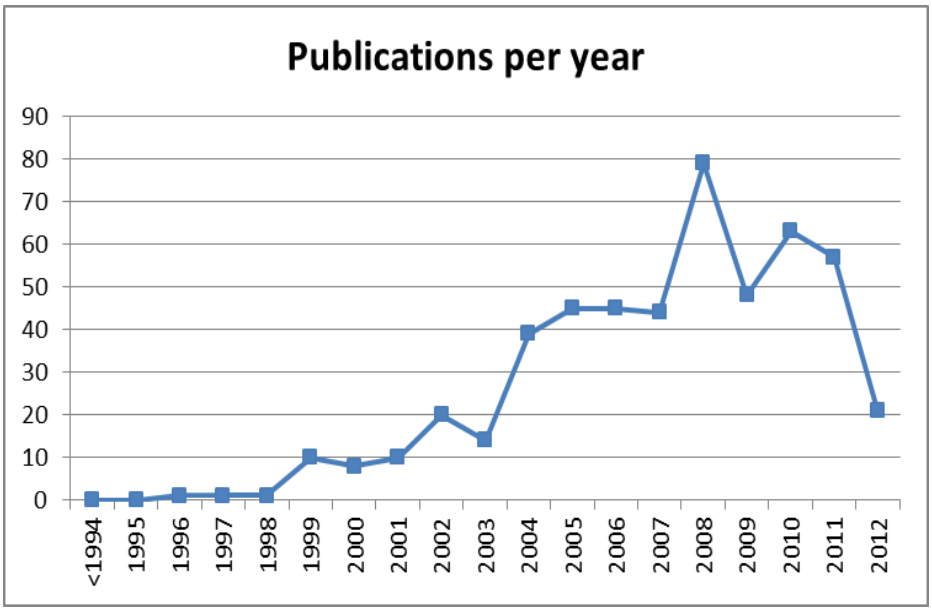


Figure 2.2: Related papers published per year in the past eighteen years (1994-2012)

Figure 2.2 presents the number of published papers throughout the period from 1994 to 2012 in the examined journals. We note that starting from 1998 the number of papers published around the topics of this review is increasing

with a low rate until the small drop in 2003. From 2004 until 2007 there is a significant increase in relevant publications that reaches 45 publications per year. 2008 was the year with the most publications retrieved whereas the number of publications although a bit lower, remained around 50 publications or more from 2009 to 2011. The significant drop in 2012 is partly attributed to the fact that the literature review took place in the first half of the year, thus not including anything that was published from March onwards. The field that was examined is pretty broad, ranging from literature on the use of digital objects either for learning or scientific purposes, to repositories and metadata applications. This fact does not allow us to attribute the high number of publications to a solo reason but it shows that in general, the field of e-learning, focusing on metadata, learning resources and repositories attracts a great deal of research effort.

50% of the journals produced the 84.19% of the publications retrieved, having from 11 to 64 papers identified in the context of this study (Figure B.2, Appendix B). “The Electronic Library” was the journal with the most papers retrieved (64) whereas “Electronic Library & Information Systems” was the second one (51). “Performance Measurement & Metrics”, “Online Information Review”, “D-Lib Magazine” and “Interdisciplinary Journal of E-Learning and Learning Objects” followed, with 35, 34, 33 and 33 papers respectively.

Overall, seventeen out of the thirty-two journals contributed no more than 10 papers each. The main reasons behind that were that (a) some of the journals were focused on more technical issues that were beyond our scope and (b) some of them were not accessible to us in their full content so some of the papers identified were not analyzed in the end. To analyze the contribution of each journal to our analysis, Table B.2 (Appendix B) presents the ranking of journals based on the rate per year of each one, for the years that our research identified relevant papers being published to them. “Active” period in this case is considered the period between the years of the oldest relevant publication until the year of the newest relevant publication retrieved. For example, in the case of the “D-Lib Magazine”, the oldest relevant publication for us was dated in 1996 whereas the newest was dated in 2011, yielding an “active” period of 16 years. This analysis will really highlight the journals which display significant activity in our fields of interest.

In Table B.2, we present the high ranking journals are mainly the ones that had many papers identified in Figure B.1 (Appendix B) as well. Despite that, in the case of “IEEE Transactions on Learning Technologies” that is a fairly recent journal; its rate is among the top ones, indicating the high relevance of the journal to the areas of interest of this study. Another case is the one of “Code4lib Journal” for which eight really relevant publications were retrieved for the past five years, raising its rate to 1.6, meaning that only recently this journal has started dealing with the issues addressed in our analysis. Other

journals such as “Journal of Knowledge Management” and “Journal of Library Innovation” are ranked high just because the really small “active period” that is only 1 year and therefore even with 2 or 3 papers their rate is deceivingly high.

It has to be noted that this “active period” publication rate cannot be considered an absolute measure of comparison. For example, the number of published issues per year is different among journals. Nevertheless, we use this rate as an indicator of the volume of relevant research that each journal publishes, considering only the point in time after which, each journal started publishing papers on our topics of interest. In total, 605 papers were identified during a first search in a total of 39 journals. Out of the 39 journals, 32 of them contained relevant papers to which we had access and therefore could be classified. Therefore, from the total of 605 papers identified, 506 (83.6%) were finally analyzed on most of the aspects of the metadata schema and are represented in the figures and tables of this section.

2.3.2 Distribution per research type

Figure B.2 and Table B.3 (see Appendix B) present the distribution of the classified papers, according to their research type. The results demonstrate that most types of research are being carried out in the areas examined. Additionally, the field doesn’t seem to be that mature yet. Only 13% of the published papers test a proposed theory and 23% of the papers apply a theory for the needs of a specific implementation (a total of 36%). A reservation exists due to the fact that papers characterized as Applied (Implementation) are classified in this category because they are not testing some particular theory. In some cases, such papers simply present a prototype implementation and cannot be considered a complete contribution. A big part of the papers is examining phenomena to result to observations for the purpose of theory building which indicates that the field is still under development. It will be interesting to examine in the future how research type has changed over time, in order to assess if theoretical and descriptive papers were published in the early years, as well as if the practical applications and theory testing papers are more recent. The full set of the results discussed above is presented in the respective table in “*Appendix A: Distributions per Journal*”.

2.3.3 Distribution per research method

Figure B.3 and Table B.4 (see Appendix B) present the distribution of the examined metadata literature, according to the research method used. The results demonstrate that the methods applied more often are Case Story (in almost 21% of the papers) and Experiment (Almost 18% of the papers). In addition, almost 14% of the papers contain some form of data collection which shows a tendency of the research field to look at qualitative and quantitative evidence. A significant number of cases identified related to surveys (12%)

carried out, which show the need for a qualitative overview of the metadata domain. Another 12% of the cases represent complete products that are presented by their developer which were mainly software and infrastructures related to metadata storage or annotation.

Related to the literature review studies identified (10%) we see that the field of metadata is covered in an adequate percentage related to studies that analyze the research carried out in the field. Overall, the results also indicate that about one quarter of metadata research is not based on data (Argument and Case Story papers together reach a percentage of 25.4%). This might be an indication that data from existing metadata applications have to be collected and researched. Finally, the score of 12% for Product Description, combined with the 18% of Experiments (real or simulated) might be an indication of the practical focus of metadata research. The full set of the results discussed above is presented in the respective table in "*Appendix A: Distributions per Journal*".

2.3.4 Distribution per research claim

Figure B.4 and Table B.5 (see Appendix B) present the distribution of the examined papers, according to their claims. These results provide a useful insight: 17% of the examined papers claim generality of their results. Additionally, another 50% claims validity for the specific case described. That is, 67% of the literature published in the examined journals does not claim either generality or at least validity of its results for the specific case examined. The full set of the results discussed above is presented in the respective table in "*Appendix A: Distributions per Journal*".

2.3.5 Distribution per focus unit

Figure B.5 and Table B.6 (see Appendix B) present the distribution of the e-market papers, according to their focus unit. The results show that literature mostly focuses on the whole sector (38.9%) and secondly on a specific method that is proposed (22.4%). An important percentage looks at the research on the organization level (18.94%) whereas a significant number of studies look at the concepts discussed on the individual or group level, focusing on the needs of specific communities (0.41% and 15.27% respectively). Finally, the small percentage of papers focusing on the society (4.07%) might be an indication that research has not yet studied the effects of metadata, digital resources and repositories on a societal level. These could be an indication that further research is necessary, both on an individual basis as well as on a society basis (the societal perspective). On the other hand, the fact that this review is carried out on mostly metadata-related journals (that is, with a rather technical focus) could be explaining this observation. The full set of the results discussed above is presented in the respective table in "*Appendix A: Distributions per Journal*".

2.3.6 Distribution per target audience

Figure B.6 and Table B.7 (see Appendix B) present the distribution of the retrieved papers, according to their target audience. The results demonstrate that in the examined set of journals, the papers are split between the ones focusing on researchers (44%) and the ones focusing on practitioners/implementers (50%). A small percentage of papers were identified as aiming managers. This may be an indication that metadata, metadata quality and repositories research is not yet connected with the market, and this issue requires further investigation.

Apart from these initial findings, it is true that characterizing a paper based on the target audience is difficult. This means that in many of the examined cases the focus was not 100% clear and a decision had to be made to characterize it each paper in one of the categories. To see if the previous results can be trusted, we can also look at the length of the papers examined as one indicator of its focus. As Groenlund (2004) states, shorter length of a paper may indicate a practitioner or manager focus. We found that the arithmetic mean of pages is 14.1 per paper whereas the median was 13 pages per paper. In addition to that, we examined (using a more intuitive measure) the classification of papers in "Short" (less than 5 pages), "Medium-Short" (6 to 10 pages), "Medium" (11 to 15 pages), "Medium-Long" (16 to 20 pages) and "Long" (more than 20 pages). It was found that 2% of the papers were classified as short, 19.2% as medium-short, 35.4% as medium, 16% as medium-long and 13.8% as long. Most papers were medium or long whereas in 69 cases we could not classify them, because they were found in html format (Table B.8 and Figure B.7 in Appendix B).

Another possible indicator of the target audience is the number of references. We would assume that a low number of references is an indicator of a practitioner or a manager and not a researcher (Groenlund, 2004). It was found that the average references per paper were 26.4 with a median of 23. This fact shows that literature in the examined fields is targeted towards researchers and less towards managers or practitioners which may also show a need for literature with such focus.

Looking at the references contained in the retrieved papers we attempt another intutitive classification based on the number of them. In the following table we created ranges of references (from no reference to 10, from 11 to 20, etc.) and we calculated the number of papers within these intervals. As it is evident, most of the papers contain up to 30 references (69.7%) whereas almost half of the papers (48.6%) contain from 11 to 30 references which shows that the papers are mainly theoretical and targeted towards researchers. An interesting observation comes from the fact that 6.9% of the papers (35 of them) contain more than 60 references, that usually point to literature reviews (Table B.9).

2.3.7 Distribution per origin (institution type)

Figure B.8 and Table B.10 (see Appendix B) present the distribution of the retrieved papers according to the affiliation of their primary author. Most of the published papers retrieved come from universities (78%) whereas libraries (9.88%) and research institutes (8.5%) are also important sources of literature. Companies and Research Institutes combined, account for more than 9% of the published literature which shows that the field examined has started moving towards the market. The full set of the results discussed above is presented in the respective table in "*Appendix A: Distributions per Journal*".

2.3.8 Distribution per discipline

Figure B.9 and Table B.11 (see Appendix B) present the distribution of the literature according to the discipline where their primary author belongs to. The results indicate that research carried out in the examined fields is published from people with Information Science (34%), Librarianship (24.9%) and Computer Science (23.1%) backgrounds. This classification cannot be considered as absolute as based on the definitions provided some of these backgrounds may overlap. Despite that fact, it seems that most of the research carried out on metadata comes from people with information science background and less technical. Lastly, a significant amount of research was published from authors with a background on Education that mainly dealt with learning objects and metadata. The full set of the results discussed above is presented in the respective table in "*Appendix A: Distributions per Journal*".

2.3.9 Distribution per geographical region

Table B.12 (see Appendix B) presents the geographical origin the retrieved research. As it is obvious, the major contributors of literature in this area are the United States of America, Great Britain and Canada. It is quite encouraging that all the literature retrieved so far has come from authors from 53 different countries, showing the interest on these topics, globally.

Grouping the results on a continent level (Table B.13, Appendix B), we see that America (38.5%) and Europe (34.4%) are the pioneers of research in the field of metadata and related applications. Asia follows with 17.6% whereas research on these topics is not yet widespread in Oceania (6.7%) and Africa (2.8%).

2.4 Conclusions

The review analyzed papers that were published during the past fifteen years (1997-2012) in thirty-two (32) well-known scientific journals that publish research related to metadata, digital repositories and libraries. To store and

classify the information retrieved from the papers, two frameworks were used: a framework classifying e-market papers around concepts and issues discussed, and a framework classifying them according to their research characteristics. In addition, a basic metadata schema was developed to facilitate the data collection and analysis process. In total, 605 papers have been identified. From them, 506 were classified using the proposed frameworks. A number of interesting observations were made, producing some implications for current and future research. Overall, the field of metadata is a balanced and mature research field. There is a balanced choice of research types engaging Descriptive studies (14%) and Theoretical studies (16%) along with studies proposing new theories (31%) or testing existing ones (13%). Several research methods are engaged in similar degrees and a significant number of research papers claim either validity or generality of their results (67%) which is judged as really high. More studies that focus on society as a whole are needed, whereas the number of studies looking at methods and organizations is evenly distributed. The research carried out is both targeted towards researchers and practitioners/implementers and research on metadata is dominated by universities whereas libraries and research institutes also play an important role. Information Scientists, Librarians and Computer Scientists are the major contributors to metadata research carried out whereas United States, Great Britain and Canada are the driving countries of research in metadata with America and Europe also being the continents that concentrated the most publications.

2.5 Literature Review Outcomes for Thesis

In this paragraph, the main literature review outcomes that guided the development of the Metadata Quality Assurance Certification Process are presented in the form of short statements. These statements came out of the examination of the 506 papers retrieved. The ordering of the statements does not represent their importance but it rather follows the structure of this chapter:

- A digital repository/library is the ecosystem of processes, metadata, people, services and tools that facilitate the storage, discovery, retrieval and preservation of digital objects for a given community of users,
- The involvement of domain experts and creators of digital resources in metadata annotation is crucial. A mixed group of domain experts and metadata experts has to be involved from metadata design to metadata deployment in each digital repository project,
- Documentation that supports the metadata annotation process is essential to allow metadata providers to gain a comprehensive understanding of the particularities and requirements of each domain,

- Domain experts have to be included in the metadata design process so that specific requirements of the domain are embedded in metadata application profiles developed,

- The cost of producing metadata for digital objects is significant for any digital repository project. The research agenda on any metadata quality assurance process has to take this cost into account,

- Metadata quality approaches need to cover the entire lifecycle of a digital repository, addressing all the different phases of development with respective quality assurance methods,

- A major issue related to existing metadata quality frameworks proposed is that they lack generalizability and adaptability to different contexts than the ones they were originally designed for,

- Metadata quality assurance processes need to be assessed in terms of their effectiveness, through frameworks that will also allow them to be continuously developed and adapted to the ever changing parameters of digital repositories.

Some of the previous arguments were also confirmed by the statistical analysis of the 506 papers. More specifically, if we look at the statistics of the literature review in relation to the qualitative outcomes, we see that:

- Many papers claim anecdotal value and show lessons learned rather than presenting a compelling and structured argument. This fact indicates the need for a comprehensive quality assurance process that covers the entire digital repository lifecycle,

- There is a general lack of research that can be applied to cases other the one that they examine, showing the need metadata quality assurance processes that are transferable across domains. This notion is enhanced by the fact that there is small percentage of papers focusing on the society (4.07%), which shows that most efforts so far are domain-specific,

- Finally, a small percentage of papers were identified as aiming managers. This is an indication that metadata and metadata quality research is not yet connected with the market and the specific costs associated. This calls for further research on the application costs and viability of metadata quality assurance processes.

3. Metadata Quality Assurance Certification Process Description

The aim of this chapter is to present the Metadata Quality Assurance Certification Process in detail. The main phases of the proposed process will be presented, along with the tools that are used to apply it as well as the people that are involved. The Metadata Quality Assurance Certification Process (MQACP) is a process that attempts to improve the quality of the metadata produced by the curators of a collection since its early development. To do this, it involves metadata and domain experts that using the MQACP tools, try to improve metadata quality. More specifically, a metadata application profile, based on an existing metadata standard, is used along with a metadata authoring tool through which domain experts use to provide metadata for digital resources. Additionally, tools for the collection of data during the experiments are deployed, such as questionnaires related to the metadata elements that are being used as well as peer-review forms containing specific metadata quality metrics that aim to assess metadata quality.

The quality metrics deployed throughout MQACP are mainly completeness of metadata elements but also metrics such as the ones proposed by Bruce & Hillman (2004) including appropriateness, correctness, objectiveness, accuracy and consistency of metadata records. The main results that the process yields include a metadata application profile produced by domain experts, quality reviews of the metadata records of a given repository using the metrics described above as well as guidelines that support the metadata providers with good and bad examples of metadata records. Through these outcomes the metadata records are revisited and revised, trying to acquire higher quality metadata for the repositories involved. The application of MQACP follows specific phases through which every repository with digital content goes through. More specifically, MQACP is applied in relation to the repository maturity in each stage of development, starting from the testing phase of each repository through its regular operation when all systems and services are finalized and operating smoothly.

An overview of the proposed Metadata Quality Assessment Certification Process (MQACP) is presented Figure 3.1. Some of the main concepts that are used within the process are explained below:

- **Phase**: A period in the lifecycle of a Learning Object Repository. Each phase is recognized and characterized by specific milestones that are reached. In the case of MQACP, phases are separated using two criteria. First of all, the number of resources that are populated in the repository

and secondly the versions of tools that are available for the content providers. In the description of each phase, these milestones are explained thoroughly,

- **Step**: Specific actions that are introduced by the proposed MQACP. These actions are closely related with each different phase of the repository lifecycle, meaning that specific actions take place during specific phases. All the steps are related to the use of metadata in the envisaged repositories and clarify the metadata related actions that take place during each phase,

- **Quality Assurance Method**: Methods that are deployed in each step to introduce the quality assurance approach suggested. These methods are also linked to the specific phases but they can also be used in different phases if needed. For example, the Metadata Understanding Session is usually held in the Metadata Design phase, but in cases where the content providers face problems with the comprehension of metadata, it can be deployed in a latter stage as well,

- **Quality Tool/Instrument**: Tools and instruments that each method uses to measure its impact to the overall metadata quality. The Quality Tools are designed to support each method and generate tangible outcomes for the analysis of the metadata quality. An in depth analysis of the metrics and criteria used within these tools will follow,

- **Outcome**: Tangible products of each step of the MQACP. Although not shown in Figure 3.1, each phase of the MQACP includes specific outcomes that serve as input to the next phases of the process. More specifically, either in the form of countitative results or in the form of recommendations towards the content providers, upon the completion of each phase of the MQACP, the repository manager receives tangible input related to metadata quality,

- **Quality Actor**: Individuals that are involved in each step, either supporting the application of methods or being subjected to them (i.e. metadata experts facilitate the metadata peer review exercises but content providers are actually providing their input during the exercises).

The proposed process can be adopted by any initiative/project/organization that is planning the launch of a new LOR, from the very beginning of its development. Parts of the process may also be adopted in the case of existing repositories to improve metadata quality, taking into consideration that the results may be less significant than the ones presented here.

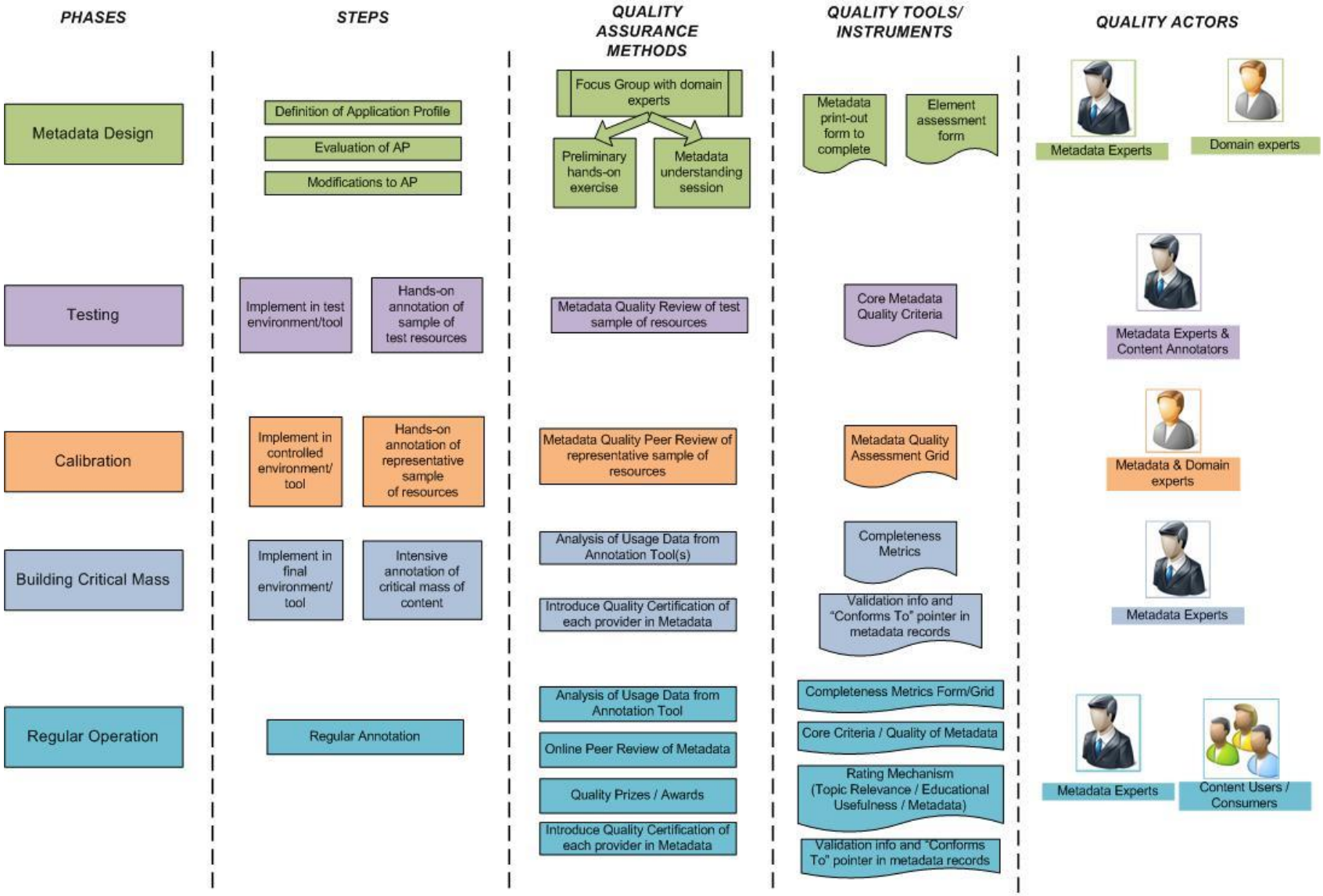


Figure 3.1: Overview of Metadata Quality Assessment Certification Process (MQACP)

In Figure 3.1, the tools that can be deployed in any of the phases B, C and D are not linked to them with one to one relationships. Each of the tools can be deployed in anyone of the phases. The only parameter that has to be considered each time so that the results yielded are relevant and accurate, is the number of resources on which they are applied. For instance, the metadata quality review exercise is usually applied on 10-20 resources in the "testing phase", on 5% of the resources in the "calibration phase" and on 30% of the resources during the Critical Mass Phase.

3.1. Instruments/Tools

To apply the MQACP, a variety of instruments is used to collect and analyze the data. In the case of the focus group meetings (Metadata Understanding Sessions) related to metadata design a base metadata schema is used as a starting point for the development of the application-oriented metadata profile. During the same sessions, questionnaires are used for assessing the understanding of each metadata element in the base standard as well as their relevance to the application domain (Figure 3.2).

Element	Is this element easy to understand?					Is it useful for describing Organic.Edunet content resources?					Should it be mandatory / recommended / optional???
1. General											
1.1 Identifier											
1.1.1 Catalog	1	2	3	4	5	1	2	3	4	5	Mand. - Rec. - Optional
1.1.2 Entry	1	2	3	4	5	1	2	3	4	5	Mand. - Rec. - Optional
1.2 Title	1	2	3	4	5	1	2	3	4	5	Mand. - Rec. - Optional
1.3 Language	1	2	3	4	5	1	2	3	4	5	Mand. - Rec. - Optional
1.4 Description	1	2	3	4	5	1	2	3	4	5	Mand. - Rec. - Optional
1.5 Keyword	1	2	3	4	5	1	2	3	4	5	Mand. - Rec. - Optional
1.6 Coverage	1	2	3	4	5	1	2	3	4	5	Mand. - Rec. - Optional
1.7 Structure	1	2	3	4	5	1	2	3	4	5	Mand. - Rec. - Optional
1.8 Aggregation Level	1	2	3	4	5	1	2	3	4	5	Mand. - Rec. - Optional
2. Life Cycle											
2.1 Version	1	2	3	4	5	1	2	3	4	5	Mand. - Rec. - Optional
2.2 Status	1	2	3	4	5	1	2	3	4	5	Mand. - Rec. - Optional
2.3 Contribute											
2.3.1 Role	1	2	3	4	5	1	2	3	4	5	Mand. - Rec. - Optional
2.3.2 Entity	1	2	3	4	5	1	2	3	4	5	Mand. - Rec. - Optional
2.3.3 Date	1	2	3	4	5	1	2	3	4	5	Mand. - Rec. - Optional

Figure 3.2: Screenshot of the form used for application profile design

The first two questions (respective columns) are answered using a Likert scale from 1 to 5 (with "1" being the lowest) and they concerned each metadata element of the application profile, asking if:

- The element is easy to understand
- The element is useful for describing digital objects in the target repository

The third question has to do with deciding if each element should be considered as "mandatory", "recommended" or "optional". This meant that each "mandatory" element would have to be filled out for every digital object in the targeted repositories, whereas "recommended" ones where not

obligatory but desired, etc. The domain experts indicate their preference, which is then crosschecked with the existing status of the elements in the application profile, to examine if revisions to the application profile are needed.

In the case of peer reviews conducted online, to support expert-driven peer review of metadata, a Metadata Quality Assessment Grid is being used with the metrics being evaluated from a scale from 1 (low) to 5 (high). The metrics used in this quality grid are adapted from Bruce & Hillman (2004) and are described briefly below:

- **Completeness**: Number of metadata elements filled out by the annotator in comparison to the total number of elements in the AP,
- **Accuracy**: In an accurate metadata record, the data contained in the fields, correspond to the resource that is being described,
- **Consistency**: Consistency measures the degree to which the metadata values provided are compliant to what is defined by the metadata AP,
- **Objectiveness**: Degree in which the metadata values provided, describe the resource in an unbiased way, without undermining or promoting the resource,

METADATA RECORD EVALUATION FORM			Evaluator	Nikos Palavitsinis			
Metadata record identifier:	*1*	*Metadata record URI in Portal:*	http://portal.organicedunet.eu/index.php?option=com_metavi&lold=lo-6ed4af09-cce6-11de-8531-				
			low → high				
1. In which degree is this metadata record completed?			1	2	3	4	5
Number of element values provided by the annotator in comparison to the total number of applicable element values. All mandatory and recommended elements must be completed. Extra points are gained for optional elements provided. Points are subtracted if recommended elements are missing							
2. Please identify the overall accuracy of the metadata values provided			1	2	3	4	5
In an accurate metadata record, the data contained in the fields, correspond to the object that is being described. Can you get the same information for the resource when looking at the resource itself and/or the metadata values? This question involves the task of checking the resource itself							
3. Are the metadata values provided consistent with the metadata standard used?			1	2	3	4	5
Consistency measures the degree to which the metadata values provided are compliant to what is defined by the metadata standard used in the specific application. Do they follow the definition of the element and the expected values?							
4. Do the metadata values describe the resource in an objective, unbiased way?			1	2	3	4	5
Degree in which the metadata values provided, describe the resource in an unbiased way, without undermining or promoting the resource in any way							
5. Are the metadata values provided, appropriate for the targeted use in the Organic.Edunet Portal?			1	2	3	4	5
Are the metadata values appropriate for helping users to find resources in the Organic.Edunet Portal? Multiple ontology terms, as well as most (if not all) of the educational elements, are important criteria here							
6. Please define the degree of correctness of the language used			1	2	3	4	5
Is the language used in the metadata, syntactically and grammatically correct?							
7. Please provide an overall score for the metadata of this resource, based on your ratings in questions 1 to 6. The overall quality of the metadata record:			1	2	3	4	5
8. Do you consider the quality of the metadata record for this resource of a desired level so as to be published in the Organic.Edunet Web Portal?			YES	NO			
				X			
9. Comments *Explanation of the review provided.Especially if the metadata record is rejected. Suggestions for improvement.*			**The metadata record should be revised in terms of the correctness of the language used. More elements such as... should be provided. Special attention should be given to the actual content of the resource as it is not 100% reflected in the metadata. These revisions are necessary prior to making the resource available through the Web Portal**				

Figure 3.3: Metadata peer-review grid

- **Appropriateness**: Degree to which the metadata values provided are facilitating the deployment of search mechanisms on top of the repositories,

- **Correctness**: The degree to which the language used in the metadata is syntactically and grammatically correct,
- **Overall Rating**: The overall score of the metadata record taking into account all the criteria above (not the average).

In Figure 3.3, the "Metadata Record Evaluation Form" is presented, containing the metrics described above. In addition to these metrics, three more questions are defined that help acquire a complete view of the metadata quality. In the first one, the reviewer is asked to provide an overall score for the metadata record and in the second one, to answer if the record is considered fitting for publication on the repositories in hand (Yes/No question). In the third one, the reviewer is provided with some space to provide any additional comments on the record itself, pointing out specific problems in the metadata.

For the Testing Phase of each repository, a simpler version of this form is used from metadata experts to carry out an initial assessment of the test metadata that are provided by domain experts. This form contains the following statements with a scale from 1 to 5, to allow the metadata expert to quickly review the resources and provide some basic recommendations to the content providers:

- The mandatory metadata fields are completed
- The recommended metadata fields are completed
- The optional metadata fields are completed
- The language used within the metadata appropriate and correct
- The resource is objectively represented by its metadata

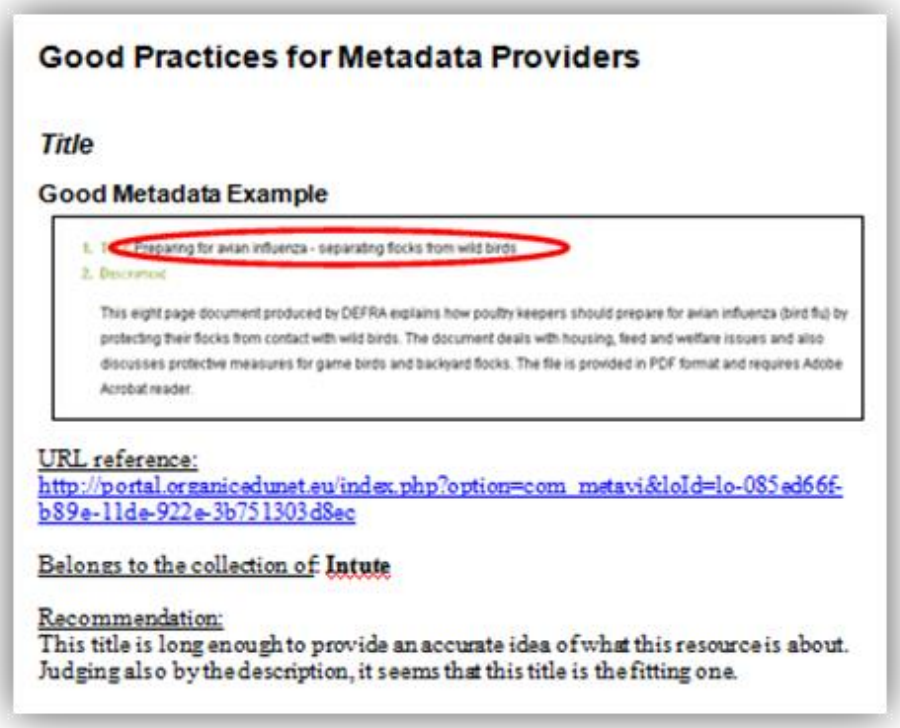
Collection Name	Record Name	Scientific Name	Classification	Common Names
JME, Material for Educa	Arms of the ancient bird	1	0	0
JME, Material for Educa	Illustration of an airplane buoyancy	0	0	0
JME, Material for Educa	Platform edge with securing line	0	0	0
JME, Material for Educa	Bones of a terrestrial crocodile	1	0	0
JME, Material for Educa	Picture of the ancient bird	1	0	0
JME, Material for Educa	View in the special exhibition Drache	0	0	0
JME, Material for Educa	Cambrian of Pathway On the traces	0	0	0
JME, Material for Educa	Cambrium of Pathway On the traces	0	0	0
JME, Material for Educa	Carboniferous of Pathway On the tra	0	0	0
JME, Material for Educa	Carboniferous of Pathway On the tra	0	0	0
JME, Material for Educa	Cretaceous of Pathway On the trace	0	0	0
JME, Material for Educa	Cretaceous of Pathway On the trace	0	0	0
JME, Material for Educa	Devon of Pathway On the traces of li	0	0	0
JME, Material for Educa	Devonian of Pathway On the traces	0	0	0
JME, Material for Educa	Wing of the ancient bird	1	0	0

Figure 3.4: Example of metadata completeness export from repositories

In the case of the experiments that measure metadata completeness, data are exported from each repository in .csv format and are analyzed using MS Excel

to deduct completeness for all elements. In these exports (Figure 3.4), each column represents a metadata element and each row represents a record in the repositories under examination. Their intersection is marked with "0" when no value is provided (incomplete), with "1" when a value is provided (completed) and more if 2 values are provided, 3, 4, etc.

Finally, after each one of the MQACP experiments that may yield problematic metadata elements, a version of a short guide with "Good and Bad Metadata Practices" is circulated to the content providers, containing examples of the correct and mistaken use of the problematic elements (Figure 3.5).


Good Practices for Metadata Providers

Title

Good Metadata Example

1. Preparing for avian influenza - separating flocks from wild birds
2. Description

This eight page document produced by DEFRA explains how poultry keepers should prepare for avian influenza (bird flu) by protecting their flocks from contact with wild birds. The document deals with housing, feed and welfare issues and also discusses protective measures for game birds and backyard flocks. The file is provided in PDF format and requires Adobe Acrobat reader.

URL reference:
http://portal.organicedunet.eu/index.php?option=com_metavi&loId=lo-085ed66f-b89e-11de-922e-3b751303d8ec

Belongs to the collection of **Intute**

Recommendation:
This title is long enough to provide an accurate idea of what this resource is about. Judging also by the description, it seems that this title is the fitting one.

Figure 3.5: Screenshot of the guide for metadata practices for content providers

In Table 3.1 below, some cases of examples that are provided within the "Good & Bad Metadata Practices" guide, are provided.

Title	*"Please use a more comprehensive title. For example the CRC acronym. can be refined as Cooperative Research Centre just to provide the user with a way to understand what this learning resource is about."*
Keyword	*"More keywords needed. Just one keyword is not enough, and even so. the keyword text here is misleading. These keywords should be provided separately as "turkey" and "poultry" along with some others, and not as one "turkey poultry"."*
Typical Age Range	*"...why is it that simple pictures of pigs in the snow with no scientific details on them cannot be used for children that are less than 10 years old? Couldn't these pictures be used in the context of a primary class?"*

Context	*"Since the age range is from 15 years old to undefined. it only makes sense that the Educational context cannot be limited to higher education but should also consider high school. Be very careful because in this sense, these two elements should not conflict."*

Table 3.1: Examples of comments in the "Good & Bad Metadata Practices" guide

3.2. Actors/Roles

For the application of MQACP, a series of specific actors are needed, to serve the roles that are defined below:

- **Metadata Expert (ME)**: The ME is the person that will apply the MQACP on the repositories under examination. The ME will design and carry out the MQACP methods, using the tools provided and will be the one that will collect the results, analyze them and present them to the repository manager. These people have a deep and profound knowledge on metadata and their use to describe digital content. In some cases, metadata experts from the side of the content providers may also be involved in the MQACP, not as designers and implementers of it, but in the same role as Domain Experts (described below),

- **Technical Expert (TE)**: The TE is the person that supports the application of MQACP related to all the technical needs. More specifically, the Technical Expert is the person that is usually in charge of the metadata authoring tool development and is capable of providing reports on the use of metadata elements. This person has the technical knowledge needed to support the metadata annotation from the domain experts,

- **Domain Expert (DE)**: The DEs are at the core of the MQACP implementation. They are the ones that invest the biggest amount of time and effort in the process. They are the ones that generate metadata annotations and reviews on metadata annotations that serve as an indication of metadata quality. Domain experts are the experts of the field/domain that the repository under examination offers resources on. These people have the knowledge that is needed to provide a meaningful set of metadata for the digital resources they describe, in order to make them findable from users,

3.3. Metadata Design Phase

In this phase, the metadata standard or specification to be used in the envisaged LOR is selected and the necessary modifications are made to "profile" it to meet the application context. More specifically, a metadata standard is chosen to fit the generic needs of the application domain and it's profiled and adapted based on the limitations and requirements of the field

it's applied to. The purpose of this phase is to elaborate on a commonly accepted metadata specification that will be adopted from all content providers. In cases where metadata are created from scratch for a set of digital resources, this process is easier. When looking at populating repositories with existing metadata records, or a mix between new and existing ones, the constraints of the existing metadata heavily influence the design of the new application profile. More specifically, obligations for the new set of metadata have to be decided based on the existing ones, to avoid problems during the content population process but also balance the costs involved in human effort.

Figure 3.6: Overview of main components of the Metadata Design Phase

3.3.1. Quality Assurance Methods

During this phase, the following methods/tools are deployed:

- A metadata understanding session is held in which metadata and domain experts are presented with a metadata standard and they are asked to provide their input related to its easiness, usefulness and appropriateness for the application domain,
- A preliminary hands-on exercise is organized where the experts are asked to use the existing standard to describe a small sample of resources.

3.3.2. Quality Assurance Tools

The tools that are used in this phase, are:

- A metadata application profile, documented in detail, both in a document and in a presentation, to be used by the domain experts that will be asked to evaluate the AP,
- The Metadata Understanding Session questionnaire that is used for assessing the proposed metadata elements,

3.3.3. Actors/Roles

Domain experts and metadata experts take part in the exercises/experiments to provide their assessment of the metadata elements proposed to them. They also participate in the hands-on annotation of the test set of resources that allows them to reflect on the proposed metadata elements in terms of their envisaged use. Finally, in this phase, the metadata expert that is in charge of applying MQACP is actively involved by coordinating the Metadata

Understanding Session, possibly with one or more assistants. The ME will also collect the input from the questionnaires and outcomes of the discussion and will "feed" them into the metadata design process.

3.3.4. Outcomes

During this phase, the following outcomes are generated through the process deployed:

- Quantitative assessment of the proposed metadata elements, coming from domain and metadata experts,
- Qualitative feedback that came out of the hands-on annotation of resources on paper,
- Completed paper-based metadata records from the domain experts that will be used to better define their needs related to the metadata.

3.3.5. Phase Milestone

This phase is considered as completed, once the metadata application profile that will be used for the annotation of the digital content is drafted, revised and completed. The application profile and technical bindings are given to the technical team that uses them to deploy the metadata authoring tool for the content providers to populate their collections with. In the case of collections that are already populated, the completed application profile is given to the technical team to make the necessary adaptations to the existing tools so that they reflect the requirements of the domain and metadata experts.

3.4. Testing Phase

In this phase, a test implementation of the content/repository management system can be used for hands-on experimentation with metadata. As it can be shown in Figure 3.7, during this phase, the only experiment that takes place is the quality review of a set of metadata records from metadata experts. Domain experts provide metadata for a limited set of resources, using the application profile that was discussed and elaborated in the previous phase. This process allows the domain experts to get accustomed to the application profile and the metadata experts to get some preliminary feedback on the use of metadata.

Figure 3.7: Overview of main components of the Testing Phase

3.4.1. Quality Assurance Methods

During this phase, the following methods are deployed:

- A hands-on annotation experiment takes place to ensure that metadata experts but mainly content annotators work with,
- A first review of the metadata records that are generated from the content management tool is carried out by the metadata expert.

3.4.2. Quality Assurance Tools

During this phase, the following tools are used:

- Test implementation of the metadata on an initial environment/tool, which brings up issues that maybe the metadata design phase has not considered so far,
- A set of core metadata quality criteria that are used by the metadata experts to carry out a quick review of the metadata provided by domain experts and identify basic problems with the metadata creation process.

3.4.3. Actors/Roles

Metadata experts and content annotators are involved in the processes deployed. The main role is the one of content annotators that use the test tool to upload/annotate a sample of resources which will serve as a test bed for the metadata application profile. Metadata experts are mostly involved in reviewing the metadata of the sample of resources.

3.4.4. Outcomes

During this phase, a "Good & Bad Metadata Practices" guide is generated that serves as a reference document for the rest phases of content population drawing from mistakes and good practices in the metadata creation process. In addition, a first set of recommendations for the metadata authoring tool is generated from the domain experts that used the tool to provide metadata annotations. This feedback directly affects the metadata offered through the tool and therefore it is relevant to mention it as part of the MQACP.

3.4.5. Phase Milestone

For this phase to be considered completed, the test implementation of the metadata annotation tool has to be online and the first set of resources have to be described with metadata on the tool. In addition, the Metadata Expert has to have carried out an initial assessment of these records and has to provide a set of guidelines and suggestions to the content providers to use during the next phase, through the "Good & Bad Metadata Practices" guide.

3.5. Calibration Phase

During this phase, the various technical components (web front-end, content management system, etc) are put together and part of the content is available online. Content providers are still involved in the process and more specifically continue to annotate resources using the tool(s) deployed. A larger

body of resources is now uploaded on the tool and a metadata peer review exercise takes place on a representative sample of resources.

Figure 3.8: Overview of main components of the Calibration Phase

As Figure 3.8 shows, the main task carried out during this phase is the peer review of the metadata records from the domain experts that are involved from the side of each content provider.

3.5.1. Quality Assurance Methods

During this phase, a peer review on a representative sample of resources from the deployed repositories is carried out. Resources under review are selected so as to reflect content coming from all content providers. Domain experts are assigned to a number of resources from different content providers and are given the peer-review grid described in Chapter 3.1 in order to assess a number of parameters related to the quality of the metadata records examined.

3.5.2. Quality Assurance Tools

During this phase, a metadata quality assessment grid is used to collect the peer-reviews coming from the domain experts. This grid is based on the metadata quality metrics proposed by a series of studies on information quality as these were aggregated in Chapter 2.2.4.

3.5.3. Actors/Roles

Domain experts are the key actors of this phase, as they are the ones that support the whole peer review process of the resources coming from content providers. After receiving the completed reviews, metadata experts have the responsibility of analyzing the results and also examining the submitted reviews in detail so that this feedback can be incorporated in the "Good & Bad Metadata Practices" guide produced in the previous phase.

3.5.4. Outcomes

During this phase, the following outcomes are generated through the process deployed:

- Peer review results that indicate the quality of the metadata available on the tool,
- Targeted feedback to content providers that comes from analyzing the peer review forms,

- Input to the documentation related to metadata practices aiming to aid content providers in their future tasks

3.5.5. Phase Milestone

For this phase to be considered completed, a 20% of the resources have to be uploaded and annotated with metadata in the respective repositories. In addition to this, peer-reviews have to be collected for a representative sample of resources and feedback has to be sent to the content providers in order to commence a process of revisions on the content uploaded so far. Finishing with the Calibration Phase, all the content providers have to be fully aware of the metadata requirements of their tasks. Through the "Good & Bad Practices" guide, the content providers are equipped with the necessary documentation to go into the Critical Mass Phase were the majority of the content will be uploaded to the repositories.

3.6. Critical Mass Phase

Critical mass is the phase during which the tool(s) have reached a high maturity level and are ready to accept large numbers of content with their respective metadata. The application profile used is now completed and final, so not a great deal of changes can take place and in addition a significant number of metadata records are available for the metadata experts to review and analyze.

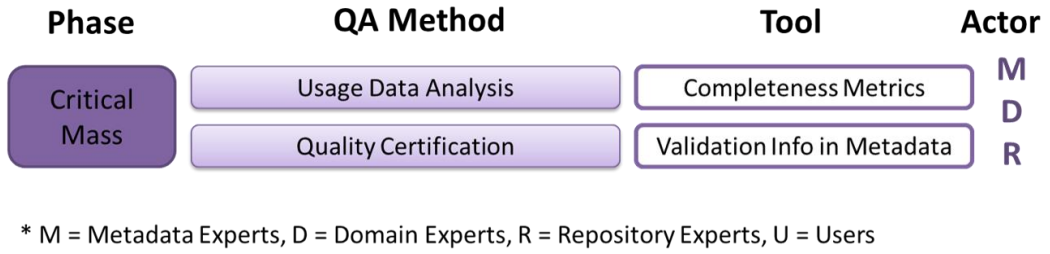


Figure 3.9: Overview of main components of the Critical Mass Phase

3.6.1. Quality Assurance Methods

During this phase, the following methods are deployed:

- Analysis of the usage data coming from a metadata authoring tool using specific completeness metrics. By analyzing this data, metadata experts can see to which extent the application profile is used and identify problematic elements,
- Metadata quality certification concept: Each content provider that uploads a resource and provides metadata for it has to "validate" it in order for the resource to become available online. The notion of validation verifies that when a content provider validates the resource uploaded, at the same time, he/she acknowledges that this resource and the accompanying metadata are of the desired quality.

3.6.2. Quality Assurance Tools

The main tool that is used in this phase is the metadata annotation tool through which, automatic completeness metrics are extracted for all the metadata records published in the federation of repositories. In addition to this, the exports from the annotation tool are analyzed in Microsoft Excel to export useful recommendations for the metadata annotation and enrichment. Another useful tool for this phase is the certification that each content provider attaches to the resources contributed to the federation. The notion of the certification acts as an incentive for the content providers to really curate the content they are contributed also in terms of metadata.

3.6.3. Actors/Roles

In this phase the main actors that facilitate the quality assurance methods are the metadata experts that analyze the data from the tool(s) used, based on a set of predefined completeness and multiplicity metrics. Metadata experts analyze the completeness of each one of the metadata elements that are selected in the metadata application profile and produce a set of recommendations for the content providers to revise the metadata for their resources accordingly. Such recommendations may prompt content providers to assign more keywords to the resources, use specific groups of elements more, etc.

3.6.4. Outcomes

During this phase, the following outcomes are generated through the process deployed:

- Minor revisions related to the application profile used may come out of this phase, mainly because of misconceptions on the use of some elements,
- Recommendations that come out of the metadata usage analysis which are directed back to the content providers to request that they improve the metadata they have provided

3.6.5. Phase Milestone

This phase is considered to be completed once the biggest part of the prospective resources are uploaded to the repositories (more than 70%-80%). The content management tools are deployed and only small changes can be made to accommodate any last-minute requests coming from the users. At this phase, the biggest part of the work for creating and populating a repository has already been carried out.

3.7. Regular Operation Phase

During regular operation, the metadata elements used in the tool(s) are considered to be final. The tools themselves and the content providers are now annotating resources regularly but not necessarily intensively like in the previous phase. This period covers the remainder of the LOR lifecycle.

Figure 3.10: Overview of main components of the Regular Operation Phase

As Figure 3.10 shows, in this phase, the methods that are deployed are more automated than in previous phases and are focused more on social contributions of the users of the digital content rather than the content providers themselves, which was the case in the previous phases.

3.7.1. Quality Assurance Methods

During this phase, the following methods are deployed:

- Regular analysis of the usage data for the annotation tool is carried out to make sure that completeness for all metadata elements is within the desired levels,
- An online peer review mechanism for metadata assists the process of ensuring high metadata quality. This review will be seamlessly incorporated on the service/tool provided to prompt the user to give reviews on the metadata of the resources as he/she accesses them,
- Informal mechanisms may be set up to support quality of metadata for the resources provided. These could include prizes and awards for resources that have been rated to have high quality in metadata based on the scores of the online review mechanisms.

3.7.2. Quality Assurance Tools

The tools that are being used for the remaining operational time of the repository are the following:

- The tools used to analyze the usage data coming from the content management tool. These tools are set up in a way so that automated reports are generated regularly to make sure that metadata for the old and new content are up to date,
- The metadata quality grid is still being used from content providers to regularly peer-review metadata records coming from other content providers. The goal here is to make this process automated and online, so that the users of the content can peer-review the metadata that accompany the resources, offering an extra quality assurance mechanism.

3.7.3. Actors/Roles

The actors that are actively involved in the quality assurance methods of this phase are larger in numbers than in the previous ones and contain the actual

users of similar services/tools. During the operation phase, the role of metadata experts is limited to analyzing the usage data from the tool whereas the role of the content users (consumers) is increasingly more important. For the online peer review mechanism to function properly, users have to support it with reviews they submit while they use the content available.

3.7.4. Outcomes

During this phase, the following outcomes are generated through the process deployed:

- Recommendations for content providers that come mainly from analyzing usage data from the environment/tool for content annotation,
- Insight related to metadata records from online peer review mechanism

3.7.5. Phase Milestone

This phase of the repository lifecycle starts when the majority of the content is deposited in it by content providers, described fully with metadata. Small additions to the existing content are being made continuously but no significant change on the content or metadata of the repository takes place. The metadata quality assurance mechanisms are still in place and operating in parallel. In the case of a substantial amount of content being deposited in the repository, we can assume that the repository will go again into the Critical Mass Phase, requiring a new examination of the application profile used and the metadata quality for the new content coming in.

3.8. Conclusions

This chapter presented in detail the Metadata Quality Assurance Certification Process that was designed to improve metadata quality in the repositories it is applied. For each phase, specific methods and tools that are used were presented and the main actors that need to be involved were discussed. More importantly, the outcomes that each phase produces were discussed, along with some basic criteria that have to be fulfilled in order to move to the next phase of the MQACP. In the following chapters we will present its application to different contexts, providing evidence that MQACP indeed improved the metadata quality of the repositories involved. From these different contexts, we expect to also identify requirements for the application of MQACP, in terms of the individuals needed to apply it, and the respective cost in terms of man hours. We expect that this application will act as a feedback loop for MQACP, to allow us to improve methods, refine tools, redefine roles and rethink the parameters of the experiments altogether.

4. MQACP Application on a Learning Federation

In this chapter, we present the results of the application of the MQACP on a real case of a federation of repositories, to assess its effectiveness on the overall metadata quality and attempt to address the research questions that were set forth in the beginning of the thesis. The proposed MQACP was applied on a real initiative, called Organic.Edunet which aimed at populating both a large but also high quality pool of learning resources related to Organic Agriculture & Agroecology and make them accessible through a web portal (www.organic-edunet.eu). More than 10.000 resources had to be described with quality metadata in multiple languages, supporting multilingual browsing through at least one part of the content. One part of this process involved the creation of an IEEE LOM Application Profile (AP), the Organic.Edunet metadata AP that would be used to annotate these resources with metadata and publish them on more than 10 LORs that were connected to the Organic.Edunet network.

4.1. Content Background

Before looking into the experiment carried out and the results yielded in the case of the learning federation, it is important to look at the details of the collections included in the Organic.Edunet federation. The biggest part of the content came from content providers within the initiative consortium and is described in Table 4.1. In Table 4.2, the same amount of content is presented, categorized per resource type that is image, text or video.

Provider[1]	**Type**[2]	**Quantity & Definition**[3]	**Format & Quality**[4]
AUA	Text, presentations, educational games, best practice guides	About 300 learning objects of various formats	Word, Powerpoint, PDF, HTML Flash, Stand alone applications; high-quality teaching material
USAMVB	Text, presentations, guides	About 100 learning objects of various formats	Word, PowerPoint, PDF, HTML; high-quality teaching material
Intute	Text, presentations, guides, other	About 4,450 learning objects of various formats	Word, Powerpoint, PDF, HTML Flash, Other; teaching material of various qualities
UMB	Text, presentations, lecture slides	About 100 learning objects of various formats	HTML, Word, PDF, Powerpoint, Flash; high-quality teaching material

[1] Short name of the participant who provides the content
[2] E.g. Text, image, movie, sound, music etc.
[3] E.g. 1,000 film clips, 2 million pages, 20,000 books etc.
[4] E.g. Format - JPEG, MPEG, Quicktime, HTML, PDF etc., Quality – Resolution, sampling rate, colour/greyscale etc.

Provider[1]	Type[2]	Quantity & Definition[3]	Format & Quality[4]
Miksike	Worksheets, methodological guidelines, informational texts etc.	About 300 learning objects of various formats	HTML, PDF, Word; high-quality teaching material
BCE	Text, presentations, lecture slides, movies, graphs	About 1000 learning objects of various formats	HTML, Word, PDF, Powerpoint, Flash, JPG, MPEG; high-quality teaching material, coloured, of varying resolution
MOGERT	Text, presentations, information slides, graphs	About 300 learning objects of various formats	Word, PowerPoint, PDF, HTML, SWI, JPG, MPEG; high-quality teaching material; coloured, of varying resolution
EULS	Text, presentations, graphs	About 30 learning objects of various formats	Word, PowerPoint, PDF; high-quality teaching material
	TOTAL	**6.770 objects**	

Table 4.1: Sources of content by Organic.Edunet content providers

Looking at the overall resources that would be contributed from the Organic.Edunet partners, we see that most of the content items are images (3.660) whereas 2.675 textual resources would also be contributed. Finally, Organic.Edunet partners would contribute a limited 245 videos to the learning federation.

Media Format	AUA	USAMVB	UMB	Miksike	BCE	Mogert	EULS	Intute
Text	60	45	30	80	400	100	10	1950
Image	165	55	50	190	550	180	20	2450
Video	75	0	20	30	50	20	0	50

Table 4.2: Categorisation of content resources per resource type

In Organic.Edunet, a significant amount of resources was contributed from external institutions that provided the content to the Organic.Edunet partners that annotated it with metadata. In Table 4.3, the breakdown of the content per provider is presented.

Provider	Type	Quantity and definitions
Soil Association	Text, PDF, graphs	About 250 learning objects of various formats
SEAE, Sociedad Española de Agricultura Ecologica	Text, presentations	About 400 learning objects of various formats
FAO Corporate Document Repository	Informational texts, papers, books	A selection of about 500 learning objects of various formats
FAO WAICENT Information Finder	Papers, informational texts etc.	A selection of about 500 learning objects of various formats (from the full list of available objects)
LEAD Virtual Research and Development Center	Guidelines, methodologies, papers	About 500 Learning resources of various formats
Organic ePrints	Texts (research papers, book chapters, reports, etc.)	A selection of about 800 learning objects of various formats
Organic Agriculture Information Access	Text (papers, reports)	About 400 learning objects of various formats
	TOTAL	**3.350 objects**

Table 4.3: Sources of public content in Organic.Edunet

4.2. Tools Used

Apart from the standard tools that are used to process the results of the MQACP, in each application case, some specific tools are deployed that really important for the interaction of the metadata annotators with the repositories involved. In the case of the educational repositories experiment, these are:

- The Organic.Edunet Metadata AP (Palavitsinis *et al.*, 2009a) that was based on the IEEE LOM Standard (IEEE LOM, 2002),
- The Confolio Tool that is a metadata authoring and publication tool which was used by content providers of the Organic.Edunet federation to describe their resources with metadata (Palavitsinis *et al.*, 2010).

4.3. Metadata Design Phase

During the Metadata Design Phase, a Metadata Understanding Session (MUS) was organized, focusing on the proposed metadata application profile, asking the domain experts to provide their evaluation of it, once a small explanation for all elements is provided.

Duration:	**2 hours**
Date:	**January 2009**
Annotated Objects:	**Not applicable**
Involved people:	**20 metadata IEEE LOM experts, organic educators & researchers**

Table 4.4: Metadata Understanding Session Parameters

This experiment took place during a focus group meeting in January 2009, where twenty (20) domain experts of the project were given an evaluation sheet with the three questions:

- Is this element easy for you to understand?
- Is this element useful for describing Organic.Edunet content resources?
- Should this element be mandatory, recommended or optional?

The experts that participated in the experiment followed the process described below:

1. A metadata expert presented an element from the metadata application profile by stating its definition, use and providing a simple example of its use,
2. Any question from the participants was asked to the metadata expert, to clarify completely the meaning and use of this element,
3. All participants were asked to decide on their answers on all three questions,
4. The next element was presented

Each element was assessed in terms of the questions posed with a score from one to five, five being the highest. The detailed results of this study are presented in Table 4.5. The name of each category of element in the Organic.Edunet IEEE LOM AP is only provided once and in the next elements of the same category it is replaced by "…". A first generic observation that can be made from this table is that the majority of elements were rated on an average of 3.5 or more related to both their easiness to understand and usefulness in the context of the educational repositories. Also educational elements in specific seem to be rated quite lower than the other elements of the AP which is problematic, taking into account that the primary use of the digital objects they describe, is education.

No	IEEE LOM AP	Easiness *Rating*	Usefulness *Rating*	Element Obligation	Obligation from Users
1	General / Identifier / Catalog	3.92	4	Mand	Mand
2	.../ Identifier / Entry	3.75	4	Mand	Mand
3	.../ Title	4.13	4.25	Mand	Mand
4	.../ Language	4.13	4.25	Mand	Mand
5	.../ Description	4.20	4.06	Mand	Mand
6	.../ Keyword	4.18	3.44	Recom	Mand
7	.../ Coverage	3.6	3.5	Recom	Optional
8	.../ Structure	2.82	2.93	Recom	Optional
9	.../ Aggregation Level	2.76	2.8	Optional	Optional
10	Life Cycle / Version	4	3.4	Optional	Optional
11	.../ Status	3.8	3.71	Optional	Optional
12	Contribute / Role	3.53	3.56	Recom	Recom
13	.../ Entity	3.71	3.75	Recom	Recom
14	.../ Date	3.82	3.4	Recom	Recom
15	Meta-Metadata / Identifier / Catalog	3.25	2.91	Recom	N/A
16	.../ Identifier / Entry	3.08	2.91	Recom	N/A
17	.../ Contribute / Role	3.86	3.29	Recom	Recom
18	.../ Contribute / Entity	3.71	3.21	Recom	Recom
19	.../ Contribute / Date	4.14	3.36	Recom	Recom
20	.../ Metadata Schema	3.67	3.08	Recom	Recom
21	.../ Language	4.27	3.45	Recom	Mand
22	Technical / Format	4.5	4.47	Recom	Mand
23	.../ Size	4.5	4.27	Recom	Mand
24	.../ Location	4.2	3.57	Recom	Mand
25	.../ Requirement / OrComposite / Type	3.65	3.06	Optional	Recom
26	.../ Requirement / OrComposite / Name	3.65	2.94	Optional	Optional
27	.../ Requirement / OrComposite / Minimum Version	3.65	2.88	Optional	Optional
28	.../ Requirement / OrComposite / Maximum Version	3.63	2.75	Optional	Optional
29	.../ Installation Remarks	3.75	3.13	Optional	Optional
30	.../ Other Platform Requirements	3.69	2.93	Optional	Optional
31	.../ Duration	3.92	3.58	Optional	Optional
32	Educational / Interactivity Type	3.2	3.5	Optional	Recom
33	.../ Learning Resource Type	3.75	3.93	Recom	Recom
34	.../ Interactivity Level	2.73	2.64	Optional	Optional
35	.../ Semantic Density	2.33	2.67	Optional	Optional
36	.../ Intended End User Role	3.56	3.33	Recom	Recom
37	.../ Context	3.75	3.47	Recom	Recom
38	.../ Typical Age Range	3.75	3.13	Recom	Recom
39	.../ Difficulty	3.33	3.25	Optional	Optional
40	.../ Typical Learning Time	3.27	2.67	Optional	Optional
41	.../ Description	2.92	2.90	Recom	Optional

No	IEEE LOM AP	Easiness *Rating*	Usefulness *Rating*	Element Obligation	Obligation from Users
42	.../ Language	3.43	3.25	Optional	Optional
43	Rights / Cost	4	4.08	Recom	Recom
44	.../ Copyright and Other Restrictions	3.86	3.92	Mand	Mand
45	.../ Description	3.64	3.54	Recom	Recom
46	Relation / Kind	3.75	3.7	Optional	Optional
47	.../ Resource / Identifier / Catalog	3.63	4.14	Optional	Optional
48	.../ Resource / Identifier / Entry	3.56	3.5	Optional	Optional
49	.../ Resource / Description	3.22	3.63	Optional	Optional
50	Annotation / Entity	3.91	3.14	Optional	Optional
51	... / Date	4.18	3.29	Optional	Optional
52	... / Description	4.18	3.57	Optional	Optional
53	Classification / Purpose	3.7	2.88	Recom	Optional
54	... / Taxon Path / Source / Taxon / Id	3.14	3	Recom	Optional
55	... / Taxon Path / Source / Taxon / Entry	3.14	2.83	Recom	Optional
56	... / Description	3.36	3.2	Optional	Optional
57	... / Keyword	3.55	3.5	Optional	Optional

Table 4.5: Results from evaluation of metadata elements from domain experts.

Table 4.5 presents and overview where it is evident that most elements were considered as easy to understand with 91.2% of the elements having an average rating of easiness coming from all the domain experts of 3 or more. 73.7% of the elements were also considered useful for the context of use. From this last question, we kept the relatively high percentage of people that were neutral in reference to the statement of the question, possibly indicating problems with the annotation of the actual resources in the next lifecycle stages of the repositories.

	Results				
Question	[0,1]	(1,2]	(2,3]	(3,4]	(4,5]
Is the element easy for you to understand?	0%	0%	8.8%	71.9%	19.3%
Is this element useful for describing educational resources?	0%	0%	26.3%	61.4%	12.3%

Table 4.6: Overall evaluation of the elements based on aggregated results

Table 4.7 brings up issues related to whether or not some of the elements should be mandatory, recommended or optional. More specifically, many of the elements changed their initial status because of the experiment carried out for the AP design. Still the aim of this experiment was mostly to expose the involved stakeholders to the notion of metadata and the potential elements that could be used. On 74% of the elements the domain experts agreed on the proposed obligation coming from the application profile, choosing the same one when evaluating the elements. Overall, the domain experts asked for more

mandatory elements, less recommended and more optional than the ones originally proposed.

	Mandatory		Recommended		Optional	
Question	Before	After	Before	After	Before	After
Should this element be mandatory, recommended or optional?	6	11	26	17	24	29
Percentile change in overall number of mandatory / recommended or optional elements	+83%		-34%		+21%	

Table 4.7: Overall status of the elements before and after the evaluation process

4.4. Testing Phase

During the Testing Phase of the Learning Federation, a small scale peer-review exercise was organized, where metadata experts obtain a set of metadata records that are provided by the domain experts in order to review it in terms of metadata quality. The main parameters of this experiment are presented in Table 4.8.

Duration:	**1 week**
Date:	**April 2009**
Annotated Objects:	**500 objects (≈5% of total expected number)**
Resources Reviewed:	**60 (15 per expert)**
Involved people:	**2 metadata experts**

Table 4.8: Peer-Review Experiment Parameters

In April 2009, a group of metadata experts carried out a small scale review of a set of metadata records. More specifically, a group of two (2) metadata experts took a random sampling of resources from each content provider, recording common mistakes in the metadata provided. The mistakes that were identified were documented in a guide titled "Good & Bad Metadata Practices" and circulated to the content providers to take into account while populating the learning repositories of the project with new resources. Mistakes were focused on ten (10) specific metadata elements, because mistakes from content providers concerned mainly these elements. Despite that, such guide has to include at all times references to all the mistakes the assigned metadata experts retrieve from the records they examine.

During this phase, a re-examination of the obligation of the elements in the application profile used, took place, where the metadata experts, guided by the results of the Metadata Understanding Session, as well as the small peer-review experiment, revisited some of the elements, changing their obligation and adding values to their vocabularies wherever this was necessary. Overall, as Tables 4.9 and 4.10 illustrate, changes in the status or vocabularies of 15 elements out of a total of 57 elements that were evaluated were proposed.

Some changes were also made regarding the vocabularies of the AP. These changes came up through the "comments" section of the questionnaire were all the participants gave unstructured comments on the use of the AP. Based on these comments, small adjustments were made to some of the vocabularies.

Category	Element	Previous Status	Users' Decision	Final Status
General	Structure	Recommended	Optional	Optional
Meta-metadata	Schema	Recommended	Mandatory	Mandatory
	Language	Recommended	Mandatory	Mandatory
Technical	Format	Recommended	Mandatory	Mandatory
	Size	Recommended	Mandatory	Mandatory
	Location	Recommended	Mandatory	Mandatory
	Minimum Version	Mandatory	Optional	Optional
	Maximum Version	Mandatory	Optional	Optional
Educational	Description	Recommended	Optional	Optional
Rights	Cost	Recommended	Mandatory	Mandatory
Annotation	Entity	Optional	Recommended	Recommended
	Date	Optional	Recommended	Recommended
	Description	Optional	Recommended	Recommended

Table 4.9. Changes in the status of the Organic.Edunet Application Profile elements

Category	Element	Changes	Values removed	Values added
Technical	Name	Revised operating systems' list	ms-windows	ms-windows XP, ms-windows Vista, other ms-windows versions, Linux
		Revised browser's list	-	Google chrome, Mozilla Firefox
Educational	Context	Revised environment's list	-	Post-graduate education, Pre-graduate education (*under higher education*)

Table 4.10: Changes in the vocabularies of the Organic.Edunet Application Profile

4.5. Calibration Phase

After allowing the domain experts for some time to work with the metadata authoring tool of the repositories and get acquainted with the new version of the application profile, we conducted a second peer-review experiment in a larger scale, involving the domain experts themselves. The main parameters of this experiment are shown in Table 4.11.

Duration:	**3 weeks**
Date:	**June 2009**
Annotated Objects:	**1.000 objects (≈10% of total expected number)**
Resources Reviewed:	**105 resources (≈5 per expert)**
Involved people:	**20 domain experts**

Table 4.11: Peer-Review Experiment Parameters

For this experiment, a peer-review mechanism for the metadata records of each content provider was set up. More specifically, a group of 20 domain experts was chosen that also had experience with metadata annotation. These experts came from all the content providers and they were assigned five (5) resources each for which they had to provide a review of their metadata using a pre-defined metadata peer-review grid.

This grid asserted metadata metrics based on a Likert scale from 1 to 5 (1 being the lowest). Specific examples were provided for each case, i.e. if completeness is rated with "1", then this means that most of the mandatory elements are not completed, few of the recommended are and almost none of the optional ones. In this way, the reviewers can be more confident and precise about their reviews.

In total, 105 reviews were received for corresponding resources, as some of the reviewers provided extra ones. Each reviewer was assigned resources coming from other providers, to keep the reviews impartial and in the end the review grids were centrally collected and analyzed to make deductions for the overall quality of metadata. In the end of this process, based on the analysis that was carried out, instructions were circulated to all partners, depending on the scores of the metrics used in the grids. These instructions regarded corrections to the metadata records of the LOs and in addition, specific comments were sent to partners based on free-text notes that were recorded on the peer-review grids. Table 4.12 contains an overview of the scores that the reviewers provided for all 105 resources.

Score	Completeness	Accuracy	Consistency	Objectiveness	Appropriateness	Correctness	Overall score
5	40%	51%	50%	69%	41%	69%	40%
4	45%	32%	28%	21%	33%	21%	37%
3	5%	10%	15%	6%	18%	9%	19%
2	9%	3%	1%	2%	5%	0%	0%
1	1%	1%	0%	0%	2%	1%	1%
no	1%	3%	6%	3%	1%	1%	3%

Table 4.12: Overview of the scores provided for 105 resources that were reviewed

As it can be noted from Table 4.12, most of the resources for all of the metrics were deemed as either excellent (score 5), or as pretty good (score 4). Some of

the resources were reported as mediocre (with a score of 3) and were sent to content providers to edit/enrich. The percentages of evaluations that did not provide a score for one or more of the criteria were negligible. Looking at each metric separately we can see that:

- 85% of the records were rated with scores of 4 or 5 out of 5, related to their completeness which is really high,
- 83% of the records were ranked as accurate related to the actual resource they describe,
- 78% of the records were consistent as regards the use of the elements with the envisaged use described in the Organic.Edunet application profile,
- Objectiveness was high as well with 90% of the records being reviewed with either 4 or 5 out of 5, showing that all records were completed in an unbiased way,
- 74% of the records were deemed to be filled out in the appropriate way in reference to the envisaged use within the Learning Federation of Organic.Edunet,
- Finally, correctness of the language used both in terms of syntax and grammar was really high with 90% of all records being rated with 4 or 5 out of 5.

Looking at some input that is not depicted in Table 4.12 but was gathered through the additional questions in the Metadata Quality Assessment form, we saw that from 105 records, 29 (28%) were not deemed to be of the desired quality (in terms of metadata) as that they could be published in the Learning Federation. Looking at the overall rating of the resources we see that 23% of them were rated with 3 or below in an average, so the high number of unfitting records for publication can be attributed to this. Given the fact that more than one quarter of the records examined were not fit for publication, the content providers were sent explicit instructions with the end of this phase, to review their records and correct any mistakes, based on a revised version of the “Good and Bad Metadata Practices” guide.

4.6. Critical Mass Phase

During the Critical Mass Phase, we performed an assessment of the completeness of the records that were populated in the Learning Federation. In Table 4.13 the basic parameters of the experiment are presented.

Duration:	**1 week**
Date:	**September 2009**
Annotated Objects:	**6.653 objects (≈60% of total expected number)**
Resources Analyzed:	**6.653**
Involved people:	**1 metadata expert**

Table 4.13: Completeness Assessment Parameters

For the next check-point in the lifecycle of the Learning Federation, it was deemed necessary to look at the way domain experts worked with the metadata elements that were defined in the metadata AP. What was expected here was that all the elements that were perceived as easy to understand and useful in the first experiment during the Metadata Understanding Session, would in turn be used in the actual content management tool. At this point, it is essential to point out that for a resource to be uploaded on the content management tool and described with metadata, there were not mandatory fields, coming from the interface of the tool. Each user could upload the resource which was assigned a URL automatically and other than that, it could be saved without imposing any other restrictions to the user. This means that each user could decide to provide from none to all metadata elements for any resource.

This part of the experiment was carried out when a first big sample of resources was uploaded and annotated with metadata in the Learning Federation. More specifically, 6.653 resources (almost 60% of the total number of resources that would be populated in the end of the project) were described with metadata and uploaded by content providers at the time of the experiment. A snapshot of the metadata completeness for all metadata elements of the application profile was taken at the time from the content management system of the LORs and all this data were analyzed to find out to which extent each metadata element was used. In this experiment, not all aspects of the quality of the metadata were examined but we mainly focused on their completeness. Some initial conclusions were made, including the fact that most mandatory elements were used at a satisfying level but when it came to the recommended, or even worse, the optional ones, their usage was well below 20% in most cases.

No	Mandatory Elements	Records	%
1	General / Title	6639	99,8%
2	.../ Language	6248	93,9%
3	.../ Description	6307	94,8%
4	Rights / Cost	1043	15,7%
5	Rights / Cost Copyright And Other Restrictions	1066	16,0%

No	Recommended Elements	Records	%
6	General / Keyword	850	12,8%
7	LifeCycle / Contribute / Role	763	11,5%
8	Educational / Intended End User Role	853	12,8%
9	.../ Context	678	10,2%
10	.../ Typical Age Range	252	3,8%
11	Rights / Description	511	7,7%
12	Classification	785	11,8%
No	**Optional Elements**	**Records**	**%**
13	General / Coverage	10	0,2%
14	.../ Structure	523	7,9%
15	LifeCycle / Status	22	0,3%
16	Educational / Interactivity Type	22	0,3%
17	.../ Interactivity Level	22	0,3%
18	.../ Semantic Density	14	0,2%
19	.../ Difficulty	9	0,1%
20	.../ Typical Learning Time	0	0%
21	.../ Language	22	0,3%
22	.../ Description	102	1,5%

Table 4.14: Completeness elements deployed in Confolio tool

As it is shown in Table 4.14, rights-related elements and all the recommended elements were heavily underused in the Learning Federation. The situation with the optional ones was even worse with completeness levels around 0.1 to 0.3%. Finally, it should be pointed out that during these measurements we are not looking at all the 57 elements that were included in the original application profile from the domain experts as some of them are filled out automatically and therefore looking at their completeness would not make sense (e.g. “General/Identifier”) and also because some of them make sense when all their sub-elements are completed, so instead of looking into each and every sub-element we look at the parent element to examine completeness (e.g. “LifeCycle/Contribute/Role” element).

Based on the aforementioned findings, recommendations were sent to all content providers, mainly targeted on the use of more keywords for each resource, more classification terms that were based on a pre-defined ontology on Organic Agriculture and finally the use of the educational elements of IEEE LOM to show the educational usage of the resources. Allowing for some time for the content providers to enhance their metadata, the next experiment that took place was designed mainly to check whether or not the requested metadata enhancements were carried out.

4.7. Regular Operation Phase

During the Regular Operation Phase of the Learning Federation, the total number of the targeted resources was populated within the repositories and enough time was given to the content providers to correct the metadata records based on recommendations that came out from the previous experiments. The parameters of the last experiment conducted are presented in Table 4.15.

Duration:	**1 week**
Date:	**September 2010**
Annotated Objects:	**11.000 objects (100%)**
Resources Analyzed:	**11.000**
Involved people:	**1 metadata expert**

Table 4.15: Completeness Assessment Parameters

During this phase, another analysis of the log files took place, for a total of approximately 11.000 resources. The purpose of this analysis was twofold. First of all it aimed at checking the overall metadata completeness and also at validating the experiments carried out so far, to see if they had the expected impact on the metadata quality. A second snapshot of the metadata completeness for all metadata elements of the application profile was taken on October 2010 from the Confolio tool. The data were analyzed to see if, at least in terms of completeness, the metadata quality was affected. Indeed, findings showed that the metadata completeness was significantly improved for all categories of elements.

In Table 4.16, the results for completeness of metadata records are presented in two distinct phases: The results from the earlier stages of "Critical Mass" and the results from the stage of "Regular Operation" for the Learning Federation. Comparing the completeness level in the two phases it becomes apparent that great improvement has been depicted in the resulting metadata because of the mechanisms that were setup. In the case of mandatory elements, all of them reached completeness levels of almost 100% with the exception of "Rights/Cost" that was high enough with 82.4% as well. In the recommended elements, all the elements were completed in 63.9% of the cases or more. Finally, with the exception of the "Educational/Typical Learning Time" and the "Educational/Description" element which confused many of the content providers, all the other optional elements were completed in more than 36.9% of the cases.

No	Mandatory Elements	Critical Mass		Regular Operation		
		Records	%	Records	%	Diff.
1	General / Title	6639	99.8%	10.968	100%	+0.2%
2	.../ Language	6248	93.9%	10.964	99.9%	+6%
3	.../ Description	6307	94.8%	10.968	100%	+5.2%
4	Rights / Cost	1043	15.7%	9.037	82.4%	+66.7%
5	.../ Cost Copyright & Other Restrictions	1066	16.0%	10.963	99.9%	+82.4%
No	**Recommended Elements**	**Critical Mass**		**Regular Operation**		**Diff.**
		Records	**%**	**Records**	**%**	
6	General / Keyword	850	12.8%	10.959	99.9%	+87.1%
7	LifeCycle / Contribute / Role	763	11.5%	8.467	77.2%	+65.7%
8	Educational / Intended End User Role	853	12.8%	9.037	82.4%	+69.6%
9	.../ Context	678	10.2%	8.885	81%	+70.8%
10	.../ Typical Age Range	252	3.8%	7.084	63.9%	+60.1%
11	Rights / Description	511	7.7%	10.135	92.4%	+84.7%
12	Classification	785	11.8%	8.151	73.6%	+61.8%
No	**Optional Elements**	**Critical Mass**		**Regular Operation**		**Diff.**
		Records	**%**	**Records**	**%**	
13	General / Coverage	10	0.2%	9.055	82.6%	+82.4%
14	.../ Structure	523	7.9%	9.048	82.5%	+74.6%
15	LifeCycle / Status	22	0.3%	4.349	39.7%	+39.4%
16	Educational / Interactivity Type	22	0.3%	4.085	36.9%	+36.6%
17	.../ Interactivity Level	22	0.3%	4.107	37.1%	+36.8%
18	.../ Semantic Density	14	0.2%	4.097	37%	+36.8%
19	.../ Difficulty	9	0.1%	4.113	37.1%	+37%
20	.../ Typical Learning Time	0	0%	45	0.4%	+0.4%
21	.../ Language	22	0.3%	5.794	52.3%	+52%
22	.../ Description	102	1.5%	1.627	14.7%	+13.2%

Table 4.16: Completeness comparison between Critical Mass & Regular Operation Phases

Completing the experiments carried out, it is evident that the completeness of the metadata records hosted in the respective repositories was significantly improved. Looking at some qualitative aspects that were measured through the peer-review experiments, we can also say with some degree of certainty that the same records are also of high quality related to the full set of quality metrics that was deployed.

4.8. Cost Implications

In Table 4.17, the cost analysis of the application of MQACP on the case of the Learning Federations is presented. In the table, the timing of each experiment

is presented along with the method used to conduct and the estimated time that this occupied the experts involved. The estimation is calculated in hours so that it can be reused independent of currencies and other country specific variables. In the end of the table, the total time needed is provided, along with the requirements in people.

Date/Period	No of Experts	Method	Time per item	No of items	Total time
12/2008	18 domain & 2 metadata experts	Focus group meeting for metadata design	2 hours per session	20 people	40 hours
5/2009	4 metadata experts	Metadata quality assessment from experts	20 minutes per record	60 records	20 hours
		Good & Bad Metadata Practices	8 hours	1 guide	8 hours
6/2009	1 metadata expert	Prepare process	5 minutes per record	105 records	8.8 hours
		Collect results	3 minutes per form	105 forms	6.9 hours
		Analyze & interpret	N/A	N/A	2 hours
	20 domain experts	Peer review conducted online	20 minutes per record	105 records	35 hours
9/2009	1 technical expert	Modify DB script to store results in excel	N/A	6.653 records	8 hours
		Export data from repositories	10 minutes per collection		1 hour
	1 metadata expert	Metadata completeness measurement	N/A		2 hours
9/2010	1 technical expert	Export data from repositories	10 minutes per collection	11.000 records	1 hour
	1 metadata expert	Metadata completeness measurement	N/A	N/A	2 hours
TOTAL	49 people			**TOTAL**	134.7 hours

Table 4.17: Cost of the MQACP application on the Learning Federation

In the case of experiments that involved experts, the time is constant and the total time spent rises as the number of experts involved increases. Such a case is the metadata understanding session experiment and the metadata quality

assessment from experts. Other methods such as the creation of a guiding document for the annotators are absolute, meaning that the time needed does not change significantly, independent of the number of experts involved. For other methods such as the analysis of the results of the online peer review, the timing needed was not known in advance and was calculated as the process was carried out from the metadata expert that analyzed the results. Finally, for the parts of the cost analysis where technical work was carried out, the respective technical experts were interviewed to document the exact time they needed to respective scripts to support the MQACP application. This initial estimation does not indicate much on its own, as there are no previous findings from literature that could be compared to this. It remains to see how much it will cost to apply the MQACP in the two remaining cases and then compare the results.

4.9. Conclusions

In the following table, an overview of all the methods that were used in the case of the Learning Federation is provided. These experiments were carried out in different phases of the educational repositories' development, involving different actors and concerning different sets of resources each time. All of them were personalized to fit the approach and particularities of the specific repositories.

Experiment	No of participants / records	Expertise	Date	Results	Tool
Metadata Understanding Session	20	Domain & metadata	1/2009	Quantitative & Qualitative	Questionnaire with Likert scale and multiple choice
Metadata Record review	4 / 60	Metadata	4/2009	Qualitative	N/A
Metadata Record review	20 / 105	Domain	6/2009	Quantitative & Qualitative	Questionnaire with Likert scale and free text comments
Log files analysis	6.600 records	Metadata	12/2009	Quantitative	Microsoft Excel
Log files analysis	11.000 records	Metadata	9/2010	Quantitative	Microsoft Excel

Table 4.18: Details of the Metadata Quality Assessment Certification Process Methods

Overall, the selected approach proved to be really efficient for the completeness of the metadata records as in most cases the improvement that was achieved through the experiments carried out is more than noticeable. Key elements of the application profile that were very crucial to the service provided and were not used in the beginning (like Educational elements),

were later used in a satisfying degree, allowing for advanced search mechanisms to be deployed on the content. At the same time, following this process of metadata quality assurance, greater visibility of the content was achieved with all the contents coming from the content providers being accessed more times than in the beginning of the process.

In total, the MQACP for the case of the educational repositories was deployed over a period of 20 months when a total of five (5) experiments were carried out, reaching an average metadata completeness of 67% per element with an improvement in metadata completeness between the two measurements of an average of 32%.

5. MQACP Application on a Cultural Federation

Completing the first experiment, we identified the need to validate the results of applying the MQACP to another case of a federation of repositories. In this chapter, we present the process as it was validated in the case of the cultural collections, along with the main results. A discussion concludes the chapter, focusing on the main outcomes of the experiments comparing them to the Learning Federation's case. The second case in which the MQACP was applied was the one of the cultural collections of Natural Europe (www.natural-europe.eu). A federation of six (6) repositories is populated with resources related to natural history that account for the digitized collections of respective Natural History Museums (NHMs) around Europe. A total of approximately 15.000 resources are populated in the Natural Europe federation. The metadata describing the content would be bilingual for all partners (English and native language) and specific requirements were also imposed on the content providers coming from the fact that one part of the metadata for the resources would be exposed to Europeana (www.europeana.eu), a Europe-wide cultural heritage digital library. In this case, the needs of the repositories, dictated for a repository with high quality metadata that would allow for the resources of the respective content providers to be easily accessible.

5.1. Content Background

In Table 5.1, the content coming from all six (6) content providers of Natural Europe is described. As the table shows, the content covers various themes of natural history from fossils to minerals and plants. In addition to this, Table 5.2 shows the distribution of the content in the main types offered, that is text, images and video.

Provider[5]	Type[6]	Quantity & Definition[7]	Format & Quality[8]
Natural History Museum of Crete (NHMC)	Text, images, videos	4.010 cultural objects on fossils, minerals & rocks, vertebrates, invertebrates, plants	Word, PDF, high-resolution JPG, MOV, MPEG4
Museu Nacional de História Natural (MNHNL)	Text, images, videos	2.060 cultural objects on botanic, zoology and geology	Word, PDF, high-resolution JPG, MOV, MPEG4

[5] Short name of the participant who provides the content
[6] E.g. Text, image, movie, sound, music etc.
[7] E.g. 1,000 film clips, 2 million pages, 20,000 books etc.
[8] E.g. Format - JPEG, MPEG, Quicktime, HTML, PDF etc., Quality – Resolution, sampling rate, colour/greyscale etc.

Provider[5]	Type[6]	Quantity & Definition[7]	Format & Quality[8]
Jura-Museum Eichstätt (JME)	Text, images, videos	1.650 cultural objects on fossils and living fossils	Word, PDF, high-resolution JPG, MOV, MPEG4
Arctic Centre (AC)	Text, images, videos	450 cultural objects on arctic flora and fauna and arctic research	Word, PDF, high-resolution JPG, MOV, MPEG4
Estonian Museum of Natural History (TNHM)	Text, images, videos	1.780 cultural objects on geology, botany and zoology	Word, PDF, high-resolution JPG, MOV, MPEG4
Hungarian Natural History Museum (HNHM)	Text, images, videos	4.210 cultural objects on zoology, botany and minerals	Word, PDF, high-resolution JPG, MOV, MPEG4
	TOTAL	**14.160 objects**	

Table 5.1: Sources of content by Natural Europe content providers

As in the case of the learning federations, in the cultural federation, we look at all the content based on the content type, to compare the tasks of metadata annotation for the two cases. In total, 5.870 text resources, 7.300 images and 1.550 videos would be uploaded on the Natural Europe federation.

Media Format	NHMC	MNHNL	JME	AC	TNHM	HNHM
Text	510	1.300	830	250	880	2.100
Image	2.500	1.000	800	200	800	2.000
Video	1.000	300	20	30	100	100

Table 5.2: Categorisation of content resources per resource type

5.2. Tools Used

Apart from the common set of tools that are used for collection, analysis and presentation of results, another set of tools that are related to metadata design and management are needed in the application of the MQACP. Therefore, in the case of the cultural repositories we used:

- The Europeana Semantic Elements (Haslhofer & Isaac, 2011) as a base schema upon which the Natural Europe-ESE Metadata Application Profile (AP) was developed with the aid of domain experts.

- The MultiMedia Authoring Tool (MMAT) that is a metadata authoring and publication tool which was used by content providers of the Natural Europe federation to populate the existing repositories with cultural content (Makris *et al.*, 2011).

5.3. Metadata Design Phase

During the Metadata Design Phase, domain experts were exposed to the Natural Europe-ESE application profile during a focus group meeting where they were presented with its elements and provided their input both through a questionnaire and also through observations and questions that were noted from the metadata experts.

Duration:	**2 hours**
Date:	**March 2011**
Annotated Objects:	**Not applicable**
Involved people:	**11 museum researchers & 1 metadata expert**

Table 5.3: Metadata Understanding Session Parameters

This experiment took place during a focus group meeting in March 2011, where 11 domain experts and 1 metadata expert were presented with the proposed application profile for their digital cultural collections and were given an evaluation sheet to provide their input. This input included answering the following questions:

- The element is easy to understand
- The element is useful for describing LOs in Natural Europe

Finally. the partners were asked if they though that that each element should be mandatory. recommended or optional. Additionally. during this phase. a preliminary metadata hands-on exercise took place where the content providers/subject matter experts. used the existing application profile to describe a set of their resources in paper. This exercise generated discussions on the Natural Europe ESE AP which needed many refinements based on the input of the content providers. Below. the results of the metadata application profile questionnaire are presented in a scale from 1 (low) to 5 (high) showing the preliminary understanding of the metadata elements from the content providers. Looking at Table 5.4, we see that apart from a few cases, most of the elements were rated in average with scores of 4 and above by the domain experts, both in terms of easiness as well as in terms of their usefulness for the application context. In addition to that looking at the agreement in the obligation of the elements between metadata and domain experts we see that for 55% of the elements the respective experts agreed in terms of being mandatory, recommended or optional.

No	Dublin Core	Easiness *Rating*	Usefulness *Rating*	Element Obligation	Obligation from Users
1	dc.title	4.82	4.64	Mand	Mand
2	dc.description	5	4.91	Mand	Mand
3	dc.rights	4.73	4.73	Mand	Mand

No	Dublin Core	Easiness *Rating*	Usefulness *Rating*	Element Obligation	Obligation from Users
4	dc.subject	3.55	4.7	Mand	Mand
5	dc.coverage	2.2	3.25	Mand	Optional
6	Europeana.URI	4.45	4.8	Mand	Mand
7	Europeana.country	5	4.91	Mand	Mand
8	Europeana.language	4.6	4.64	Mand	Mand
9	Europeana.provider	3.55	3.67	Mand	Mand
10	Europeana.data.provider	3.8	3.63	Mand	Mand
11	Europeana.CollectionName	4.82	4.56	Mand	Mand
12	Europeana.Rights	4.64	4.67	Mand	Mand
13	meta-metadata	3.22	3.25	Mand	Optional
14	dc.type	4.27	4.64	Recom	Mand
15	dc.language	5	4.91	Recom	Mand
16	dc.contributor	2.7	2.75	Recom	Recom
17	dc.creator	4.36	4.55	Recom	Mand
18	dc.publisher	3.82	4.2	Recom	Mand
19	dc.provider	3.64	3.64	Recom	Mand
20	dc.identifier	3.55	4.3	Optional	Mand
21	dcterms.created	4	4.1	Optional	Mand
22	dcterms.issued	3.36	3.4	Optional	Optional
23	dc.source	4.45	4.55	Optional	Mand
24	dc.relation	2.91	3.33	Optional	Optional
25	dc.format	4.45	3.8	Optional	Optional
26	dcterms.provenance	2.55	3.2	Optional	Optional
27	Europeana.object	4.2	5	Optional	Mand
28	Europeana.isShownBy	3	3.75	Optional	Mand
29	Europeana.isShownAt	3.2	3.63	Optional	Recom

Table 5.4: Results from evaluation of metadata elements from domain experts.

The overall analysis of the results agrees with the initial observation from the previous table. More specifically, for 82.8% of the elements, the domain experts agreed that they are easy to understand with scores of 3 out of 5 or more. Additionally, for 96.5% of the elements, the domain experts agreed that they would be useful to have when describing their digital cultural resources (Table 5.5).

	Results				
Question	[0,1]	(1,2]	(2,3]	(3,4]	(4,5]
Is the element easy for you to understand?	0%	0%	17.2%	34.5%	48.3%
Is this element useful for describing cultural resources?	0%	0%	3.4%	37.9%	58.6%

Table 5.5: Overall evaluation of the elements based on aggregated results

Table 5.6 shows that there were lots of differences in the obligations indicated by the designers of the metadata application profile in comparison to what the domain experts requested. Overall, domain experts asked for more mandatory and less recommended and optional elements than the original application profile.

	Mandatory		**Recommended**		**Optional**	
Question	Before	After	Before	After	Before	After
Should this element be mandatory, recommended or optional?	13	21	6	2	10	6
Percentile change in overall number of mandatory / recommended or optional elements	+62%		-66%		-40%	

Table 5.6: Overall status of the elements before and after the evaluation process

5.4. Testing Phase

During the Testing Phase of the cultural repositories, a peer-review exercise was organized on August 2011, where metadata experts obtain a set of metadata records that are provided by the domain experts in order to review it in terms of metadata quality. The main parameters of this experiment are presented in Table 5.7.

Duration:	**1 week**
Date:	**August 2011**
Annotated Objects:	**1000 objects (≈10% of total expected number)**
Resources Reviewed:	**100 (50 per expert)**
Involved people:	**2 metadata experts**

Table 5.7: Peer-Review Experiment Parameters

During the Testing Phase, a first working version of the MMAT was made available to the content providers. Curators in charge of the collections worked with the tool to provide us with their opinion on each element of the Natural Europe ESE AP. The results from the questionnaires provided new insight for the revisiting of the Natural Europe ESE AP. Based on the comments coming from the content providers, revisions were made to the MMAT and the application profile and a hands-on annotation of 10% of the resources in the cultural federation took place. 1.000 resources were annotated with metadata giving the opportunity to content providers to work with metadata and the tool in real situations. Once the process was completed a small group of metadata experts reviewed a sample of these 1.000 resources (almost 10% of them) based on the Metadata Quality Assessment Grid. The outcomes of this review were sent to all the content providers as a future roadmap but were also used to draft the first version of a metadata good and bad practices guide. The major revisions during this stage were the changes made in the application profile used. Table 5.7, shows the first version of the

application profile and maps it to the adapted one based on the domain experts input. As it can be seen below, some of the elements (14) changed their name to become easier to understand by the domain experts, i.e. switching from the traditional DC names to something else. Three (3) elements were broken down into more elements to store complex data about the cultural objects, so "dc.title" was substituted by "CHO Title" that provides the title of the cultural heritage object itself (the physical object) and "Object Titles" that stores the title of its digital counterpart, introducing in this was an FRBR[9] approach to the Natural Europe application profile. On the contrary, some elements were grouped and formed one element in the new application profile, covering the multiple element names with vocabularies for the new elements. As an example, in the case of "Object Creators", this element replaced the "dc.contributor", "dc.creator", "dc.publisher", "dc.provider" by incorporating the respective values in the form of a controlled vocabulary.

No	First version	Second version	No
1	dc.title	CHO Title	1
		Object Titles	2
2	dc.description	Object Descriptions	3
3	dc.rights	Copyrights	4
4	Europeana.Rights	Access	5
5	dc.subject	Classification	6
		Scientific Name	7
		Common Names	8
6	dc.coverage	CHO Temporal Coverage	9
7	dc.source	Object URL	10
8	Europeana.URI		
9	Europeana.country	CHO Spatial Coverage	11
10	Europeana.provider	CHO Creators	12
11	Europeana.data.provider		
12	Europeana.CollectionName		
13	meta-metadata	N/A	
14	dc.type	CHO Types	13
15	Europeana.language	Object Languages	14
16	dc.language		
17	dc.contributor	Object Creators	15
18	dc.creator		
19	dc.publisher		
20	dc.provider		
21	dc.identifier	Object Identifiers	16
22	dcterms.created	CHO Significant Dates	17
23	dcterms.issued	Object Creation Dates	18

[9] http://www.loc.gov/cds/downloads/FRBR.PDF

No	First version	Second version	No
24	dc.relation	Related Objects	19
25	dc.format	Object Formats	20
		CHO Mediums	21
26	dcterms.provenance	Object Extends	22
27	Europeana.object	Object Content Type	23
28	Europeana.isShownBy	Object Thumbnail URL	24
29	Europeana.isShownAt	Object Context URL	25
	N/A	CHO Keywords	26

Table 5.8: Mapping between first and second version of Natural Europe application profile

Coming out of the Testing Phase, the application profile now has 26 elements which were incorporated into the MMAT tool, meaning that all the content providers are now working with the new version of the application profile. From this point on, small changes can only take place within the elements themselves as the population of the repositories with the majority of the digital content begins.

5.5. Calibration Phase

More than 3.000 resources from the cultural collections were annotated with metadata during the end of the 1st year of the repository lifecycle, entering in this way what we call its "Calibration Phase". After the successful completion of the previous phase, a peer review exercise was organized on March 2012, where content providers were given a sample of resources coming from providers other than themselves to evaluate based on a set of predefined quality metrics. In total, 89 reviews were received for 89 different metadata records and corresponding resources, provided by ten (10) domain experts. The main parameters of this experiment are provided in Table 5.9.

Duration:	**3 weeks**
Date:	**March 2012**
Annotated Objects:	**3.000 objects (≈20% of total expected number)**
Resources Reviewed:	**99 resources (≈10 per expert)**
Involved people:	**10 domain experts**

Table 5.9: Peer-Review Experiment Parameters

In Table 5.10, the aggregated scores of the peer review experiment are presented. For each metric, the percentage of reviewers that gave the corresponding score are mentioned.

Score	Completeness	Accuracy	Consistency	Objectiveness	Appropriateness	Correctness	Overall score
5	16%	22%	36%	73%	32%	62%	26%
4	42%	43%	40%	13%	29%	22%	38%

Score	Completeness	Accuracy	Consistency	Objectiveness	Appropriateness	Correctness	Overall score
3	19%	13%	19%	4%	13%	13%	30%
2	6%	1%	3%	6%	14%	1%	3%
1	13%	9%	1%	4%	11%	1%	2%
no	3%	11%	0%	0%	0%	1%	0%

Table 5.10: Overview of the scores provided for all the quality metrics for 99 resources that were reviewed

From Table 5.10, the following conclusions can be drawn:

- 58% of the records were rated with scores of 4 or 5 out of 5, related to their completeness,
- 65% of the records were ranked as accurate related to the actual resource they describe,
- Consistency with the ESE CHO Application Profile is high with almost 76% of the reviews scoring them with 4 or 5 out of 5. This indicates that all elements are used in the appropriate manner and in accordance with the Natural Europe ESE Application Profile (AP),
- Objectiveness was both ranked really high with 86% of the records being reviewed with either 4 or 5 out of 5, which shows that all records were completed in an objective, unbiased way,
- 61% of the records were deemed to be filled out in the appropriate way in reference to the way that metadata would be used in the cultural repositories.
- Correctness in the language used in the metadata descriptions was also really high, rated almost at 84% for scores of 4 or 5 out of 5.

Looking at some input that was gathered through the reviews that is not present in Table 5.10, we saw that 43 out of 99 (43%) resources were not deemed to be of the desired quality (in terms of metadata) so that they could be published on the federation of cultural repositories. The reviewers indicated this but answering a Yes/No question in the end of each review form. This phenomenon may be attributed to the low scores in three of the metrics, namely completeness, accuracy and appropriateness but also to the fact that looking at the overall ratings provided by the reviewers, 35% of the records were rated with 3 or below which can explain a big portion of the records that were deemed as unfit for publication. Following these results, content providers were provided with targeted input related to their collections and the document containing good and bad practices for metadata annotation was revised. To check whether or not, the results that we collected from our domain experts, matched the actual situation in the cultural repositories, we took a closer look at the metadata completeness in the repositories themselves. To this end, we analyzed the annotated objects' metadata in terms of their completeness. The parameters of the experiment that took place on May 2012 are presented in Table 5.11.

Duration:	**1 week**
Date:	**May 2012**
Annotated Objects:	**3.417 objects (≈20% of total expected number)**
Resources Analyzed:	**3.417**
Involved people:	**1 metadata expert**

Table 5.11: Completeness Assessment Parameters

Table 5.12 shows the completeness of each element in the cultural repositories. Grouping the elements based on their obligation, we see that for most of the mandatory elements (except from Access), the completeness is really high. In the recommended elements, only "Object Languages" is problematic in terms of completeness whereas in the optional elements, problems are noticed for "Scientific Name", "Classification", "Common Names", "CHO Temporal Coverage", "CHO Mediums", "CHO Creators". "Object Creation Dates" and "Object Context URL".

No	Mandatory Elements	Records	%
1	CHO Titles	3.414	99,9%
2	CHO Keywords	3.306	96,8%
3	Object Titles	3.417	100,0%
4	Object Descriptions	3.066	89,7%
5	Object URL	3.408	99,7%
6	Object Thumbnail URL	3.318	97,1%
7	Object Content Type	3.417	100,0%
8	Copyrights	3.417	100,0%
9	Access	2.063	60,4%
No	**Recommended Elements**	**Records**	**%**
10	CHO Types	2.667	78,1%
11	Object Creators	3.009	88,1%
12	Object Languages	1.106	32,4%
No	**Optional Elements**	**Records**	**%**
13	Scientific Name	11	0,3%
14	Classification	10	0,3%
15	Common Names	0	0,0%
16	CHO Significant Dates	2.452	71,8%
17	CHO Temporal Coverage	725	21,2%
18	CHO Spatial Coverage	2.554	74,7%
19	CHO Mediums	227	6,6%
20	CHO Creators	0	0,00%
21	Object Creation Dates	560	16,4%
22	Object Identifiers	2.112	61,8%
23	Object Context URL	555	16,2%
24	Related Objects	1.797	52,6%
25	Object Formats	3.059	89,5%
26	Object Extents	1.989	58,2%

Table 5.12: Completeness for all elements of AP used in Metadata Authoring Tool.

Taking the elements that were problematic, we compared them to the results of the Metadata Understanding Session that was held in the Metadata Design Phase, keeping in mind their mappings to the first version of the application profile. In the case of "Access", "Object Languages" and "Object Creation Dates" there is no clear connection between the understanding and usefulness the domain experts provided and their limited use within the repositories. Despite that, for all the other elements that were not used as much as expected, there were problems either with the easiness to understand them or their perceived usefulness which might have led to their limited use afterwards.

No	Mandatory Elements	%	Easiness to Understand	Usefulness
1	Access	60,37%	4.64	4.67
2	Object Languages	32,37%	5	4.91
3	Scientific Name	0,32%	3.55	4.7
4	Classification	0,29%		
5	Common Names	0,00%		
6	CHO Temporal Coverage	21,22%	2.2	3.25
7	CHO Mediums	6,64%	4.45	3.8
8	CHO Creators	0,00%	3.68	3.65
9	Object Creation Dates	16,39%	4	4.1
10	Object Context URL	16,24%	3.2	3.63

Table 5.13: Problematic elements' completeness compared to Metadata Understanding Session ratings

To solve the aforementioned problems, we issued clear instructions and examples of use for the content providers and revisited the really problematic elements such as the classification group ("Scientific Name", "Classification" and "Common Names"), refining their description and explaining their envisaged use in the repositories.

5.6. Critical Mass Phase

After completing the set of the two experiments in the Calibration Phase, we decided to check on the metadata quality of the cultural collections once more. To this direction, usage data were exported from MMAT on September 2012 and they were analyzed in terms of completeness. The main parameters of the experiment are shown in Table 5.14 and the results are showcased in Table 5.15.

Duration:	**1 week**
Date:	**September 2012**
Annotated Objects:	**9.402 objects (≈62% of total expected number)**
Resources Analyzed:	**9.402**
Involved people:	**1 metadata expert**

Table 5.14: Completeness Assessment Parameters

Table 5.15 shows the completeness of each element after the last measurement and the guidelines that were provided to the content providers. In this table we see that now all mandatory elements are completed in a satisfying degree, having solved the problem with the "Access" element and also the problem with the "Object Languages" has improved significantly. Finally, looking at the optional elements as a total we see that from an average completeness of 33.6% they rose to an average of 50.8% with only two elements presenting completeness below 35% which dropped the average significantly.

No	Mandatory Elements	Records	%
1	CHO Titles	9.253	98.4%
2	CHO Keywords	9.141	97,22%
3	Object Titles	9.402	100%
4	Object Descriptions	9.323	99,2%
5	Object URL	9.349	99,4%
6	Object Thumbnail URL	8.291	88,2%
7	Object Content Type	9.402	100%
8	Copyrights	9.402	100%
9	Access	9.402	100%
No	**Recommended Elements**	**Records**	**%**
10	CHO Types	7.362	78,3%
11	Object Creators	8.095	86,1%
12	Object Languages	5.796	61,7%
No	**Optional Elements**	**Records**	**%**
13	Scientific Name	5.629	59,9%
14	Classification	3.717	39,5%
15	Common Names	3.931	41,8%
16	CHO Significant Dates	4.332	46,1%
17	CHO Temporal Coverage	1.689	18%
18	CHO Spatial Coverage	6.590	70,1%
19	CHO Mediums	309	3,3%
20	CHO Creators	4.627	49,2%
21	Object Creation Dates	4.469	47,5%
22	Object Identifiers	5.996	63,8%
23	Object Context URL	3.418	36,4%
24	Related Objects	8.291	88,2%
25	Object Formats	9.402	100%
26	Object Extents	4.436	47,2%

Table 5.15: Completeness for all elements of AP used in Metadata Authoring Tool.

Coming out of the Critical Mass Phase, the cultural repositories are populated with more than half of the envisaged resources and most of the elements are completed in a satisfying degree, taking into account (a) their original status during the Calibration Phase and (b) their obligation, meaning that we did not expect or demand for the optional elements to be completed as much as the

mandatory ones. To make sure that the results that we got from analyzing completeness in the repositories were consistent, we conducted another small peer-review experiment with the participation of four domain experts that reviewed a small sample of 34 resources. The main parameters of this experiment are shown in Table 5.16 and the results in Table 5.17.

Duration:	**3 weeks**
Date:	**September 2012**
Annotated Objects:	**9.402 objects (≈62% of total expected number)**
Resources Reviewed:	**34 resources (≈9 per expert)**
Involved people:	**4 domain experts**

Table 5.16: Peer-Review Experiment Parameters

Looking at Table 5.17, we see that the results compared to the previous peer-review experiment have significantly improved for all the quality metrics which was also reflected on the completeness measurement that was carried out during the same period. Overall we see that completeness was improved in terms of having fewer resources rated with 1 out of 5 (dropped from 13% to 0%). Accuracy went from 65% to 80%, Consistency went from 76% to 80% as well and Objectiveness remained steady around 85%. Appropriateness went from 61% to 76% and Correctness went from 84% to 88%.

Score	Completeness	Accuracy	Consistency	Objectiveness	Appropriateness	Correctness	Overall score
5	38%	59%	68%	79%	47%	79%	53%
4	29%	21%	12%	6%	29%	9%	29%
3	26%	9%	12%	6%	15%	12%	9%
2	6%	12%	9%	9%	6%	0%	9%
1	0%	0%	0%	0%	3%	0%	0%
no	0%	0%	0%	0%	0%	0%	0%

Table 5.17: Overview of the scores provided for all the quality metrics for 34 resources that were reviewed

5.7. Regular Operation Phase

During the Regular Operation Phase, the tools of the cultural collections are completed and the application profile is now stable and not likely to change. The content providers are completing the upload and annotation of the digital resources they are contributing, also making sure to revise the problematic ones from previous phases. To assess the quality of the metadata in the repositories in terms of their completeness we organized another experiment on May 2013, one year after the first measurement. The parameters of the experiment are presented in Table 5.18 and its results in Table 5.19.

Duration:	**1 week**
Date:	**May 2013**
Annotated Objects:	**11.375 objects (≈80% of total expected number)**
Resources Analyzed:	**11.375**
Involved people:	**1 metadata expert**

Table 5.18: Completeness Assessment Parameters

Looking at the comparison between the two measurements in Table 5.19, we see that all the mandatory elements were completed in 97.2% of the cases or more, which is really high, also taking into account that the number of mandatory elements is almost double the one in the case of the Learning Federation (5). For the limited set of recommended elements (3), all of them are completed in satisfying percentages with "Object Languages" showing a big improvement between the two measurements. Finally, in the case of optional elements, with the exception of "CHO Mediums" (17.9%) and "CHO Temporal" (18.5%), all the other elements were used in more than 48.1% of the cases. Here, the case of "CHO Significant Dates" is a problematic one, as this was the only element of the application profile that dropped in terms of completeness, significantly (23.7%).

No	**Mandatory Elements**	**Critical Mass Phase**		**Regular Operation**		**Diff**
		Record	**%**	**Record**	**%**	
1	CHO Titles	3.414	99.9%	11.323	99.5%	-0.4%
2	CHO Keywords	3.306	96.8%	11.268	99.1%	+2.3%
3	Object Titles	3.417	100%	11.375	100.00%	0%
4	Object Descriptions	3.066	89.7%	11.247	98.9%	+9.2%
5	Object URL	3.408	99.7%	11.361	99.9%	+0.2%
6	Object Thumbnail	3.318	97.1%	11.055	97.2%	+0.1%
7	Object Content Type	3.417	100%	11.375	100%	0%
8	Copyrights	3.417	100%	11.375	100%	0%
9	Access	2.063	60.3%	11.375	100%	+39.7%
No	**Recommended Elements**	**Critical Mass Phase**		**Regular Operation**		**Diff**
		Record	**%**	**Record**	**%**	
10	CHO Types	2.667	78.1%	8.524	74.9%	-3.2%
11	Object Creators	3.009	88.1%	9.931	87.3%	-0.8%
12	Object Languages	1.106	32.4%	8.926	78.5%	+46.1%
No	**Optional Elements**	**Critical Mass Phase**		**Regular Operation**		**Diff**
		Record	**%**	**Record**	**%**	
13	Scientific Name	11	0.3%	7.401	65.1%	+64.8%
14	Classification	10	0.3%	6.444	56.7%	+56.4%
15	Common Names	0	0%	5.659	49.8%	+49.8%
16	CHO Significant Dates	2.452	71.8%	5.475	48.1%	-23.7%
17	CHO Temporal	725	21.2%	2.101	18.5%	-2.7%
18	CHO Spatial Coverage	2.554	74.7%	8.117	71.4%	-3.3%
19	CHO Mediums	227	6.6%	2.039	17.9%	+13.3%

20	CHO Creators	0	0%	6.607	58.1%	+58.1%
21	Object Creation Dates	560	16.4%	7.559	66.5%	+50.1%
22	Object Identifiers	2.112	61.8%	7.775	68.4%	+6.6%
23	Object Context URL	555	16.2%	4.789	42.1%	+25.9%
24	Related Objects	1.797	52.6%	5.353	47.1%	-5.5%
25	Object Formats	3.059	89.5%	10.724	94.3%	+4.8%
26	Object Extents	1.989	58.2%	10.138	89.1%	+30.9%

Table 5.19: Completeness comparison for all elements between Critical Mass and Regular Operation Phases

Once more, to cross-check the results from the completeness measurement, a peer-review experiment was carried out with the help of domain experts to assess multiple parameters of metadata quality, in addition to completeness. The main parameters of the experiment that took place on May 2013, are presented in Table 5.20 and the results in Table 5.21.

Duration:	**3 weeks**
Date:	**May 2013**
Annotated Objects:	**9.402 objects (≈62% of total expected number)**
Resources Reviewed:	**90 resources (≈9 per expert)**
Involved people:	**10 domain experts**

Table 5.20: Peer-Review Experiment Parameters

As Table 5.21 shows, Completeness has increased since the previous peer-review (66% VS 57%) for the resources rated with 4 or 5 out of 5. With the same two ratings in mind, Accuracy (84% VS 80%), Consistency (84% VS 80%), Objectiveness (95% VS 85%) and Appropriateness (85% VS 76%) have increased their values. Correctness dropped slightly (87% VS 88%), actually staying at around the same level.

Score	Completeness	Accuracy	Consistency	Objectiveness	Appropriateness	Correctness	Overall score
5	40%	60%	74%	89%	61%	74%	52%
4	26%	24%	10%	6%	24%	13%	30%
3	18%	8%	7%	1%	4%	9%	10%
2	13%	3%	4%	1%	6%	0%	3%
1	3%	4%	4%	3%	4%	3%	5%
no	0%	0%	0%	0%	0%	0%	0%

Table 5.21: Overview of the scores provided for all the quality metrics for 90 resources that were reviewed

Completing the experiments carried out, we saw that the completeness of the metadata records was greatly improved in all but one cases. In addition, when a subset of the same records was examined in terms of other qualitative metrics the records were found to be improved in all the metrics except from one that remained more or less steady.

5.8. Cost Implications

In Table 5.22, the cost analysis of the application of MQACP on the case of the Cultural Federations is presented. In the table, the timing of each experiment is presented along with the method used to conduct and the estimated time that this occupied the experts involved.

Date/Period	No of Experts	Method	Time per item	No of items	Total time
3/2011	11 domain & 1 metadata experts	Focus group meeting for metadata design	2 hours per session	12 people	24 hours
8/2011	2 metadata experts	Metadata quality assessment from experts	20 minutes per record	100 records	33.3 hours
		Good & Bad Metadata Practices	8 hours	1 guide	8 hours
11/2011	1 metadata expert	Prepare process	5 minutes per record	100 records	8.3 hours
		Collect results	3 minutes per form	99 forms	5 hours
		Analyze & interpret	N/A	N/A	2 hours
	10 domain experts	Peer review conducted online	20 minutes per record	99 records	33 hours
6/2012	1 technical expert	Modify DB script to store results in excel	N/A	3.417 records	2 hours
		Export data from repositories	10 minutes		1 hour
	1 metadata expert	Metadata completeness measurement	N/A		2 hours
9/2012	4 metadata experts	Peer review conducted online	20 minutes per record	34 records	11.3 hours
9/2012	1 technical expert	Export data from repositories	10 minutes	9.402 records	1 hour
	1 metadata expert	Metadata completeness measurement	N/A		2 hours
5/2013	10 domain experts	Peer review conducted online	20 minutes per record	90 records	30 hours
5/2013	1 technical expert	Export data from repositories	10 minutes	11.375 records	1 hour
	1 metadata expert	Metadata completeness measurement	N/A		2 hours
TOTAL	45 people			**TOTAL**	166.5 hours

Table 5.22: Cost of the MQACP application on the Cultural Federation

Having the experience from the Learning Federation in mind, we see that in this case, the total time spent to apply MQACP is higher, with 166.5 hours in comparison to 134.7 in the case of the Learning Federation. People involved were less as in this case, 45 people were involved in the experiments whereas for the Learning Federations, 49 people participated in the experiments. Nevertheless, we see that the average time per person involved in each experiment is close with 2.8 hours per person in the case of Learning Federations and 3.7 for the Cultural Federations. In reality, the biggest difference between the two experiments lies in the fact that in the Cultural Federarion, more peer review experiments took place, adding more hours to the overall cost. More specifically, in the Cultural Federation, 323 records were peer-reviewed, whereas in the Learning Federation, only 165 records were reviewed, which dropped the overall time by more than one hour per expert. Theoretically, the experiment would have worked even with less resource peer-reviewed, so this is not considered a major difference in terms of cost.

5.9. Conclusions

In Table 5.23, an overview of all the methods that were used in the case of the cultural repositories is provided. As it can be seen, the actual methods that were needed in the case of the cultural repositories were more than the ones in the Learning Federation case (8 VS 5), which came to happen mainly because of the extra metadata record reviews that were held.

Experiment	No of participants / records	Expertise	Date	Results	Tool
Metadata Understanding Session	12	Domain & metadata	3/2011	Quantitative & Qualitative	Questionnaire with Likert scale and multiple choice
Metadata Record review	2 / 100 (records)	Metadata	8/2011	Qualitative	N/A
Metadata Record review	10 / 99 (records)	Domain	3/2012	Quantitative & Qualitative	Questionnaire with Likert scale and free text comments
Log files analysis	2/ 3.417 (records)	N/A	5/2012	Quantitative	Microsoft Excel
Metadata Record review	4 / 34 (records)	Domain	9/2012	Quantitative & Qualitative	Questionnaire with Likert scale and free text comments
Log files analysis	2 / 9.402 (records)	N/A	9/2012	Quantitative	Microsoft Excel
Metadata Record review	10 / 90 (records)	Domain	5/2013	Quantitative & Qualitative	Questionnaire with Likert scale and free text comments

Experiment	No of participants / records	Expertise	Date	Results	Tool
Log files analysis	2 / 11.375 (records)	N/A	5/2013	Quantitative	Microsoft Excel

Table 5.23: Details of the Metadata Quality Assessment Certification Process Methods that were used

Looking at the results yielded in the case of the cultural repositories, it is evident that the use of the experiments of Table 5.23, greatly improved the quality of the metadata records in the repositories, both in terms of completeness but also in terms of the other metrics that were assessed. Looking at the actual completeness of recommended and optional elements in the cultural collections, it is evident that there's still room for improvement but the fact that all nine of the mandatory elements are completed in almost 100% of the cases, is a positive step towards enhancing the overall quality of the records.

Focusing more on the other quality metrics used within this work, we present in Table 5.24, the aggregated results from the peer-review experiments that were carried out during the entire 26-month period that the MQACP was applied on the cultural federations. Looking at completeness (which was also proven high through the practical results) we see that it remained steady for the reviewers, although the actual records changed a lot through the course of the time. This may be attributed to their changing expectations through their involvement with metadata annotation. Looking at other metrics we see that Accuracy, Consistency, Objectiveness, Appropriateness and Correctness were in general improving throughout the experiments with some minor deviations of 1-2%. The overall score was also greatly improved from March 2012 to September 2012, remaining relatively high after that. Overall we see that the biggest improvement in the perceived quality of the records took place between March and September 2012, which was mainly during the Calibration Phase of the repository lifecycle.

Metric / Value	Complete (4 or 5)	Accurate (4 or 5)	Consistent (4 or 5)	Objective (4 or 5)	Appropriate (4 or 5)	Correct (4 or 5)	Overall Score
March 2012	68%	65%	76%	86%	61%	84%	64%
September 2012	67%	80%	80%	85%	76%	88%	84%
May 2013	66%	84%	84%	95%	85%	87%	82%

Table 5.24: Aggregated results from Quality Peer Review experiment.

In total, the MQACP for the case of the cultural repositories was deployed over a period of 26 months, when a total of eight (8) experiments were carried out, reaching an average metadata completeness of 74.2% per element with an improvement in metadata completeness between the two measurements of an average of 16%.

6. MQACP Application on a Research Federation

In the following chapter, we present a different case than the first two cases where the MQACP was applied and tested. We understood although the domain of application was different in both initial cases, they still presented similarities that may affect the results of the experiments. The most fundamental ones were that in both projects the content in the federations was more or less in the vicinity of 15.000 resources and also all metadata for them were created from scratch, in parallel to the application of MQACP. To address these similarities, we chose a case where the resources were significantly more and also where metadata for the content already existed and the task for the content providers was to enrich the metadata rather than create it from scratch. Therefore, the third case in which the MQACP was tested is VOA3R (http://voa3r.eu), a federation of nineteen (19) institutional repositories with scientific and research data to a total of more than 2.500.000 resources. For a core set of nine (9) content providers, 71.316 objects would be enriched in at least two languages (English and native). The main focus of the application of MQACP is this last set of 70.000 resources, offering an interesting case study in terms of the magnitude of the resources examined. The two federations of the Cultural and Learning repositories examined in Chapters 5 and 6 are evolving in parallel through the different stages that MQACP.

6.1. Content Background

In order to provide a better image of the collections of VOA3R, Table 6.1 presents the main characteristics of the collections that bring their content. As it can be seen, this case is completely different than the other two that were examined so far, since one of the repositories examined possess the same resources that in the other two cases, were contributed by all the content providers, e.g. the case of ICROFS.

Provider[10]	**Type**[11]	**Quantity & Definition**[12]	**Format & Quality**[13]
SLU	Theses, papers, articles, reports, books, book chapters, conference papers, proceedings, factsheets, article reviews, data sets	7.453 objects on Landscape Planning, Horticulture and Agricultural Science, Natural Resources and Agricultural Sciences, Forest Sciences, Veterinary Medicine and Animal Science	PDF, Word, HTML, Excel

[10] Short name of the participant who provides the content
[11] E.g. Text, image, movie, sound, music etc.
[12] E.g. 1,000 film clips, 2 million pages, 20,000 books etc.
[13] E.g. Format - JPEG, MPEG, Quicktime, HTML, PDF etc., Quality – Resolution, sampling rate, colour/greyscale etc.

Provider[10]	Type[11]	Quantity & Definition[12]	Format & Quality[13]
ICROFS	Papers	12.097 objects on research on organic food and farming systems	PDF, Word, HTML
CULS	Papers, news items	2.111 objects on agriculture and rural development; agricultural economics, management, agribusiness, agrarian policy, information and communication technologies, information systems, e-business, social economy and rural sociology	PDF, Word, HTML
ACTA-INFO	Conference papers	981 objects on ICT in agriculture	PDF, Word, HTML
UHASSELT	Research papers, reports	12.624 objects on oceanography, Water, Fisheries, Environment, Aquatic science	PDF, Word, HTML
INRA	Papers, books, book chapters, working papers, reports	11.619 objects on agriculture, environment and food	PDF, Word, HTML
ARI	Monographies	343 objects on agriculture	PDF, Word, HTML
CINECA	Papers, data	23.988 objects on agriculture and aquaculture	PDF, Word, HTML, Excel
AUA	Graduate, post-graduate and PhD theses, paper abstracts, images	100 objects on viticulture and organic agriculture	PDF, Word, HTML, JPEG
	TOTAL	**71.376 objects**	

Table 6.1: Sources of content by VOA3R content providers

In the analysis of the experimental context that will follow, the expected amount of resources for the VOA3R federation will be calculated based on the VOA3R content providers (71.376) and not the external ones (2.500.000). In Table 6.2, the breakdown of the content per different type is provided. We see that in the case of VOA3R, most of the content that was contributed was texts, differentiating this collection significantly from the other two.

Media Format	SLU	ICROFS	CULS	ACTA-INFO	UHASSELT
Text	7.453	12.097	2.111	981	12.624
Image	0	0	0	0	0
Video	0	0	0	0	0

Media Format	INRA	ARI	CINECA	AUA	
Text	11.619	343	23.988	100	
Image	0	0	0	250	
Video	0	0	0	0	

Table 6.2: Categorisation of content resources per resource type

Finally, a significant difference of the VOA3R experiment in comparison to the other two, is that no metadata annotation took place from scratch. All the metadata that were contributed to the repository federations, came from heterogeneous repositories. This phenomenon heavily influenced and made more difficult, the process of designing the metadata application profile to be used in the VOA3R Federation. In many cases, a few metadata elements were initially completed, making it really difficult for the metadata experts to decide on a minimum set of mandatory elements that all collections must have in all of their records.

6.2. Tools Used

The specialized tools that were used in the case of the Research/Scientific Repositories are the following:

- The VOA3R Metadata AP (Diamantopoulos et al., 2011) that was based on the Dublin Core standard (Weibel & Koch. 2000), enhanced with some additional elements to fit the application domain of scientific publications and research data,

- The Confolio Tool that is a metadata authoring and publication tool which was used by content providers of the VOA3R federation to describe their resources with metadata (Ebner et al. 2009).

6.3. Metadata Design Phase

During the Metadata Design Phase, a focus group meeting was organized in July 2011 during which, 16 domain experts were called to provide their opinion on the tentative application profile that would be used to describe the research data and scientific resources of the content providers.

Duration:	**2 hours**
Date:	**June 2011**
Annotated Objects:	**Not applicable**
Involved people:	**16 domain experts**

Table 6.3: Metadata Understanding Session Parameters

The process applied was the following:

1. Experts were presented with the proposed metadata application profile based on the Dublin Core standard,
2. Experts used printed forms to describe one resource each. using the proposed metadata application profile,
3. Experts used a questionnaire to provide their input related to (a) how easy it was for them to understand the metadata elements presented and (b) how useful they thought the specific metadata elements were for their institutional collections. Finally, they were also asked to identify for each element, the obligation they thought it should have in the final application profile as well as any comments related to the elements.

As it is shown in Table 6.3, thirty-five (35) elements were evaluated by the domain experts. The experts wanted to have 12 of them be mandatory, 6 of them as recommended and 17 of them as optional elements in the final version of the application profile. The original distribution of the elements was 13 mandatory, 14 recommended and 8 optional. This indicates a general tendency of content providers to ask for more optional elements than recommended, whereas mandatory remained more or less the same.

No	Dublin Core	Easiness *Rating*	Usefulness *Rating*	Obligation from AP	Obligation from Users
1	dcterms.title	4.8	5	Mand	Mand
2	dcterms.alternative	4.27	4	Optional	Optional
3	dcterms.creator	4	4.83	Recom	Mand
4	ags.creatorPersonal	4.29	3.93	Mand	Mand
5	ags.creatorCorporate	4.5	3.93	Mand	Optional
6	dcterms.contributor	3.93	3.93	Recom	Optional
7	dcterms.publisher	4.93	4.73	Recom	Mand
8	dcterms.date	4.47	4.93	Mand	Mand
9	dcterms.identifier	4.13	4.8	Recom	Recom
10	dcterms.language	4.93	5	Mand	Mand
11	dcterms.format	4.33	4.57	Recom	Mand
12	dcterms.source	3.33	3.69	Mand	Mand
13	dcterms.type	4.53	4.8	Mand	Recom
14	meta-metadata.catalog	2.67	3.64	Mand	Optional
15	meta-metadata.entry	2.64	3.93	Mand	Optional
16	mm.contribute.role	3	3.46	Mand	Optional
17	mm.contribute.entity	3	3.92	Mand	Optional
18	mm.contribute.date	3.07	3.75	Mand	Optional
19	mm.metadata schema	3.14	3.93	Mand	Mand
20	mm.language	3.86	4.25	Recom	Optional
21	dcterms.rights	3.87	4.33	Recom	Recom
22	dcterms.accessrights	4.67	4.73	Recom	Mand

No	Dublin Core	Easiness *Rating*	Usefulness *Rating*	Obligation from AP	Obligation from Users
23	dcterms.license	4.53	4.29	Recom	Recom
24	ags.rightsStatement	4.13	4.08	Recom	Optional
25	ags.termsOfUse	4.07	3.92	Recom	Optional
26	dcterms.relation	3.63	3.93	Optional	Optional
27	dcterms.conformsTo	4	3.38	Optional	Optional
28	dcterms.references	4.21	4.08	Optional	Optional
29	dcterms.isReferencedBy	4.36	3.64	Optional	Optional
30	ags.isTranslationOf	4.29	3.43	Optional	Optional
31	ags.hasTranslation	4.27	3.57	Optional	Optional
32	dcterms.subject	4.63	5	Recom	Mand
33	dcterms.description	4.4	4.64	Recom	Recom
34	dcterms.abstract	4.87	4.46	Optional	Recom
35	dcterms.blbiographicCitation	4.71	4.27	Recom	Mand

Table 6.4: Results from evaluation of metadata elements from subject matter experts for research repositories

Looking at the overall table, apart from the case of "meta-metadata" elements, all other elements were highly rated in both dimensions that they were evaluated. The case of the "meta-metadata" element is not expected to create any problems in the content population process as these elements are usually completed automatically. Table 6.4, presents and overview where it is evident that most elements were considered as easy to understand with 88.6% of the elements having an average rating of easiness coming from all the domain experts of 3 or more. From this percentage, 62.9% concerned ratings of 4 and above which is higher than the corresponding experiments in the other two cases examined (Learning and Cultural). 100% of the elements were also considered useful for the context of use. This last question showed a tendency of the domain experts to keep all the elements proposed within the application profile, showing that we would result with a big number of elements in the final application profile which can in turn lead to low completeness of the corresponding elements.

	Results				
Question	[0,1]	(1,2]	(2,3]	(3,4]	(4,5]
Is the element easy for you to understand?	0%	0%	11.4%	25.7%	62.9%
Is this element useful for describing educational resources?	0%	0%	0%	48.6%	51.4%

Table 6.5: Overall evaluation of the elements based on aggregated results

Adding to the fact that most elements were considered useful, the domain experts also agreed with the metadata experts that suggested the application profile on the obligation of the elements of the AP in 49% of the cases. Table 6.5 shows that there were lots of differences in the obligations indicated by the

designers of the metadata application profile in comparison to what the domain experts requested. Overall, domain experts asked for less recommended and more optional elements than the original application profile.

	Mandatory		Recommended		Optional	
Question	Before	After	Before	After	Before	After
Should this element be mandatory, recommended or optional?	13	12	14	6	8	17
Percentile change in overall number of mandatory / recommended or optional elements	-8%		-57%		+112%	

Table 6.6: Overall status of the elements before and after the evaluation process

6.4. Testing Phase

During this phase of the research collections' federation, a test implementation of the repositories was in place, giving to the content providers a platform through which they uploaded a small set of their resources in August 2011. Table 6.6 shows the main parameters of the peer-review experiment that was organized at the time.

Duration:	**1 week**
Date:	**August 2011**
Annotated Objects:	**25.000 objects (≈35% of total expected number)**
Resources Reviewed:	**65 (≈7 per expert)**
Involved people:	**9 domain experts**

Table 6.7: Peer-Review Experiment Parameters

After the upload of the initial set of resources, mainly through harvesting processes and less through manual insertion of metadata, a peer-review took place. 65 resources from all content providers were assigned to 9 reviewers to review and evaluate them using the Metadata Quality Assessment Grid.

Score	Completeness	Accuracy	Consistency	Objectiveness	Appropriateness	Correctness	Overall score
5	20%	31%	26%	60%	14%	65%	18%
4	11%	43%	42%	20%	17%	22%	31%
3	6%	9%	12%	5%	14%	6%	14%
2	20%	2%	0%	3%	22%	2%	18%
1	40%	6%	2%	3%	26%	2%	15%
no	3%	9%	18%	9%	8%	5%	3%

Table 6.8: Overview of the scores provided for all the quality metrics for 65 resources that were reviewed

As it is shown in Table 6.8, the results from this initial experiment were not encouraging. More specifically, the main outcomes were the following:

- Only 31% of the metadata records were rated to be completed at a satisfying level (4 or 5 out of 5).
- Almost 74% of the metadata records were considered to be accurate of the resource they described.
- 68% of the metadata records were considered to be in accordance with the VOA3R application profile.
- 80% of the metadata records were considered objective and un-biased as to the resource they describe.
- Only 31% of the metadata records were considered to be really appropriate for the envisaged use in the VOA3R portal.
- 87% of the metadata records contained wording and descriptions that were syntactically and grammatically correct.

Looking at some input that was gathered through the reviews that is not present in Table 6.8, we saw that 40% of the resources and corresponding metadata records were not considered of high quality enough to be published online. Based on the reviews conducted from subject-matter experts, a first version of a guide for annotators was created, namely the "Good & Bad Metadata Practices Guide" that provided good and bad cases of metadata records to help content providers to work with metadata in the future.

6.5. Calibration Phase

During the Calibration Phase of the institutional collections, 50.000 resources were populated to the project repositories until December 2011. In this Phase, a second peer review was organized where 61 reviews were provided for a selected sample of each content provider's resources with the help of 9 reviewers.

Duration:	**1 week**
Date:	**December 2011**
Annotated Objects:	**50.000 objects (≈70% of total expected number)**
Resources Reviewed:	**61(≈7 per expert)**
Involved people:	**9 domain experts**

Table 6.9: Peer-Review Experiment Parameters

The results related to the quality of the metadata records but also the resources themselves were more encouraging than the previous one but still not at the desired level. Table 6.10 presents the overall results, followed by the conclusions drawn.

Score	Completeness	Accuracy	Consistency	Objectiveness	Appropriateness	Correctness	Overall score
5	8%	51%	26%	67%	30%	74%	25%
4	28%	31%	46%	26%	48%	21%	44%
3	26%	10%	23%	3%	8%	0%	28%
2	31%	3%	5%	3%	11%	0%	3%
1	7%	2%	0%	0%	3%	2%	0%
no	0%	3%	0%	0%	0%	3%	0%

Table 6.10: Overview of the scores provided for reviewed resources

Looking at the overall results, the following conclusions can be drawn:

- 36% of the records were considered to be completed at a desired degree whereas the majority of them were not.
- 82% of the records represented accurate descriptions of the resource they depicted.
- 72% of the metadata records were consistent with the metadata application profile adopted, being rated with 3 out of 5 and more.
- 93% of the metadata records were objective related to the description of the resource they provided.
- 78% of the metadata records are appropriate for the envisaged use in the institutional repositories' federation.
- 95% of the metadata records are correct in terms of the grammatical and syntactical use of the languages in which they are provided.

Finally, related to the Yes/No question that is included in every questionnaire, about whether or not the domain experts consider the metadata record to be of the desired quality. 20% of the metadata records and their corresponding resources were not deemed to be fit for publishing online. Based on the input gathered both from the peer review exercise. a revised version of the metadata application profile was produced, based on which the population of the critical mass of resources will take place. Following the adaptation of the application profile, a new version of the "Good & Bad Metadata Practices Guide" was produced, to help content annotators in their tasks.

6.6. Critical Mass Phase

During the Critical Mass Phase, a usage data analysis on the metadata records of the content management tool (Confolio) took place, looking at a sample of 51.057 resources that were populated in the institutional repositories until April 2012. In Table 6.11, the main parameters of the experiment are presented.

Duration:	**1 week**
Date:	**April 2012**
Annotated Objects:	**51.057 objects (≈70% of total expected number)**
Resources Analyzed:	**51.057**
Involved people:	**2 metadata experts**

Table 6.11: Completeness Assessment Parameters

The findings of this analysis are presented in Table 6.12. The metadata element set that was examined is the one current in use in the institutional repositories. This is a subset of the application profile which was exposed to the domain experts of the project mainly because at the time of this study there was no consensus related to the final element set that would be used in the federation.

No	Element Name	Total Records	Occurrences	%	Obligation
1	dc.identifier	51.057	50.942	99.8%	Recom
2	dc.title	51.057	51.051	100%	Mand
3	dc.language	51.057	43.510	85.2%	Mand
4	dc.description	51.057	48.076	94.2%	Recom
5	dc.subject	51.057	46.414	90.9%	Recom
6	dc.coverage	51.057	21.090	41.3%	Optional
7	dc.type	51.057	49.854	97.6%	Mand
8	dc.date	51.057	42.141	82.5%	Mand
9	dc.creator	51.057	48.827	95.6%	Recom
10	dc.contributor	51.057	21.961	43.0%	Recom
11	dc.publisher	51.057	32.078	62.8%	Recom
12	dc.format	51.057	47.298	92.6%	Recom
13	dc.rights	51.057	24.389	47.8%	Recom
14	dc.relation	51.057	45.932	90.0%	Optional
15	dc.source	51.057	24.166	47.3%	Optional

Table 6.12: Completeness Analysis from Confolio, for the cultural collections.

The results of Table 6.12 were satisfying for the content providers, related to a core set of elements used in the federation, but still many elements remained underused or completely absent in most repositories.

To address the problems for these elements but also the new ones that were not yet part of the application profile, another focus group meeting for the metadata application profile took place. A new version of it was presented to the domain experts on May 2012, asking for their input. This is a major difference from the other two experiments (cultural collections and learning collections), with this exercise taking place for the second time instead of once. The need for the extra time that this exercise took place came mainly

from the fact that the elaboration of an application profile in the case of the institutional repositories' federation was more difficult. Content providers had their own, legacy application profiles that greatly varied from the final one, requiring a more complicated harmonization process to reach a unanimous decision on the metadata application profile.

Duration:	**2 hours**
Date:	**May 2012**
Annotated Objects:	**Not applicable**
Involved people:	**13 domain experts**

Table 6.13: Metadata Understanding Session Parameters

As it is shown in Table 6.14, 48 elements of the new version of the metadata AP were evaluated from domain experts who concluded that 8 elements should be mandatory in the new AP, 11 should be recommended and 29 optional. The initial obligation coming from the AP included 9 mandatory elements, 18 recommended and 19 optional ones. One more time, results from such experiments indicate that users prefer to have more optional elements than what the metadata experts originally intended for.

No	Dublin Core	Easiness *Rating*	Usefulness *Rating*	Obligation from AP	Obligation from Users
1	Title	4.92	4.85	Mand	Mand
2	Alternative Title	3.08	3.5	Optional	Optional
3	Creator	4.58	4.83	Recom	Mand
4	Contributor	3.25	4.08	Recom	Optional
5	Publisher	4.33	4.67	Recom	Recom
6	Date	3.92	4.54	Mand	Mand
7	Language	4.54	4.46	Mand	Mand
8	Identifier	3.38	4.31	Recom	Recom
9	Format	3.83	3.77	Recom	Recom
10	Is Shown By	3.08	3.67	Optional	Recom
11	Is Shown At	3.25	3.92	Optional	Recom
12	Subject	4.31	4.46	Recom	Recom
13	Description	3.62	3.62	Recom	Optional
14	Abstract	4.54	4.62	Optional	Recom
15	Bibliographic Citation	4	4.38	Recom	Mand
16	Type	3.77	4.38	Mand	Mand
17	Rights	3.31	3.77	Recom	Optional
18	Access Rights	3.54	4.15	Recom	Optional
19	License	3.25	3.83	Recom	Optional
20	Review Status	4.15	3.92	Recom	Optional
21	Publication Status	4.31	3.83	Recom	Recom
22	Relation	2.31	2.92	Optional	Optional
23	Conforms To	2.46	2.5	Optional	Optional
24	References	3.23	3.67	Optional	Optional

No	Dublin Core	Easiness *Rating*	Usefulness *Rating*	Obligation from AP	Obligation from Users
25	Is Referenced By	3.73	3.45	Optional	Optional
26	Has Part	2.42	2.73	Optional	Optional
27	Is Part Of	2.92	2.8	Optional	Optional
28	Has Version	2.75	3	Optional	Optional
29	Is Version Of	2.83	3	Optional	Optional
30	Has Translation	3.25	3.27	Optional	Optional
31	Is Translation Of	3.25	3.18	Optional	Optional
32	Has Meta-metadata	2.67	3.27	Optional	Optional
33	Object of Interest	3.75	3.64	Optional	Optional
34	Variable	2.58	3.18	Optional	Optional
35	Method	3.33	3.55	Optional	Optional
36	Protocol	3.08	3.55	Optional	Optional
37	Instrument	3.08	3.45	Optional	Optional
38	Technique	3.25	3.27	Optional	Optional
39	Identifier	3.33	3.82	Mand	Mand
40	Type	3.73	3.82	Mand	Recom
41	Language	3.83	3.82	Recom	Recom
42	Date	4.33	4.18	Mand	Mand
43	Contributor	3.42	3.55	Recom	Recom
44	Agent/Name	3.83	4.09	Mand	Optional
45	Organization/Name	4	4	Mand	Optional
46	Person/First Name	3.92	4	Recom	Optional
47	Person/Last Name	3.83	4.09	Recom	Optional
48	Person/Personal Mailbox	3.92	3.55	Recom	Optional

Table 6.14: Results from evaluation of metadata elements

Table 6.14, presents and\ overview where it is evident that most elements were considered as easy to understand with 83.4% of the elements having an average rating of easiness coming from all the domain experts of more than three out of five. 87.6% of the elements were also considered useful for the context of use. From these results we see that the attitude of the domain experts towards keeping in the AP as many elements as possible has significantly changed especially after working with the elements in the tools deployed.

	Results				
Question	[0,1]	(1,2]	(2,3]	(3,4]	(4,5]
Is the element easy for you to understand?	0%	0%	16.7%	64.6%	18.8%
Is this element useful for describing educational resources?	0%	0%	16.7%	56.3%	31.3%

Table 6.15: Overall evaluation of the elements based on aggregated results

Adding to the fact that most elements were considered useful, the domain experts also agreed with the metadata experts that suggested the application

profile on the obligation of the elements of the AP in 65% of the cases. Overall, this Metadata Understanding Session indicated the willingness of the domain experts to reduce the elements present in the AP and it also aligned the results with Metadata Understanding Sessions held in the other two cases, as in the first round with the domain experts, the results overwhelming in comparison to the other two experiments presented already.

6.7. Regular Operation Phase

During the Regular Operation Phase of the research federation, the content providers are still using, in most of the cases, the Dublin Core metadata standard, despite the fact that an application profile was developed following the recommendations from the MQACP. The application profile developed contained additional bibliographic elements that demanded considerable effort from the content providers at the time. This effort was deemed as necessary but on the other hand, it was not covered financially within the project framework that funded the federation of the VOA3R content in the first place. More specifically, in an analysis that was carried out (VOA3R, 2013), the content providers estimated the time that the enrichment of a metadata record requires. Their estimations varied significantly, from 5 minutes to 40 minutes, in cases of collections with really poor metadata. Overall, an average of 20 minutes per record was estimated to be needed which was not considered cost effective. At the time being, the content providers had contributed more than 75.000 records, which would results to a cost of 1.500.000 minutes, which is almost 1000 days, or 178 man months. The remaining records (almost 2.400.000) of the Cultural Federation that were harvested from external content providers were not considered in this analysis.

After deciding to stay with the existing metadata standard and enrich only these records, a usage data analysis took place, looking at a sample of 74.379 resources that were populated in the institutional repositories until June 2013. In Table 6.15, the main parameters of the experiment are presented.

Duration:	**1 week**
Date:	**June 2013**
Annotated Objects:	**74.379 objects (100% of total expected number)**
Resources Analyzed:	**74.379**
Involved people:	**2 metadata experts**

Table 6.16: Completeness Assessment Parameters

The findings of this analysis are presented in Table 6.17 and being compared with previous results.

No	Element Name	Obligation	% May 2012	% June 2013	No of Records
1	dc.identifier	Recom	99.77%	100%	74.379
2	dc.title	Mand	99.99%	100%	74.379
3	dc.language	Mand	85.22%	93.3%	69.395
4	dc.description	Recom	94.16%	96.8%	71.999
5	dc.subject	Recom	90.91%	100%	74.379
6	dc.coverage	Optional	41.31%	97.8%	72.743
7	dc.type	Mand	97.64%	91.8%	68.280
8	dc.date	Mand	82.54%	96.3%	71.627
9	dc.creator	Recom	95.63%	100%	74.379
10	dc.contributor	Recom	43.01%	51.8%	38.528
11	dc.publisher	Recom	62.83%	41.2%	30.644
12	dc.format	Recom	92.64%	81.7%	60.768
13	dc.rights	Recom	47.77%	71.3%	53.032
14	dc.relation	Optional	89.96%	75.7%	56.305
15	dc.source	Optional	47.33%	100%	74.379

Table 6.17: Completeness Analysis from Confolio, for the cultural collections.

The results of Table 6.17 show a significant increase in the usage of most elements with the exception of Relation that was almost 14% lower than in May 2012 and Publisher that dropped by almost 20%. On the other side, the use of the Coverage and Source elements increased by more than 50% whereas Date and Language showed smaller but significant improvements. Looking at the June 2013 measurements we see all but two elements being completed in more than 70% of the records which is more than satisfying in terms of metadata completeness.

6.8. Cost Implications

In Table 6.18, the cost analysis of the application of MQACP on the case of the Research Federations is presented. In the table, the timing of each experiment is presented along with the method used to conduct and the estimated time that this occupied the experts involved.

Date/Period	No of Experts	Method	Time per item	No of items	Total time
6/2011	16 domain & 1 metadata experts	Focus group meeting for metadata design	2 hours	17 people	34 hours
7/2011-8/2011	1 metadata expert	Prepare process	5 minutes per record	65 records	5.4 hours
		Collect results	3 minutes per form	65 forms	3.3 hours
		Analyze & interpret	N/A	N/A	2 hours

Date/Period	No of Experts	Method	Time per item	No of items	Total time
	9 domain experts	Peer review conducted online	20 minutes per record	65 records	21.7 hours
	1 metadata expert	Good & Bad Metadata Practices	8 hours	1 guide	8 hours
12/2011	1 metadata expert	Prepare process	5 minutes per record	61 records	5 hours
		Collect results	3 minutes per form	61 forms	3.1 hours
		Analyze & interpret	N/A	N/A	2 hours
	9 domain experts	Peer review conducted online	20 minutes per record	61 records	20.3 hours
4/2012	1 technical expert	Modify DB script to store results in excel	N/A	51.057 records	5 hours
		Export data from repositories	10 minutes per collection		1,5 hours
	1 metadata expert	Metadata completeness measurement	N/A		3 hours
5/2012	17 domain & 1 metadata experts	Focus group meeting for metadata design	2 hours	18 people	36 hours
6/2013	1 technical expert	Export data from repositories	10 minutes per collection	74.379 records	1,5 hours
	1 metadata expert	Metadata completeness measurement	N/A		3 hours
TOTAL	58 people			**TOTAL**	154.8 hours

Table 6.18: Cost of the MQACP application on the Research Federation

Looking at the cost in the case of the Research Federation, we see that the total cost amounted to 154.8 hours which was less than the Cultural Federation but more than the Learning Federation. We see that the total time vested, even in a Federation with significantly more resources, did not exceed the other cases. Looking at the time per person, we saw that it was at 2.7 hours per person, really close with the 2.8 of the Learning Federation which shows that the cost is comparable in all the cases. Once more, the difference in the total cost was attributed to the smaller number of resources that were peer-reviewed in comparison to the Cultural Federation case.

6.9. Conclusions

In the following table, an overview of the methods utilized for applying MQACP in the Research Federation is presented. The methods used were six (6) in total, more than the Learning Federation (5) and less than the Cultural Federation (8).

Experiment	No of participants / records	Expertise	Date	Results	Tool
Metadata Understanding Session	16	Domain	6/2011	Quantitative & Qualitative	Questionnaire with Likert scale and multiple choice
Metadata Record review	9 / 65 (records)	Domain	8/2011	Quantitative & Qualitative	Questionnaire with Likert scale and free text comments
Metadata Record review	9 / 61 (records)	Domain	12/2011	Quantitative & Qualitative	Questionnaire with Likert scale and free text comments
Log files analysis	2/ 51.057 (records)	N/A	4/2012	Quantitative	Microsoft Excel
Metadata Understanding Session	13	Domain	5/2012	Quantitative & Qualitative	Questionnaire with Likert scale and multiple choice
Log files analysis	2 / 74.379 (records)	N/A	6/2013	Quantitative	Microsoft Excel

Table 6.19: Details of the Metadata Quality Assessment Certification Process Methods

Looking at the results yielded in the case of the research repositories, it is evident that the use of the experiments of Table 6.19 greatly improved the quality of the metadata records in the repositories, both in terms of completeness but also in terms of the other metrics that were assessed. Although that the number of metadata elements used in the case of the research repositories is not comparable to the other cases, being way fewer, still for these fifteen elements completeness was really high for all elements, regardless their obligation, averaging 86,5%.

Focusing more on the other quality metrics used within this work, we present in Table 6.20, the aggregated results from the peer-review experiments that were carried out during the entire 24-month period that the MQACP was applied on the Research Federation. Table 6.20 shows an overall improvement on the quality dimensions examined for the sample of metadata records between the two peer-review experiments. Percentages in Table 6.20 show the amount of resources that were reviewed with 4 or 5 out of 5 in the respective metrics and in the case of the column “Publish”, the ones that were deemed as fit for publishing. Appropriateness showed the biggest

improvement whereas Completeness, Accuracy, Consistency, Objectiveness and Correctness showed small positive deviations between the two measurements. Finally, the recommendation of publishing the material online, went up by 20% showing that metadata records overall were significantly improved in the meantime between the two experiments.

Metric / Value	**Complete** (4 or 5)	**Accurate** (4 or 5)	**Consistent** (4 or 5)	**Objective** (4 or 5)	**Appropriate** (4 or 5)	**Correct** (4 or 5)	**Overall Score**
March 2012	31%	74%	68%	80%	31%	86%	60%
May 2013	36%	82%	72%	88%	72%	89%	80%

Table 6.20: Aggregated results from Quality Peer Review experiment.

In total, the MQACP for the case of the research repositories was deployed over a period of 24 months, when a total of six (6) experiments were carried out, reaching an average metadata completeness of 86.5% per element with an improvement in metadata completeness between the two measurements of an average of 8.4%.

7. Conclusions & Future Work

In the next section, the main conclusions of this thesis are presented, pointing out the outcomes of the experiments carried out. Adding to that, a set of possible research extensions is described, reflecting ongoing research but also future work that has come out of this thesis.

7.1. Cross-Case Study Discussion

This section will include a short discussion on the similarities and differences between the common experiments carried out in all three case studies to try and identify meaningful patterns related to metadata quality.

7.1.1. Metadata Design Phase

To begin with, we look at the requirements coming from the side of the domain experts that provide the content. Before looking at the results in Table 7.1, one would expect that the domain experts would ask for less mandatory elements and more optional, to avoid cumbersome annotation tasks. In fact this was not true in most cases as in the Learning case and the Cultural one, the experts asked for more mandatory elements in the application profile than the ones that were proposed to them during the AP presentation. In the Research case, more or less, the mandatory elements proposed and requested were the same in numbers. A similarity in all three cases was that less recommended elements were asked from the experts whereas in only one case, the cultural one, less optional elements were requested. The other two cases, Learning and Research one, requested more optional elements as it was anticipated.

	Mandatory		**Recommended**		**Optional**	
	Before	After	Before	After	Before	After
Learning Federations *(Organic.Edunet)*	6	11	26	17	24	29
	+83%		-34%		+21%	
Cultural Federations *(Natural Europe)*	13	21	6	2	10	6
	+62%		-66%		-40%	
Research Federations *(VOA3R)*	13	12	14	6	8	17
	-8%		-57%		+112%	

Table 7.1: Overall status of the elements before and after the Metadata Understanding Session

Overall, we can say with certainty that no specific pattern was distinguished in the experiments during the metadata understanding session in regards to the obligation of the elements. This clearly contradicts the notion that the metadata annotators would be asking for less mandatory elements to avoid a demanding metadata annotation process.

Looking at another aspect of the same experiment, we focus on the easiness and usefulness evaluations of all element by the domain experts. In the following table, we group the results of the evaluations on the easiness and usefulness of the elements to see how many were deemed as really easy, or really useful each time and overall.

Case	Element is easy to understand		Element is useful for the targeted use	
	[0 to 3.5)	[3.5 to 5]	[0 to 3.5)	[3.5 to 5]
Learning	24% (11)	76% (34)	49% (22)	51% (23)
Cultural	28% (8)	72% (21)	21% (6)	79% (23)
Research	20% (7)	80% (28)	12% (3)	88% (32)

Table 7.2: Easiness and usefulness of all elements proposed in all case studies

Looking at Table 7.2 we see that in all cases, almost 75% of the elements are deemed as relatively easy to understand by the domain experts, which shows that in theory, metadata is not a hard topic for the content providers, with most of them really understanding their meaning and purpose. When it comes to their usefulness though, things change significantly. In the case of the learning federations, half of the elements were deemed as relatively useful for the targeted use, with average scores of 3.5 or more, whereas the other half scored less than 3.5. This indicated a disagreement among domain experts and the metadata experts that built the application profile as to the elements that should be used for the description of the digital resources. In the other two cases though, the domain experts agreed that most of the elements were useful for the targeted use with percentages of 80% or more. Overall, looking also at the results in comparison to the number of elements contained in the application profile evaluated, we see that for smaller application profiles as the ones in the cultural (29 elements) and the research cases (35 elements) the ratings of the domain experts are closer whereas when we have bigger application profiles like in the learning case (45 elements), problems with the easiness or the usefulness of the elements arise.

Finally, for the experiments that took place during the Metadata Design Phase, it would be useful to see how elements were rated in terms of easiness or usefulness depending on their obligation. In the following table, their easiness and usefulness is presented for the elements grouped based on their obligation from the application profile (metadata designers) or the domain experts' feedback. In Table 7.3, we see that in all cases, elements that were deemed as mandatory from the domain experts were thought to be easier to understand than the mandatory ones chosen from the metadata experts that design the application profile. For the optional elements, it seems that in most cases, the optional elements that are chosed by the domain experts are rated lower in terms of easiness to understand them, than the ones selected by the metadata designers. This means that in general, the elements that the domain

experts do not understand well, they usually prefer them to be optional. For the recommended element, the same finding is confirmed in the case of the cultural federation, whereas in the case of the other two federations, the perceived easiness for the elements selected by the experts and the ones selected by the domain experts themselves, is more or less the same.

Easiness of elements	**Mandatory**	**Recommended**	**Optional**
Learning (AP)	4	3.7	3.53
Learning (Domain experts)	4.15	3.72	3.47
Cultural (AP)	4.18	3.97	3.57
Cultural (Domain experts)	4.25	2.95	3.12
Research (AP)	3.72	4.3	4.24
Research (Domain Experts)	4.35	4.39	3.76
Usefulness of elements	**Mandatory**	**Recommended**	**Optional**
Learning (AP)	4.08	3.39	3.23
Learning (Domain experts)	3.97	3.45	3.18
Cultural (AP)	4.34	4.12	3.91
Cultural (Domain experts)	4.45	3.19	3.37
Research (AP)	4.15	4.46	3.81
Research (Domain Experts)	4.55	4.55	3.81

Table 7.3: Easiness and usefulness of all elements based on their obligation

Looking at the perceived usefulness, in all cases, the perceived usefulness for the mandatory elements is the approximately the same or higher for the ones chosen by the domain experts. Similarly, in all cases, the optional elements chosen by the domain experts were the ones with an equal or lower usefulness. The same phenomenon is observed for recommended elements. Looking at the ratings of the elements based on the selected obligation coming from metadata designers, we see that in all cases, as we go from mandatory to optional, both easiness and usefulness drop. Only in the case of research repositories, domain experts seem to understand the optional elements better than they do the mandatory ones. This phenomenon partially explains actually the need for a new metadata understanding session that was organized in a latter phase of the project. Also, it was "corrected" through the selection of the domain experts that chose the easier elements for them, as mandatory and recommended, choosing as optional the ones they could not understand or did not think useful.

7.1.2. Calibration Phase

During the Calibration phase, in all three cases, peer-reviews of samples of records from each repository were carried out. The aim of the short analysis that follows is to examine whether or not we can find any significant similarities or systematic differences that could serve as theory generators for peer-reviewing of metadata records.

Based on the metrics deployed in the Metadata Quality Grid, all records were evaluated in terms of their completeness, accuracy, consistenct, objectiveness, appropriateness and correctness. In addition, all domain experts provided an overall score for each record.

Score	Completeness	Accuracy	Consistency	Objectiveness	Appropriateness	Correctness	Overall score
5	40%	51%	50%	69%	41%	69%	40%
4	45%	32%	28%	21%	33%	21%	37%
3	5%	10%	15%	6%	18%	9%	19%
2	9%	3%	1%	2%	5%	0%	0%
1	1%	1%	0%	0%	2%	1%	1%
no	1%	3%	6%	3%	1%	1%	3%
Score	Completeness	Accuracy	Consistency	Objectiveness	Appropriateness	Correctness	Overall score
5	16%	22%	36%	73%	32%	62%	26%
4	42%	43%	40%	13%	29%	22%	38%
3	19%	13%	19%	4%	13%	13%	30%
2	6%	1%	3%	6%	14%	1%	3%
1	13%	9%	1%	4%	11%	1%	2%
no	3%	11%	0%	0%	0%	1%	0%
Score	Completeness	Accuracy	Consistency	Objectiveness	Appropriateness	Correctness	Overall score
5	8%	51%	26%	67%	30%	74%	25%
4	28%	31%	46%	26%	48%	21%	44%
3	26%	10%	23%	3%	8%	0%	28%
2	31%	3%	5%	3%	11%	0%	3%
1	7%	2%	0%	0%	3%	2%	0%
no	0%	3%	0%	0%	0%	3%	0%

Table 7.4: Overview of the scores provided for reviewed resources in all cases (from top to bottom, Learning, Cultural & Research)

Looking at the results above, we see that in the Learning case, the metadata quality of all reviewed resources in most of the metrics is really high. The other two cases, present quite some similarities. In both the cultural and research federations, the overall score of the reviews for the resources are quite similar with 25% & 26% of the resources having a score of 5, 38% and 44% of them having a score of 4 and 30% and 28% having a score of 3. Looking at the other metrics as well, we find similarities between the two cases as well. Because of the nature of the experiments, it is not easy to deduct more generic conclusions from this table, but it might be interesting to look at how the metrics themselves are used in all three cases:

- **Completeness** does not present any pattern in all three cases. The only observation that can be made regarding its use, is that most resources are reviewed with 4 out of 5 for this metric,
- **Accuracy** in all cases is rated quite high, with 65-80% of the records being reviewed with 4 or 5 out of 5. The most common value is 4 out of 5, showing that accuracy is usually rated high,

- **Consistency** presents the same behavior as accuracy. In most cases the reviewers provide a score of 4 out of 5 or more. It seems though that more frequently than accuracy, consistency also takes values around 3 out of 5,
- **Objectiveness** takes always really high values with 5 out of 5 being the most prevalent and frequent value. It is the highest rated metric from the reviewers which indicates the tendency of the metadata annotators to be really objective with the metadata they provide,
- **Appropriateness** is really similar to accuracy and consistency with most resources taking a value of 4 out of 5,
- **Correctness** is similar to objectiveness in terms of the values reviewers have provided. It is also a metric that refers to the language used in the metadata and it usually reflects the correctness of the English language used in the records. This is expected to be high as all annotators possess a good knowledge of the English language.

Finally, if we were looking to find a metric that could act as a predictor of the overall quality of a metadata record as this is expressed by a domain expert, this metric would be either consistency or appropriateness.

7.1.3. Regular Operation Phase

In all three cases, in the beginning of the Critical Mass phase, a measurement of the completeness of all metadata records in the respective repositories took place. The comparison of these three cases could not lead to any useful insight, as the repositories were not fully aligned, with heavy annotation tasks taking place. For that reason, we decided to look at the metadata record completeness at a latter stage where the repositories are more stable, like in the Regular Operation phase. In the following tables, the elements of each application profile are presented, along with their perceived usefulness and easiness. The aim of this examination is to see at which degree this initial assessment of the elements can "predict" the metadata annotation process in the long-term. And also, if after applying all the metadata quality assurance methods that were described in each case, the final completeness of the elements was high, in spite of their original evaluation which might have been problematic.

7.1.3.1. Learning Federation

In Table 7.5, the final completeness of all elements in the Learning federation is presented. In addition, the evaluation scores of the metadata understanding session are provided, to check whether or not these evaluations can serve as an indication of the expected completeness of each element.

No	Learning Federation	%	Easiness	Usefulness	Obligation
1	General / Title	100%	4,13	4,25	Mand
2	.../ Language	99.9%	4,13	4,25	Mand
3	.../ Description	100%	4,2	4,06	Mand
4	Rights / Cost	82.4%	4	4,08	Mand
5	.../ Cost Copyright & Other Restrictions	99.9%	3,86	3,92	Mand
6	General / Keyword	99.9%	4,18	3,44	Recom
7	LifeCycle / Contribute / Role	77.2%	3,53	3,56	Recom
8	Educational / Intended End User Role	82.4%	3,56	3,33	Recom
9	.../ Context	81%	3,75	3,47	Recom
10	.../ Typical Age Range	63.9%	3,75	3,13	Recom
11	Rights / Description	92.4%	3,64	3,54	Recom
12	Classification	73.6%	3,1	3,4	Recom
13	General / Coverage	82.6%	3,6	3,5	Opt
14	.../ Structure	82.5%	2,82	2,93	Opt
15	LifeCycle / Status	39.7%	3,69	3,57	Opt
16	Educational / Interactivity Type	36.9%	3,2	3,5	Opt
17	.../ Interactivity Level	37.1%	2,73	2,64	Opt
18	.../ Semantic Density	37%	2,33	2,67	Opt
19	.../ Difficulty	37.1%	3,33	3,25	Opt
20	.../ Typical Learning Time	0.4%	3,27	2,67	Opt
21	.../ Language	52.3%	3,43	3,25	Opt
22	.../ Description	14.7%	3,64	3,54	Opt

Table 7.5: Completeness of elements in the Learning federations compared with perceived easiness & usefulness

Overall, we see that for most mandatory elements as well as the recommended ones, the easiness to understand and usefulness for the envisaged use, goes hand in hand with the respective percentages of completeness. This means that elements that the content providers understand or think that are useful, they usually provide them when annotating the resources. As it can be see in Figure 7.1, this is not the case for all the elements. In this table, the ratings of easiness and usefulness, were expressed as a percentage, to compare them with the completeness of the respective elements. The blue (dotted) line is the completeness, the red (dashed) is the easiness and the green (solid) is the usefulness. In general, the three lines show the same tendency thoughout the diagram, apart from some optional elements, depicted on the right. More specifically, these elements were not completed as much as expected, although their perceived easiness and usefulness was not that low.

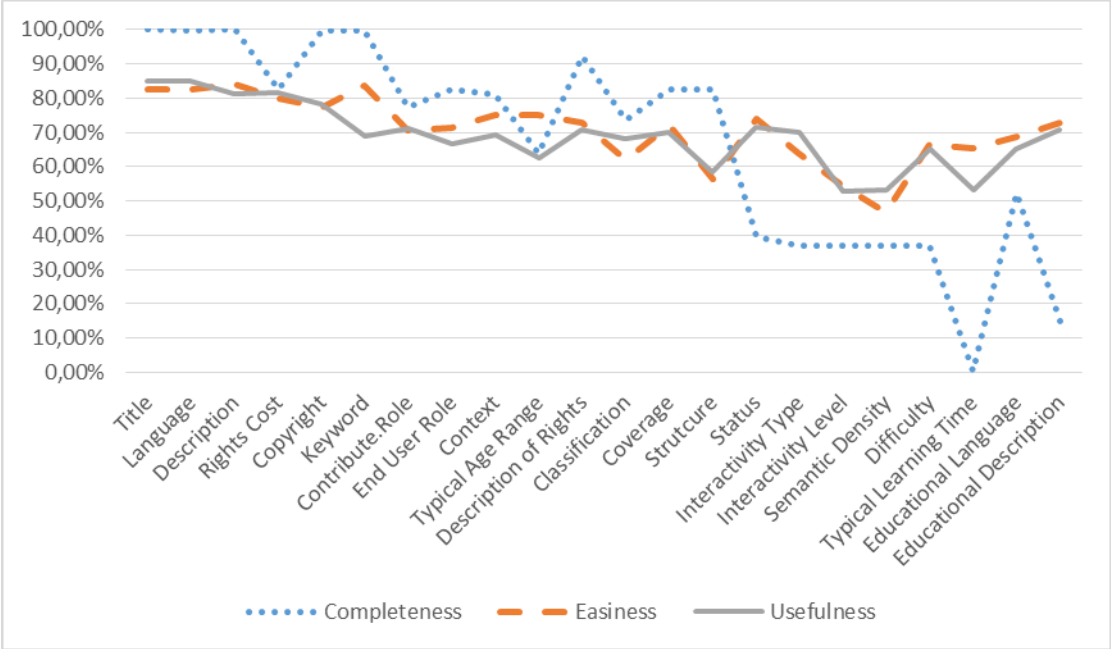


Figure 7.1: Completeness VS Easiness VS Usefulness for Learning Federations

7.1.3.2. Cultural Federation

In Table 7.6 we can see the completeness of all the elements in the Regular Operation phase of the Cultural Federation along with the perceived easiness and usefulness of each element as this was decided during the Metadata Understanding Session organized early in the lifecycle of the federation.

No	Cultural Federation	%	Easiness	Usefulness	Obligation
1	CHO Titles	99.5%	4,82	4,64	Mand
2	CHO Keywords	99.1%	3,55	4,7	Mand
3	Object Titles	100.00%	4,82	4,64	Mand
4	Object Descriptions	98.9%	5	4,91	Mand
5	Object URL	99.9%	4,45	4,8	Mand
6	Object Thumbnail URL	97.2%	3,55	4,3	Mand
7	Object Content Type	100%	4,27	4,64	Mand
8	Copyrights	100%	4,64	4,67	Mand
9	Access	100%	4,45	4,8	Mand
10	CHO Types	74.9%	4,27	4,64	Recom
11	Object Creators	87.3%	4,36	4,55	Recom
12	Object Languages	78.5%	5	4,91	Recom
13	Scientific Name	65.1%	3,55	4,7	Opt
14	Classification	56.7%	3,55	4,7	Opt
15	Common Names	49.8%	3,55	4,7	Opt
16	CHO Significant Dates	48.1%	3,36	3,4	Opt
17	CHO Temporal Coverage	18.5%	2,2	3,25	Opt

No	Cultural Federation	%	Easiness	Usefulness	Obligation
18	CHO Spatial Coverage	71.4%	2,2	3,25	Opt
19	CHO Mediums	17.9%	4,2	5	Opt
20	CHO Creators	58.1%	4,36	4,55	Opt
21	Object Creation Dates	66.5%	4	4,1	Opt
22	Object Identifiers	68.4%	4,45	4,8	Opt
23	Object Context URL	42.1%	3,2	3,63	Opt
24	Related Objects	47.1%	2,91	3,33	Opt
25	Object Formats	94.3%	4,45	3,8	Opt
26	Object Extents	89.1%	2,91	3,33	Opt

Table 7.6: Completeness of elements in the Cultural federations compared with perceived easiness & usefulness

Overall, the same observation that was made in the case of the Learning Federation can be made for this case as well. For most mandatory and recommended elements, high completeness goes along with easiness to understand and usefulness. If we generate a similar graph to the one of Figure 7.1, we will see that this is not the case for optional elements. More specifically, as it can be seen in Figure 7.2, for most of the elements, the completeness, easiness and usefulness lines, follow the same pattern. Only in a handful of cases this is not confirmed. More specifically, in the case of keywords, although it was not easy for the domain experts to understand their use, still the completeness rate was pretty high.

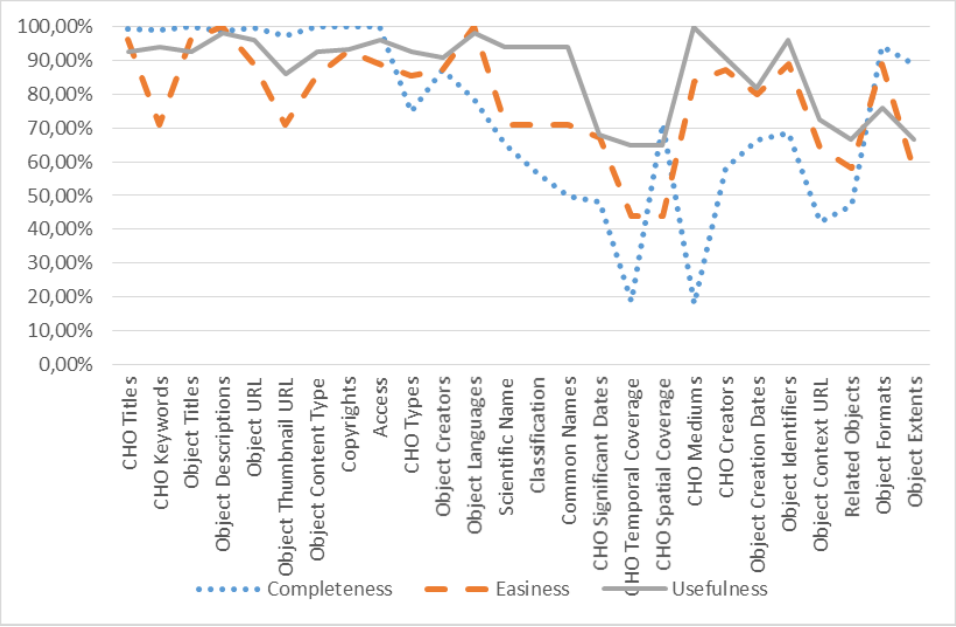


Figure 7.2: Completeness VS Easiness VS Usefulness for Cultural Federations

Similarly to the previous case, the most problems appeared for optional elements, were the actual completeness of the elements did not match the high level of understanding or usefulness that the domain experts indicated for the elements at hand. This showed us that despite the fact that some elements are considered important for the domain experts, they still do not provide them during metadata annotation.

7.1.3.3. Research Federation

In Table 7.7 the elements that were used in the Research Federations are presented along with their completeness as this was measured during the Regular Operation phase of the respective repositories. In the case of the Research repositories, a larger set of metadata elements was supposed to be annotated but in the end, the content providers kept the basic Dublin Core metadata elements as a feasibility study for the enrichment of the metadata records, indicated that this task was out of the budget limits of the initiative.

No	Element Name	Obligation	Completeness	Easiness	Usefulness
1	dc.title	Mand	100,00%	4,8	5
2	dc.language	Mand	93,30%	4,93	5
3	dc.type	Mand	91,80%	4,53	4,8
4	dc.date	Mand	96,30%	4,47	4,93
5	dc.identifier	Recom	100,00%	4,13	4,8
6	dc.description	Recom	96,80%	4,4	4,64
7	dc.subject	Recom	100,00%	4,63	5
8	dc.creator	Recom	100,00%	4	4,83
9	dc.contributor	Recom	51,80%	3,93	3,93
10	dc.publisher	Recom	41,20%	4,93	4,73
11	dc.format	Recom	81,70%	4,33	4,57
12	dc.rights	Recom	71,30%	3,87	4,33
13	dc.coverage	Optional	97,80%	4,4	3,93
14	dc.relation	Optional	75,70%	3,63	3,93
15	dc.source	Optional	100,00%	3,33	3,69

Table 7.7: Completeness of elements in the Research federations compared with perceived easiness & usefulness

In this case, we see that all the elements that were finally used, we populated in high percentages, apart from two or three of them. In general, their perceived easiness and usefulness was also high, more or less expected as this was not an application profile created from scratch, but an existing standard with which most domain experts were familiar and recognized its usefulness. In Figure 7.3, we see that completeness (dotted line), easiness (dashed line) and usefulness (solid line) were hand in hand for most of the elements. The two cases where this was not confirmed were the “Publisher” and

"Contributor" elements where the ratings were higher than the actual completeness, showing that even if the content providers thought they were needed, still they did not spend the time to complete them during the annotation process.

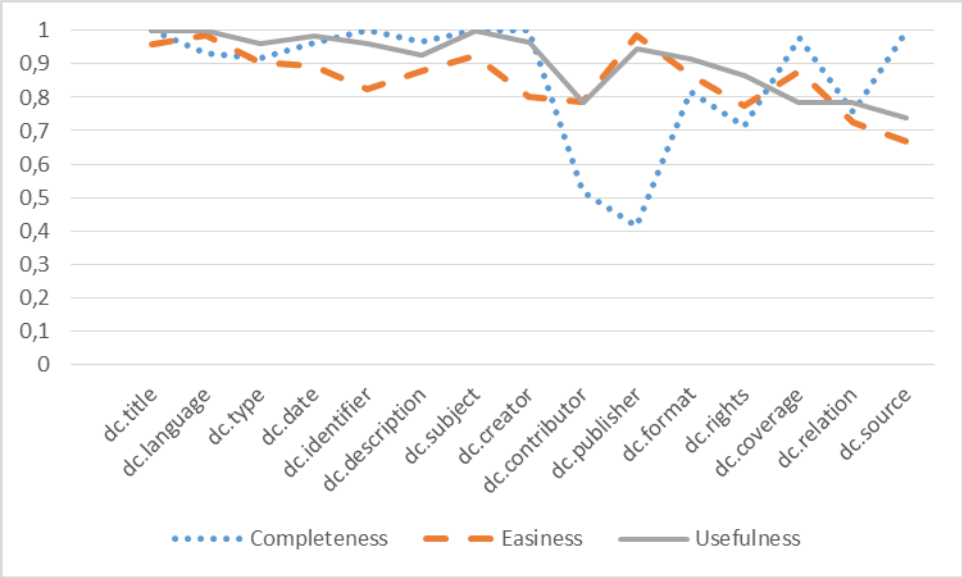


Figure 7.3: Completeness VS Easiness VS Usefulness for Research Federations

Looking at all three cases together, it seems that for mandatory and recommended elements, the perceived easiness and usefulness of the elements can act as a predictor of a high completeness of the elements (above 70%) whereas this is not true for optional elements, as we observed cases where easiness and usefulness were high, but the completeness was lower than anticipated.

7.2. Research Questions Addressed

In the introduction of this thesis, eight (8) research questions were discussed that the thesis aimed at addressing. In the following paragraph, all these questions are shortly discussed in the context of the thesis outcomes.

1. *Can we set up quality assurance methods for ensuring metadata record quality that will have a positive effect on the resulting quality of the metadata records of a given federation of repositories?*

In our research we set up a Metadata Quality Assurance Certification Process that was successfully applied to more than one federation of repositories of different contexts, proving a positive effect on the metadata quality of the resulting records.

2. *Which are the metrics that can be used to effectively assess the metadata record quality?*

In the context of MQACP, we used a set of metadata quality metrics that were retrieved from relevant literature. We used completeness, accuracy, objectiveness, correctness, consistency and appropriateness, as we consider those as representative in terms of describing metadata quality whereas at the same time enough to keep the complexity of the approach low.

3. *At what levels of the metadata record quality metrics is a repository considered to have a satisfying metadata record quality?*

The answer to this question is not an easy one. Going back to the quality definition we see that metadata quality is termed as the *totality of features and characteristics of metadata that bears on its ability to satisfy stated or implicit needs*. Drawing from this, we will have to take into consideration all the metadata quality metrics that we used within MQACP and we will have to see if the quality produced, satisfies the needs of the communities of users of the three federations examined. Looking at Accuracy, Consistency, Objectiveness, Appropriateness and Correctness we see that reviews that are above 4 out of 5, are considered as satisfying in the peer-review experiments that were carried out in an advanced phase of each repository lifecycle. We can say with some certainty, that when all these metrics are rated with 4 or more in more than 70% of the metadata records, then the overall quality of the records of each federation are in a satisfying degree of quality. Looking on the other hand to completeness, we see that in all three cases, average completeness of all metadata elements, no matter their quantity, is above 67% in all cases. Therefore, a threshold of 70% completeness in average is also considered to be satisfactory. In the case of mandatory elements, this number rises to more than 95% in all cases, so again it would be safe to expect this kind of completeness to consider the completeness metric satisfactory. In general, although the results of our research provide an indication for the expected metadata quality, they do not seem to be enough for a conclusive answer to this question.

4. *Can we introduce a quality assurance process that is transferable through different application contexts and types of repositories?*

Through the application of the MQACP in three (3) different cases of federations of repositories, we proved that the proposed process can be easily transferred to other domains with limited adaptations and more importantly with the same or comparable requirements in effort as this was measured in hours.

5. *What are the specific adjustments that have to be made to apply a quality assurance process in other contexts?*

Drawing from the experience coming from the application of MQACP in three different cases, we saw that no major adjustments were needed in order to apply it in other contexts. Apart from adapting the metadata understanding session depending on the metadata application profile used and other similar adaptations in the forms used, the MQACP can be used as is in different contexts. One significant adaptation that was carried out was the number of times each separate quality assurance method was deployed. For example, in the case of the research federations, more than one metadata understanding sessions were needed, as the content providers had a really hard time aligning the application profiles they used. In addition, in the cultural case, we selected to carry out more peer-reviews of the metadata from experts, to ensure the quality of metadata as the intermediate completeness measurements were not satisfying. Overall, whereas the MQACP itself was adapted seamlessly, some of its methods needed to be carried out more or less times based either on specificities of each case or the results of the intermediate metadata completeness measurements that took place.

6. *Are the results of the application of the same metadata quality assurance process in different repositories comparable in terms of the resulting metadata quality?*

To answer this question we will look at the results from the completeness measurement that were carried out in the Regular Operation Phase of all three federations. In the case of the Learning Federation, the average completeness for all elements of the application profile (22) regardless of their obligation (mandatory, recommended or optional) was at 67% with a median of 79.1% and a range of 99.6%. Looking closely, we see that 14 out of 22 elements were completed in more than 60% of the cases, showing that this mediocre average was affected by some really low completed elements. This is confirmed from the big range between the lowest and highest completeness. In the case of the 26 elements of the cultural case, the average completeness was at 74.1% with a median of 76.7%, and a range of 82.07%. Again, 18 of the 26 elements were completed in more than 60% of the cases, showing that the high percentage of completeness found could be even higher if one or two elements were not as neglected from the users. This is again confirmed by the high range observed. In the research federations the average completeness was 85.2% with a median of 96.3% and a range of 58.8% showing higher quality in terms of completeness which is attributed to the small number of elements that were finally deployed in the federation. Overall, the completeness results were comparable in all three cases for the mandatory elements in each application profile, as well as the recommended ones. The big difference in completeness

that is observed in the research federation comes mainly from the optional elements that were completed in many more cases by the content providers.

In addition, looking at the three cases through figures, we see that in all Figures 7.1, 7.2 and 7.3 when elements are sorted based on their completeness and plotted; the situation in the Learning and Cultural Federations is pretty much alike. Most of the elements (63-69%) are completed more than 60% of the times whereas only 7-10% of them are completed in less than 40% of the times but with really low rates. Finaly, almost 20% of the elements are completed 40% to 60% of the times. In the Research Federation, 86% of the elements are completed more than 60% of the times, whereas the remaining 14% are completed 40% to 60% of the times.

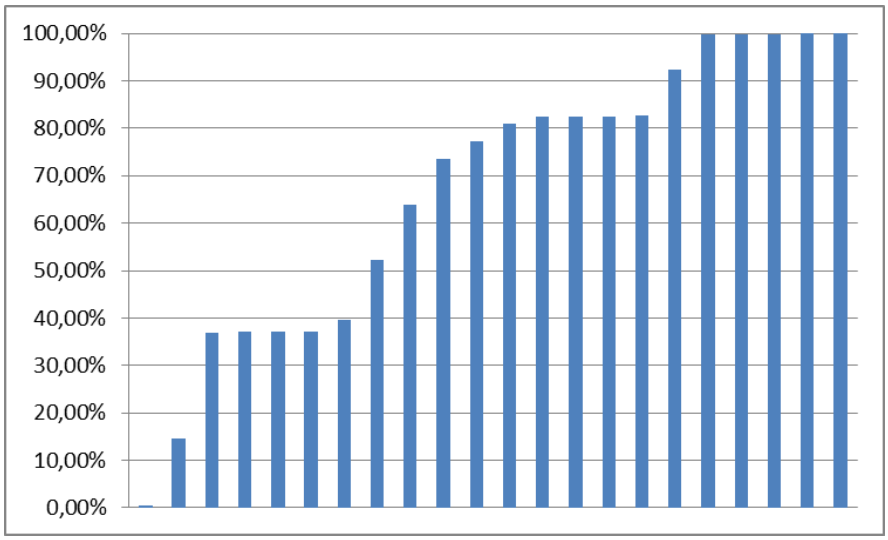


Figure 7.4: Completeness of all elements in learning federation (ascending order)

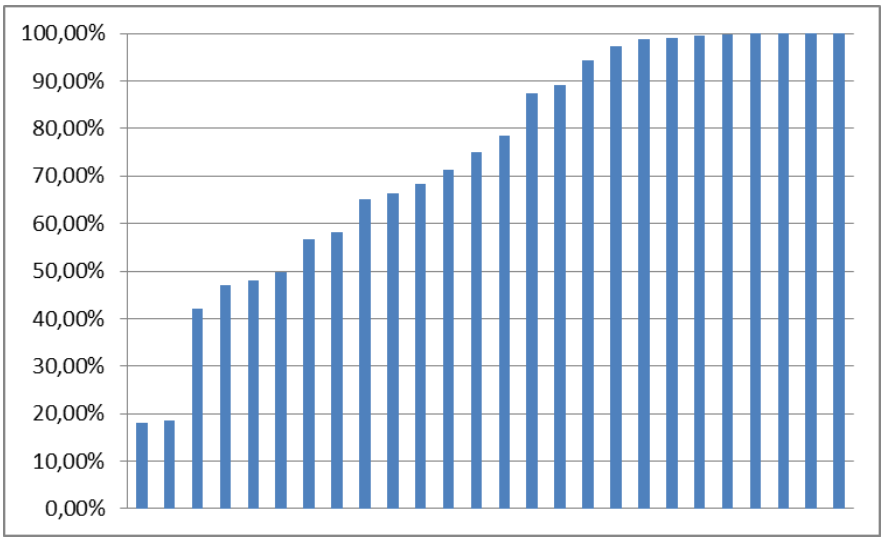


Figure 7.5: Completeness of all elements in cultural federation (ascending order)

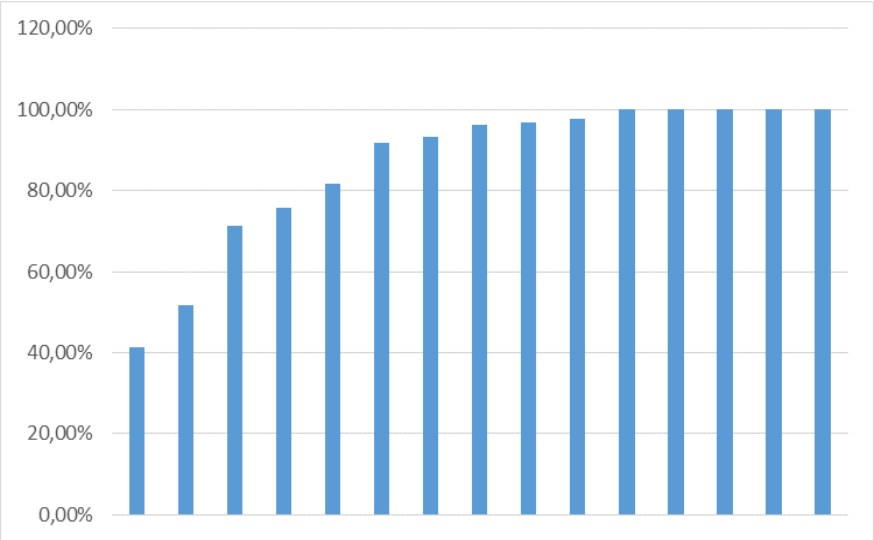


Figure 7.6: Completeness of all elements in research federation (ascending order)

7. *Is the cost involved in the application of a metadata quality assurance process comparable in terms of magnitude for different repositories?*

To answer this question, we look at the comparison of the time that was vested for the application of MQACP in each case. Table 7.1, indicates the number of quality assurance methods deployed in each case, as well as the people involved and the time vested in total. We need to clarify that the number of people involved is not absolute, but it includes the same people that may have participated in more than one methods. So, if a person participated three times in three peer reviews of metadata records, he/she is counted as three people. Overall, all the experiments were supported by a total of around twenty (20) domain experts, two (2) metadata experts and one (1) technical expert.

Case	People	Methods used	Hours
Learning	49	5	134.7
Cultural	45	8	166.5
Research	58	6	154.8

Table 7.8: Cost parameters for all experiments.

The methods used are also comparable, as in an average, around 6 or 7 of them are enough to achieve comparable results in terms of completeness for all metadata records as it was shown before. Finally, the overall time needed is again comparable, as no significant deviations were noted in each experiment. An average of 150 hours is enough to apply MQACP in each case.

8. *Is the improvement in metadata record quality comparable with the cost of the quality assurance method?*

To answer this question we must look at the overall cost of each experiment in hours, and compare it with the metadata completeness it yielded. The following table shows the metadata completeness in each federation as these were discussed before, along with the cost of each case, in hours. Finally, the difference of the average metadata completeness between the beginning of the repository lifecycle and its regular operation is presented.

Case	Average Completeness	Improvement between two measurements	Elements in AP	Hours	% of Improvement per hour spent
Learning	67%	+48.7%	22	134.7	0.36%
Cultural	74.1%	+16%	26	166.5	0.09%
Research	85.2%	+7.2%	15	154.8	0.05%

Table 7.9: Metadata completeness and cost

Looking at the improvement in completeness between the Critical Mass phase and the regular operation in relation to the cost, we see that overall, the costs vested did not yield the same improvement every time. More specifically, the lowest overall cost (learning case) brought upon the highest improvement in the resulting metadata completeness. In this case, we find that the starting point of the metadata that exist in each federation, is of great importance. In the learning federation, during the early stages of the Critical Mass phase, the average completeness was as low as 18.3% which provided room for improvement, greater than in the case of the research federations that started with an average completeness of 78%. To address this question, we see that the improvement in metadata records completeness (and not quality) is not comparable to the cost of applying MQACP. To examine the quality in general, was not possible, as many more parameters have to be examined, such as the quality of the services offered on top of these federations. What would be safe to claim though, is that the higher the starting completeness of the repository where MQACP is applied on, the lower the return of the investment is.

7.3. Research Contributions

The research carried out within the scope of this Thesis, made the following contributions:

- **Proposed a Metadata Quality Assurance Certification Process** that can be applied during the lifecycle of a repository, to monitor the quality of metadata created and contribute to their continuous improvement. To achieve that, we defined specific methods that can be

used as well as the tools that have to be deployed in the context of these methods,

- **Applied the proposed MQACP to a Learning Federation** of repositories hosting learning resources in multiple languages. More specifically, we inserted the proposed methods and tools into specific points in the repository lifecycle and managed to showcase significant improvement to the quality of the metadata created within the federation,

- **Validated the effectiveness of MQACP in new contexts** by applying it to a Cultural Federation hosting cultural content from museums of natural history as well as to a Research Federation, hosting bibliographic and research data from institutions around Europe,

- **Identified future areas of research** that came out of the application of the MQACP in real cases with real users. More specifically, extensions of our work are offered in the next section where directions of research are discussed trying to address problematic areas that came up during the MQACP application in all three cases.

7.4. Research Considerations

Through the application of MQACP on the three cases chosen, a number of issues came up that require more extensive research to be addressed. To this end, we present these issues along with some first steps that have already been taken from ongoing research to address them.

7.4.1. Metadata Training

During the application of MQACP on any given federation, one of the first steps that is carried out, is a Metadata Understanding Session with the participation of the domain experts that will also provide the content. The scope of this session is to allow the experts to understand the metadata elements that will be used to describe the content and also affect the selection of these elements through their needs. During these sessions, each domain expert is presented with the application profile to be used and he/she provides a rating of the easiness, usefulness and necessity of each element as already described. In some cases though, it was obvious to us that this approach was not always working. As we collected and analyzed the ratings coming out of theses sessions, we identified two major problems:

- For many elements, the comprehension from the experts was dramatically low, leading afterwards in low completeness for these elements in the respective repositories,

- For many elements, the necessity for the targeted use by the domain experts was also pretty low, indicating that some of the elements that we proposed to the domain experts were not considered useful by them. These elements were often underused in the repositories.

This phenomenon showed that we should focus more on the process followed during the Metadata Understanding Sessions (MUS) and try to (a) make the elements easier to understand for the domain experts and (b) allow a more active role to the domain experts in designing the application profile to be used in the respective federations.

In Palavitsinis *et al.*, (2013) we present a Creative Metadata Understanding Session which is a new outlook on the traditional MUS that was deployed in MQACP so far. We adopted and adapted an existing approach by Agro-Know Technologies' format (the AK Creativity Package), used with communities of domain experts to generate ideas on environmental issues. To this end, we increased the steps of the process and grouped them in broader categories to map in a more efficient way the components of an application profile on which we wanted to train the domain experts.

7.4.2. Creative Metadata Understanding Session

One lesson learned from the previous Metadata Understanding Sessions was that the reference to terms such as metadata or related technical terms is problematic for domain experts. To this end, in the new approach to the MUS, any reference to such terms is avoided to make sure that the participants that are mostly domain experts are not confused. To this end, in the description of the phases, both here and during the actual workshop, the terms that follow are replaced with terms that are generic and understandable:

- Metadata element: Attribute
- Element value/vocabulary: Value
- Metadata Schema: Set of attributes

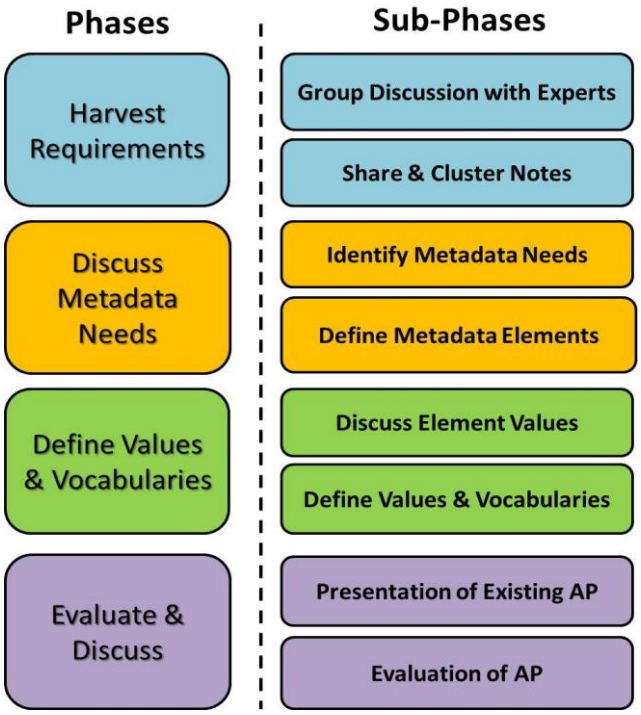


Figure 7.7: Overview of the Creative MUS Process

In Figure 7.7, an overview of the Creative MUS approach is presented, with the main phases and the sub-phases in which they are broken down. Roughly, the first phase is intended to serve as an introduction to metadata where the domain experts discuss amongst them related to their needs whereas in the second phase, the discussion becomes more specific focusing on specific elements. In the third phase, the discussion revolves around vocabularies, allowing the experts to gain a better understanding of the elements through the values they take. Finally, in the last step, the existing application profile that we need to the experts to work with is presented to them and their evaluations are collected.

Harvest Requirements

During this phase, the participants of the Metadata Understanding Session try to collect requirements related to the metadata that they would need in order to describe a digital resource. To do so, the process used, discusses about metadata in an indirect way, so as not to confuse participants with technical terms.

Group Discussion with Experts

The participants are posed with a question, related to their needs when describing or searching for digital content on their subject. More specifically, they are asked to write on a post-it and share with the group, the attributes of a digital resource they would provide for a resource they would create themselves. In the same sense, they are also asked which attributes they would

like to be able to use when searching through this content, in a search engine or portal. The two questions posed to the group are the following:

- Which attributes of the digital object you create would you consider important for a person that wants to use it? Title? Description? Other attributes?
- When searching for digital content online, which attributes of a digital object would you consider important when searching? Its size? Its title? Other?

To manage and direct the discussion, a facilitator is appointed who provides the questions above to the participants and makes sure that the timing is kept. Once the questions are given to the group, each participant is given ten (10) minutes to think the questions on their own and take notes.

Share & Cluster Notes

Once all the participants have their notes, they are asked to share on a common space, the attributes they came up with, explaining the rationale to the group. All the notes are put on the wall with no specific order, and the participants take turns to provide their view on the questions posed in the beginning. During this phase, the facilitator may also coordinate a discussion among the participants, as there can be different opinions related to the attributes that each participant chooses. To continue with, the notes are collected on the wall and the participants are asked to approach and try to create clusters of similar attributes, that is connect similar of them, others that might refer to metadata elements that may belong under the same category, etc. Coming out of this part, the participants have to have a group of clusters that would roughly correspond to the high level metadata categories of a metadata standard, i.e. the General, Lifecycle, Educational, etc. categories of the IEEE LOM standard. The purpose of this exercise is to give to the participants the perception of the actual organization of the data that may "follow" a digital object that is the metadata record itself.

Discuss Metadata Needs

During this phase, the group is introduced to the needs of organizing the information that will describe the digital object into a coherent schema of information so that the provision of this information during the creation of a new digital object is made easier to understand and carry out. No reference to metadata is made yet, to avoid confusion with the technical terms involved.

Identify Metadata Needs

The participants are asked to revisit the clusters of attributes that they created and decide on the ones that are necessary for the description of the digital object. They are asked to keep the ones that are either really important or less important for them, excluding only the ones that are of minor importance for them. The process of deciding on the retention of the attributes or not, is

carried out by the group with no intervention of the facilitator other than answer any trivial questions about the expected outcomes of the process.

Define Metadata Elements
Having kept the attributes that the participants deem as necessary, the facilitator asks them to title the clusters of attributes with one word, so that they provide a concrete title for each one. The same exercise is carried out for each attribute. In this case, next to the post-its describing the attribute, a new post-it is placed with one or two words that represent the title for this attribute and consequently the metadata element that will be created from it. This part of the process is completed with the participants having a final attribute set arranged in categories.

Define Values & Vocabularies
During this phase, the participants of the Metadata Understanding Session are starting to familiarize themselves with the process of structuring a metadata application profile by defining the values that each element can take, whether it's a free-text field or a field that contains a vocabulary, ontology, etc.

Discuss Element Values
The facilitator asks the participants to place the attributes that they have decided upon on a new space and use post-its again to define the type of values that these attributes would take. The participants are split into groups corresponding to one or two categories of elements as these are decided previously. Each group is given twenty (20) minutes to think on the values that these attributes would take. Once they are finished, the groups come together and each group shares their view of the possible values. Answers here may include specific values such as "user, teacher, author", etc. or ranges of values, i.e. "10-100" or even specific ontologies that the participants know of. As the groups share their views on the values, other groups are allowed to add upon what they hear so that they enrich the notes taken and also contribute themselves to other element categories.

Define Values & Vocabularies
Finishing with the sharing of the possible element values, the participants are asked to decide on the final value of each attribute in the cases where more than one possibility is discussed within the groups. Finishing with this phase, the participants have defined the set of attributes they need to describe a digital object, they have clustered them into groups of attributes with a specific title and they have also defined the possible values of all attributes in all groups.

Evaluate & Discuss
During this phase, the time has come to expose the participants to a completed metadata application profile that is already used to describe digital

objects. The presentation of this application profile follows, to allow the participants to make the connection between what they have defined and what is presented.

Presentation of Existing AP

During this phase the actual application profile is presented, element by element providing also examples of use for each one. The participants already have a form in their hands, where they are asked to rate each element that is presented using a 5 point Likert scale, in terms of the following (1 being the lowest):

- Is this element easy to understand?
- Is this element useful for describing digital resources?
- How important is this element?

Evaluation of Existing AP

During this phase that is almost parallel to the previous, the participants complete their evaluations of the elements as these are presented and a discussion follows that is facilitated by the workshop facilitator, related to the similarities and differences between the attributes/elements defined by the participants and the ones proposed by the actual application profile.

Following the principle of strict time management that the "Guided Brainstorming" technique dictates, we limit the work of the domain experts in each phase to the following times provided in Table 7.10. The facilitator of the Creative MUS has to make sure that the groups working with the metadata concepts finish with each assignment given to them on time so that the process moves along quickly.

Phase	Proposed Time
1. Harvest Requirements	**40'**
1.1 Group Discussion with Experts	20'
1.2 Share & Cluster Notes	20'
2. Discuss Metadata Needs	**30'**
2.1 Identify Metadata Needs	15'
2.2 Define Metadata Elements	15'
3. Define Values & Vocabularies	**35'**
3.1 Discuss Element Values	20'
3.2 Define Values & Vocabularies	15'
4. Evaluate & Discuss	**45'**
4.1 Presentation of Application Profile	30'
4.2 Evaluation of Application Profile	15'
TOTAL	**150'**

Table 7.10: Creative Metadata Understanding Session Time Slots

Looking at the time required to carry out the Creative MUS we see that the total time is increased by 30 minutes from the original MQACP which does not affect that much the overall cost of the MQACP as this was discussed

previously. This approach has not yet been tested as it was described here, so no findings can be discussed related to its performance. Our aim is to test it with real users and compare the results to the ones of the MUS to see if the two bigger problems that we have identified with the MUS approach can be addressed through the Creative MUS.

7.4.3. Measuring Quality

One of the open issues that came up during the application of MQACP, was the criteria used to assess quality. In the MQACP application it was obvious that mostly completeness was used as a quality metri. Although other metrics were deployed in the peer-review experiments, these were not cost-effective as they could not be applied easily to the total number of metadata records hosted in each federation. In the end, a small sample of resources of each federation was checked each time against these metrics, which did not allow for the MQACP to deduct absolute measures of quality for the whole collection examined but rather for a small portion of it. To this direction, Tsiflidou & Manouselis (2013) was based on the work carried out in this thesis to examine how different metadata quality metrics can be extracted using automated tools, allowing for safer measurements of the overall metadata quality of a collection. More specifically the authors mentioned above, adopted a set of metadata quality metrics as shown below, which can be calculated using specific tools, automatically for an entire collection.

Metric	Reference	Definition
Completeness	Bruce & Hillman (2004)	The percentage of records in which an element is used.
Element frequency	Ochoa et al. (2006)	This metric provides information about the number of values used in a metadata element
Entropy	Ochoa et al.(2006)	The entropy provides information about the amount of information that is included in an element.
Vocabularies Values distribution	Ochoa et al. (2006)	The metric is being used to study the frequency distribution of specific vocabulary values in controlled elements
Metadata multi-linguality for the free text elements	Vogias et al. (2013)	This metric is used to study the language attribute (eg. Lang=en) value usage frequency in free text metadata elements such as Title, Description and Keyword

Table 7.11: Metadata Quality Metrics adopted by Tsiflidou & Manouselis (in press)

The tools that they used to assess these metrics, were Google Refine[14], MINT[15] and the AK Metadata Analytics Tool (Vogias et al., 2013). In their work, Tsiflidou & Manouselis (2013) used the aforementioned metrics and tools, to examine a collection of almost 2.500 metadata records. The results have showed great promise for the use of the metrics proposed and their introduction in the MQACP as additional metrics that will define the metadata quality of any given collection with significantly lower cost rather than peer-reviewing each record based on the Metadata Quality Grid which may involve a more thorough set of metrics but is less scalable.

7.4.4. Metadata Quality Cost Considerations

The parameter of cost is a really important factor for the MQACP proposed in this thesis. Our literature review though, has retrieved a limited amount of relevant studies that discuss the cost in detail. In very few cases is the cost of creating a metadata record or enriching it, discussed. Similarly, the cost of applying a quality assurance process on the metadata records of a repository is a topic that is not discussed although the initiatives that deploy and maintain large repositories continue to multiply. To this direction, we carried out a first discussion of the costs involved in applying the MQACP on a repository through its lifecycle, determining the number of people that have to be involved and the time they need to invest to participate in the experiments carried out. From our analysis it became evident that there are more hidden costs involved that were not covered sufficiently. Costs such as the time needed for the content providers to review their resources and enrich them with new metadata as well as the time needed for the content providers to work with the supporting material and guides presented to them are not factored in to our estimation.

To this direction, we feel that more work is needed to accurately document the metadata associated costs from a curatorial point of view, starting from the selection of the content to its annotation with metadata and their continuous preservation. These costs have to be added to the costs of applying MQACP to provide a holistic view of the total costs that an institution has to take into consideration when developing a repository of resources. Although some of these considerations are already covered by work carried out in the library sciences domain, we feel that the unique nature of every different domain that digital repositories serve, calls for focused research that exceeds the library domain and includes in the whole process stakeholders from different domains. For example, the cost of training domain experts on metadata annotation goes beyond existing library approaches and should also be

[14] https://code.google.com/p/google-refine/

[15] http://mint.image.ece.ntua.gr/redmine/projects/mint/wiki

considered. As a first step, Table 7.12 offers a non-exhaustive list of the associated metadata costs that a repository manager has to take into consideration.

No	Cost Category	Included in MQACP	Explanation
1	Metadata Design	Yes	Designing an application profile based on user needs.
2	Resource Selection	No	Selecting the resources to be populated in the repository
3	Metadata Annotation	No	Adding metadata to the resources from scratch
4	Metadata Enrichment	No	Enriching/correcting problematic metadata fields
5	Peer Review of Metadata	Yes	Checking the quality of metadata through peer review processes
6	Automated Metadata Review	Yes	Checking the quality of metadata through automated means
7	Development of Training Material	Yes	Develop guidelines and manuals for the content providers
8	Training on Metadata	No	Hold training sessions on metadata annotation / spend time on reading supportive material

Table 7.12: Metadata Associated Costs for Repositories

The costs discussed in Table 7.12 are just some of the costs that are involved in the process of populating a repository with resources. Technical costs for the deployment of the metadata application profile used or the development of metadata authoring tools are also part of the development of a repository. Starting to map all of these costs and accurately describing them, will provide useful insights for the development process of repositories but also help their managers define their business value and sustainability aspects in the long-term.

References

Al-Khalifa, H. S., & Davis, H. C. (2006). “The evolution of metadata from standards to semantics in elearning applications”. Proceedings of the Seventeenth Conference on Hypertext and Hypermedia, Odense, Denmark.

Abbas, J. (2005). Creating metadata for children's resources: Issues, research, and current developments. Library trends, 54(2), 303-317.

Abdullah, A., & Zainab, A. N. (2008). The digital library as an enterprise: the Zachman approach. Electronic Library, The, 26(4), 446-467.

Abrami, P. C., Bernard, R., Wade, A., Schmid, R. F., Borokhovski, E., Tamin, R., ... & Peretiatkowicz, A. (2008). A review of e-learning in Canada: A rough sketch of the evidence, gaps and promising directions. Canadian Journal of Learning and Technology/La revue canadienne de l’apprentissage et de la technologie, 32(3).

Ackermann, E. (2008). Library assessment on a budget: Using effect size meta-analysis to get the most out of the library-related survey data available across campus. Performance Measurement and Metrics, 9(3), 192-201.

Afshari, F., & Jones, R. (2007). Developing an integrated institutional repository at Imperial College London. Program: electronic library and information systems, 41(4), 338-352.

Agostinho, S., Bennett, S., Lockyer, L., & Harper, B. (2004). Developing a learning object metadata application profile based on LOM suitable for the Australian higher education context. Australasian Journal of Educational Technology, 20(2), 191-208.

Ahmed, S. Z., & Shoeb, M. Z. H. (2009). Measuring service quality of a public university library in Bangladesh using SERVQUAL. Performance Measurement and Metrics, 10(1), 17-32.

Akpinar, Y. (2008). validation of a learning object Review Instrument: Relationship between Ratings of learning objects and Actual learning outcomes. Interdisciplinary Journal of E-Learning and Learning Objects, 4(1), 291-302.

Al-alak, B. A., & Alnawas, I. A. (2011). Measuring the Acceptance and Adoption of E-Learning by Academic Staff. Knowledge Management & E-Learning: An International Journal (KM&EL), 3(2), 201-221.

Alberta Learning (2002-2008). Learn Alberta glossary. [viewed 21 Nov 2005, not found 13 Sep 2008] http://www.learnalberta.ca/Browse.aspx?View=Glossary

Aleahmad, T., Aleven, V., & Kraut, R. (2009). Creating a corpus of targeted learning resources with a Web-Based open authoring tool. Learning Technologies, IEEE Transactions on, 2(1), 3-9.

Alijani, A. S., & Jowkar, A. (2009). Dublin Core Metadata Element Set usage in national libraries' web sites. Electronic Library, The, 27(3), 441-447.

Alimohammadi, D. (2005). Meta-tags: still a matter of opinion. Electronic Library, The, 23(6), 625-631.

Alipour-Hafezi, M., Horri, A., Shiri, A., & Ghaebi, A. (2010). Interoperability models in digital libraries: an overview. Electronic Library, The, 28(3), 438-452.

Alkhattabi, M., Neagu, D., & Cullen, A. (2010). Information quality framework for e-learning systems. Knowledge Management & E-Learning: An International Journal (KM&EL), 2(4), 340-362.

Allard, S. (2002). Digital libraries and organizations for international collaboration and knowledge creation. Electronic Library, The, 20(5), 369-381.

Allinson, J. (2008). Describing scholarly works with Dublin Core: A functional approach. Library Trends, 57(2), 221-243.

Alonzo, A.C. (2001), "Virtual heritage, archeology, and cultural heritage",Proceedings of the 2001 Conference on Virtual Reality, Archeology and Cultural Heritage, Glyfada, November 28-30, pp. 343-54.

Alva, M. E., Martinez, A. B., Suarez, M. D. C., Labra, J. E., Cueva, J., & Sagastegui, H. (2010). Towards the evaluation of usability in educative websites. International Journal of Technology Enhanced Learning, 2(1), 145-161.

Anderson, I. G. (2007). Pure dead brilliant?: Evaluating The Glasgow Story digitisation project. Program: electronic library and information systems, 41(4), 365-385.

Anido, L. E., Fernández, M. J., Caeiro, M., Santos, J. M., Rodrı´guez, J. S., & Llamas, M. (2002). Educational metadata and brokerage for learning resources.Computers & Education, 38(4), 351-374.

Anido, L., Rodríguez, J., Caeiro, M., & Santos, J. M (2003). High-level brokerage services for the elearning domain. Computer Standards & Interfaces, 25(4), 303-327.

Arazy, O., & Kopak, R. (2011). On the measurability of information quality.Journal of the American Society for Information Science and Technology, 62(1), 89-99.

Arms, W. Y., Dushay, N., Fulker, D., & Lagoze, C. (2003). A case study in metadata harvesting: the NSDL. Library Hi Tech, 21(2), 228-237.

Armstrong, C.J. (1995). Database quality: label or liable. In: P. Wressel (ed.), Proceedings of the 1st Northumbria International Conference on Performance Measurement in Libraries and Information Services (Information North, Newcastle upon Tyne, 1995) 203–6.

Arshad, A., & Ameen, K. (2010). Service quality of the University of the Punjab's libraries: An exploration of users' perceptions. Performance Measurement and Metrics, 11(3), 313-325.

Aschenbrenner, A., Blanke, T., Flanders, D., Hedges, M., & O'Steen, B. (2008). The future of repositories? Patterns for (cross-) repository architectures. D-Lib Magazine, 14(11/12).

Aschenbrenner, A., Blanke, T., Küster, M. W., & Pempe, W. (2010). Towards an open repository environment. Journal of Digital Information, 11(1).

Ashoor, M. S. (2000). Planning the electronic library–suggested guidelines for the Arabian Gulf region. Electronic Library, The, 18(1), 29-39.

Atakan, C., Atilgan, D., Bayram, Ö., & Arslantekin, S. (2008). An evaluation of the second survey on electronic databases usage at Ankara University Digital Library. Electronic Library, The, 26(2), 249-259.

Ataman, B. K. (2009). Requirements for information professionals in a digital environment: some thoughts. Program: electronic library and information systems, 43(2), 215-228.

Attig, J., Copeland, A. and Pelikan, M. (2004), "Context and meaning: the challenges of metadata for a digital image library within the university",College and Research Libraries, Vol. 65 No. 3, pp. 251-61.

Augustyniak, R. H., Aguero, D. B., & Finley, A. (2005). The IP's guide to the galaxy of portal planning: Part II. Content management. Online Information Review, 29(6), 643-655.

Averkamp, S., & Lee, J. (2009). Repurposing ProQuest metadata for batch ingesting ETDs into an institutional repository. University Libraries Staff Publications, 38.

Babalhavaeji, F., Isfandyari-Moghaddam, A., Aqili, S. V., & Shakooii, A. (2010). Quality assessment of academic libraries' performance with a special reference to information technology-based services: Suggesting an evaluation checklist.Electronic Library, The, 28(4), 592-621.

Bainbridge, D., Thompson, J. and Witten, I.H. (2003), "Assembling and enriching digital library collections",Proceedings of the 3rd ACM/IEEE-CS Joint Conference on Digital Libraries, pp. 323-34.

Baker, D. (2006). Digital library futures: a UK HE and FE perspective.Interlending & Document Supply, 34(1), 4-8.

Baker, T. (2012). Libraries, languages of description, and linked data: a Dublin Core perspective. Library Hi Tech, 30(1), 116-133.

Baker, T., Dekkers, M., Heery, R., Patel, M., & Salokhe, G. (2006). What terms does your metadata use? Application profiles as machine-understandable narratives. Journal of Digital information, 2(2).

Ball, A., Darlington, M., Howard, T., McMahon, C., & Culley, S. (2012). Visualizing research data records for their better management. Journal of Digital Information, 13(1).

Ballou, D.P., & Tayi, G.K. (1999). Enhancing data quality in data warehouse environments, Communications of the ACM42(1), pp. 73-78.

Banach, M., & Li, Y. (2011). Institutional repositories and digital preservation: Assessing current practices at research libraries. D-Lib Magazine, 17(5/6).

Bar-Ilan, J., Shoham, S., Idan, A., Miller, Y., & Shachak, A. (2008). Structured versus unstructured tagging: a case study. Online Information Review, 32(5), 635-647.

Barker, E., & Ryan, B. (2005). Case studies in implementing educational metadata standards. Centre for Educational Technology Interoperability Standards, United Kingdom. Retrieved August 26th, 2005, from http://metadata.cetis.ac.uk/usage_survey/

Barton, J., Currier, S. & Hey, J.M.N. (2003), "Building quality assurance into metadata creation: an analysis based on the learning objects and e-prints communities of practice", Proceedings 2003 Dublin Core Conference: Supporting Communities of Discourse and Practice – Metadata Research and Applications (DC-2003), Seattle, WA, USA, 28 September-2 October 2003, available at: http://eprints.rclis.org/archive/00001972/ (accessed January 2007).

Barton, M. R., & Walker, J. H. (2003). Building a business plan for DSpace, MIT libraries digital institutional repository.

Barton, M. R., & Waters, M. M. (2004). Creating an institutional repository: LEADIRS workbook.

Barwick, J. (2007). Building an institutional repository at Loughborough University: some experiences. Program: electronic library and information systems, 41(2), 113-123.

Bawden, D., & Robinson, L. (2009). The dark side of information: overload, anxiety and other paradoxes and pathologies. Journal of information science,35(2), 180-191.

Beall, J. (2005), "Metadata and data quality problems in the digital library", University of Colorado at Denver and Health Sciences Centre, Denver, available at: http://journals.tdl.org/jodi/article/view/jodi-171/68.

Beall, J. (2006). Metadata and data quality problems in the digital library.Journal of Digital Information, 6(3).

Beamer, A. (2009). Map metadata: essential elements for search and storage.Program: electronic library and information systems, 43(1), 18-35.

Beard, J., & De Vekey, J. (2004). The quality assurance of Healthinfo4u: a Web-based Internet resource for the lay public. Performance measurement and metrics, 5(1), 28-34.

Beard, J., Dale, P., & Hutchins, J. (2007). The impact of e-resources at Bournemouth University 2004/2006. Performance Measurement and Metrics,8(1), 7-17.

Beard, K. (1996), "A structure for organizing metadata collection", Proceedings of the 3rd International Conference/Workshop on Integrating GIS and Environmental Modeling, Santa Fe, NM, USA, January 21-26, available at: www.ncgia.ucsb.edu/conf/SANTA_FE_CD-ROM/main.html

Bearman, D., Rust, G., Weibel, S., Miller, E., & Trant, J. (1999). A Common Model To Support Interoperable Metadata: Progress Report on Reconciling Metadata Requirements from the Dublin Core and INDECS/DOI Communities. D-Lib magazine, 5(1), n1.

Beaudoin, M., Kurtz, G., & Eden, S. (2009). Experiences and opinions of e-learners: What works, what are the challenges, and what competencies ensure successful online learning. Interdisciplinary Journal of E-Learning and Learning Objects, 5(1), 275-289.

Beer, C., Michael, C., & Todorovic, M. (2009). Visualizing Media Archives: A Case Study. Code4Lib Journal, (6).

Bell, J., & Lewis, S. (2006). Using OAI-PMH and METS for exporting metadata and digital objects between repositories. Program: electronic library and information systems, 40(3), 268-276.

Benjelloun, R. (2005). Archimède: a Canadian solution for institutional repository. Library hi tech, 23(4), 481-489.

Benoit, G., & Hussey, L. (2011). Repurposing digital objects: Case studies across the publishing industry. Journal of the American Society for Information Science and Technology, 62(2), 363-374.

Bertot, J. C. (2004). Libraries and networked information services: issues and consideration in measurement. Performance measurement and metrics, 5(1), 11-19.

Besser, H. (1999). Network access to visual information: A study of costs and uses In IFLA Council and General Conference. 65th Conference programme and proceedings, Bangkok, Thailand, 1999. Retrieved 2006-10-06 from: www.ifla.org/IV/ifla65/papers/021-112e.htm

Bevan, S. J. (2005). Electronic thesis development at Cranfield University.Program: electronic library and information systems, 39(2), 100-111.

Bhattacharya, P. (2004). Advances in digital library initiatives: a developing country perspective. The International Information & Library Review, 36(3), 165-175.

Bin, F., & Miao, Q. (2005). Electronic publications for Chinese public libraries: challenges and opportunities. Electronic Library, The, 23(2), 181-188.

Björk, B. C. (2006). Electronic document management in temporary project organisations: Construction industry experiences. Online Information Review,30(6), 644-655.

Blustein, J., Webber, R. E., & Tague-Sutcliffe, J. (1997). Methods for evaluating the quality of hypertext links. Information processing & management, 33(2), 255-271.

Bond, S. T., Ingram, C., & Ryan, S. (2008). Reuse, repurposing and learning design–Lessons from the DART project. Computers & Education, 50(2), 601-612.

Borges, E. N., de Carvalho, M. G., Galante, R., Gonçalves, M. A., & Laender, A. H. (2011). An unsupervised heuristic-based approach for bibliographic metadata deduplication. Information Processing & Management, 47(5), 706-718.

Borgman, C. (2000),From Gutenberg to Global Information Infrastructure, MIT Press, Boston, MA

Borgman, C.L. (1996a), "Why are online catalogs hard to use? Lessons learned from information retrieval studies", Journal of the American Society for the Information Science, Vol. 37 No. 6, pp. 387-400.

Borgman, C.L., Bates, M.J., Cloonan, M.V., Efthimiadis, E.N., Gilliland-Swetland, A. J., Kafai, Y., (...), & Maddox, A. (1996). Social Aspects of Digital Libraries. Final Report to the National Science Foundation; Computer, Information Science, and Engineering Directorate; Division of Information, Robotics, and Intelligent Systems; Information Technology and Organizations Program (1996). http://is.gseis.ucla.edu/research/dl/index.html.

Borokhovski, E., Bernard, R., Mills, E., Abrami, P. C., Wade, C. A., Tamim, R., ... & Surkes, M. A. (2011). An Extended Systematic Review of Canadian Policy Documents on e-Learning: What We're Doing and Not Doing. Canadian Journal of Learning and Technology/La revue canadienne de l'apprentissage et de la technologie, 37(3).

Bountouri, L., Papatheodorou, C., Soulikias, V., & Stratis, M. (2009). Metadata interoperability in public sector information. Journal of Information Science,35(2), 204-231.

Boyle, T. (2003). Design principles for authoring dynamic, reusable learning objects. Australian Journal of Educational Technology, 19(1), 46-58.

Bradley, C., & Boyle, T. (2004). The design, development, and use of multimedia learning objects. Journal of Educational Multimedia and Hypermedia, 13(4), 371-389.

Bremer, C. (2012). Enhancing e-learning quality through the application of the AKUE procedure model. Journal of Computer Assisted Learning, 28(1), 15-26.

Brindis, R. G., Fitzgerald, S., Anderson, H. V., Shaw, R. E., Weintraub, W. S., & Williams, J. F. (2001). The American College of Cardiology-National Cardiovascular Data Registry™(ACC-NCDR™): building a national clinical data repository. Journal of the American College of Cardiology, 37(8), 2240-2245.

Brindley, G., Muir, A., & Probets, S. (2004). Provision of digital preservation metadata: a role for ONIX?. Program: electronic library and information systems, 38(4), 240-250.

Britain, S. (2004). A review of learning design: concept, specifications and tools. A report for the JISC E-learning Pedagogy Programme, 2006.

Broady-Preston, J., & Lobo, A. (2011). Measuring the quality, value and impact of academic libraries: the role of external standards. Performance Measurement and Metrics, 12(2), 122-135.

Brophy, P. (2008). Telling the story: qualitative approaches to measuring the performance of emerging library services. Performance measurement and metrics, 9(1), 7-17.

Brown, M. I., Doughty, G. F., Draper, S. W., Henderson, F. P., & McAteer, E. (1996). Measuring learning resource use. Computers & Education, 27(2), 103-113.

Bruce, T.R., & Hillmann, D.I. (2004). The Continuum of Metadata Quality: Defining, Expressing, Exploiting. In Metadata in Practice, edited by D.I. Hillmann and E.L. Westbrooks (Chicago: American Library Association) pp. 238-256.

Brusilovsky, P. (2003). Developing adaptive educational hypermedia systems: From design models to authoring tools. In Authoring Tools for Advanced Technology Learning Environments (pp. 377-409). Springer Netherlands.

Bryden, A. (2003) Open and GlobalStandards for Achieving an Inclusive Information Society. Retrieved January 13, 2005 from http://www.iso.org/iso/en/commcentre/presentations/secgen/2003/ajb2003SISTspeech.pdf

Burnett, K., Ng, K. B., & Park, S. (1999). A comparison of the two traditions of metadata development. Journal of the American Society for Information Science, 50(13), 1209-1217.

Buzzetto-More, N., & Pinhey, K. (2006). Guidelines and standards for the development of fully online learning objects. Interdisciplinary Journal of E-Learning and Learning Objects, 2(1), 95-104.

Byamugisha, H. M. (2010). Digitizing library resources for new modes of information use in Uganda. Library Management, 31(1/2), 42-56.

Bygstad, B., Ghinea, G., & Klæboe, G. T. (2009). Organisational challenges of the Semantic Web in digital libraries: A Norwegian case study. Online Information Review, 33(5), 973-985.

Byrne, A. (2003). Digital libraries: barriers or gateways to scholarly information?. Electronic Library, The, 21(5), 414-421.

Cahill, K. (2011). Going social at Vancouver Public Library: what the virtual branch did next. Program: electronic library and information systems, 45(3), 259-278.

Calero, C., Ruiz, J., & Piattini, M. (2005). Classifying web metrics using the web quality model. Online Information Review, 29(3), 227-248.

Calhoun, K. (2007). Being a librarian: metadata and metadata specialists in the twenty-first century. Library hi tech, 25(2), 174-187.

Calverley, G., & Shephard, K. (2003). Assisting the uptake of on-line resources: why good learning resources are not enough. Computers & Education, 41(3), 205-224.

Campbell, D. (2002). Federating access to digital objects: PictureAustralia.Program: electronic library and information systems, 36(3), 182-187.

Campbell, D. (2002). The use of the Dublin Core in web annotation programs. In Proceedings of the international conference on Dublin Core and metadata for e-communities(pp. 105–110).

Carey, T., Swallow, J., & Oldfield, W. (2002). Educational rationale metadata for learning objects. Canadian Journal of Learning and Technology/La revue canadienne de l'apprentissage et de la technologie, 28(3).

Carlson, J., Ramsey, A. E., & Kotterman, J. D. (2010). Using an institutional repository to address local-scale needs: a case study at Purdue University.Library Hi Tech, 28(1), 152-173.

Cassella, M. (2010). Institutional repositories: an internal and external perspective on the value of IRs for researchers' communities. Liber quarterly, 20(2), 210-225.

Catapano, T., DiPasquale, J., & Marquis, S. Building an Archival Collections Portal. Code4Lib Journal.

Caws, C., Friesen, N., & Beaudoin, M. (2006). A new learning object repository for language learning: Methods and possible outcomes. Interdisciplinary Journal of E-Learning and Learning Objects, 2(1), 111-124.

Cebeci, Z., & Erdogan, Y. (2005). Tree view editing learning object metadata. Interdisciplinary Journal of E-Learning and Learning Objects, 1(1), 99-108.

Cebeci, Z., Erdogan, Y., & Kara, M. (2008, April). TrAgLor: A LOM-Based Digital Learning Objects Repository for Agriculture. In Proc. of the 4th Int. Scientific Conference "eLearning and Software for Education (eLSE'08) (pp. 17-18).

Cechinel, C., & Sánchez-Alonso, S. (2011). Analyzing associations between the different ratings dimensions of the MERLOT repository. Interdisciplinary Journal of E-Learning and Learning Objects, 7(1), 1-9.

Cechinel, C., & Sánchez-Alonso, S. (2011). Analyzing associations between the different ratings dimensions of the MERLOT repository.Interdisciplinary Journal of E-Learning and Learning Objects, 7,1–9.

Cechinel, C., Sánchez-Alonso, S., & García-Barriocanal, E. (2011). Statistical profiles of highly-rated learning objects. Computers & Education, 57(1), 1255-1269.

Chan, L. M. (1989) "Inter-indexer consistency in subject cataloging". Information Technology & Libraries, 8(4), 349-358

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, 12(6).

Chand, P., & Arora, J. (2008). Access to scholarly communication in higher education in India: trends in usage statistics via INFLIBNET. Program: electronic library and information systems, 42(4), 382-390.

Chang, S.H."Institutional repositories: the library's new role", OCLC Systems and Services. Vol 19 No.3. (2003): 77-79

Chapman, A. (2007). Resource discovery: Catalogs, cataloging, and the user.Library trends, 55(4), 917-931.

Chapman, J., & Lown, C. (2010). Practical Ways to Promote and Support Collaborative Data Analysis Projects.

Charalabidis, Y., & Askounis, D. (2008, January). Interoperability registries in eGovernment: Developing a semantically rich repository for electronic

services and documents of the new public administration. In Hawaii International Conference on System Sciences, Proceedings of the 41st Annual (pp. 195-195). IEEE.

Chavez, R., Crane, G., Sauer, A., Babeu, A., Packel, A., & Weaver, G. (2007). Services make the repository. Journal of Digital Information, 8(2).

Chen, B.H., Hung, S.H. & Hong, J.S. (2003), "Modularization framework for digital museum exhibition",Proceedings of the 3rd ACM/IEEE-CS Joint Conference on Digital Libraries, Houston, TX, May 27-31, p. 394.

Chen, C. C., Chen, H. H., Chen, K. H., & Hsiang, J. (2002). The design of metadata for the Digital Museum Initiative in Taiwan. Online Information Review,26(5), 295-306.

Chen, K. H., & Hsiang, J. (2009). The unique approach to institutional repository: practice of National Taiwan University. Electronic Library, The, 27(2), 204-221.

Chen, Y. C., Hwang, R. H., & Wang, C. Y. (2012). Development and evaluation of a Web 2.0 annotation system as a learning tool in an e-learning environment.Computers & Education, 58(4), 1094-1105.

Chen, Y. N., Chen, S. J., Chiang, H. Y., & Liu, C. H. (2000). A case study in designing Chinese metadata. Online Information Review, 24(3), 229-234.

Chepesuik, R. (1999). Organizing the Internet: The Core of the challenge. American Libraries (January), 60–64.

Childs, S. (2005). Judging the quality of internet-based health information.Performance Measurement and Metrics, 6(2), 80-96.

Choo, C.W. (2002). Information Management for the Intelligent Organization: The Art of Scanning the Environment, 3rd ed., Information Today, Medford, NJ, pp. 158, 224, 238.

Choudhury, G. S. (2008). Case study in data curation at Johns Hopkins University. Library Trends, 57(2), 211-220.

Chu, H. C., Hwang, G. J., & Tseng, J. C. (2010). An innovative approach for developing and employing electronic libraries to support context-aware ubiquitous learning. Electronic Library, The, 28(6), 873-890.

Chu, H. C., Hwang, G. J., & Tseng, J. C. (2010). An innovative approach for developing and employing electronic libraries to support context-aware ubiquitous learning. Electronic Library, The, 28(6), 873-890.

Chu, K. W., Wang, M., & Yuen, A. H. K. Implementing knowledge management in school environment: Teachers' perception.

Chua, A.Y.K. & Goh D.H., A study of web 2.0 applications in library websites, Library & Information Science Research 32(2010) 203–211.

Clarke, S. G. D. (2008). The last 50 years of knowledge organization: a journey through my personal archives. Journal of Information Science, 34(4), 427-437.

Clavel-Merrin, G. (2010). Many roads to information: digital resource sharing and access at the Swiss National Library. Interlending & Document Supply,38(1), 54-57.

Clements, K. I., & Pawlowski, J. M. (2012). User-oriented quality for OER: understanding teachers' views on re-use, quality, and trust. Journal of Computer Assisted Learning, 28(1), 4-14.

Cleveland-Innes, M., McGreal, R., Anderson, T., Friesen, N., Ally, M., Tin, T., ... & Schafer, S. (2005). The Athabasca University eduSource Project: Building an accessible learning object repository. Australasian Journal of Educational Technology, 21(3), 367.

Cochrane, T. (2005). Interactive quicktime: Developing and evaluating multimedia learning objects to enhance both face-to-face and distance e-learning environments. Interdisciplinary Journal of Knowledge and Learning Objects, 1(1), 33–54. Retrieved from http://ijello.org/Volume1/v1p033-054Cochrane.pdf

Cohen, A., Shmueli, E., & Nachmias, R. (2011). [Chais] The Usage of Data Repositories: The Case of MAOR. Interdisciplinary Journal of E-Learning and Learning Objects, 7(1), 323-338.

Cohen, E., & Nycz, M. (2006). Learning objects and e-learning: An informing science perspective. Interdisciplinary Journal of E-Learning and Learning Objects, 2(1), 23-34.

Cole, T. W., & Shreeves, S. L. (2004). Search and discovery across collections: the IMLS digital collections and content project. Library Hi Tech,22(3), 307-322.

Coleman, A., Bracke, P., & Karthik, S. (2004). Integration of non-OAI resources for federated searching in DLIST, an eprints repository. D-Lib Magazine, 10(7/8), 1082-9873.

Collier, M. (2004). Development of a business plan for an international co-operative digital library–The European Library (TEL). Program: electronic library and information systems, 38(4), 225-231.

Collins, L. M., Hussell, J. A., Hettinga, R. K., Powell, J. E., Mane, K. K., & Martinez, M. L. (2007). Information visualization and large-scale repositories.Library Hi Tech, 25(3), 366-378.

Copeland, S., Penman, A., & Milne, R. (2005). Electronic theses: the turning point. Program: Electronic library and information systems, 39(3), 185-197.

Corrall, S. (2010). Educating the academic librarian as a blended professional: a review and case study. Library Management, 31(8/9), 567-593.

Craig, A., Goold, A., Coldwell, J., & Mustard, J. (2008). Perceptions of roles and responsibilities in online learning: A case study. Interdisciplinary journal of e-learning and learning objects, 4(1), 205-223.

Cress, U., Barquero, B., Schwan, S., & Hesse, F. W. (2007). Improving quality and quantity of contributions: Two models for promoting knowledge exchange with shared databases. Computers & Education, 49(2), 423-440.

Crow, R. (2002). The Case for Institutional Repositories: A SPARC Position Paper, Scholarly Publishing and Academic Resources Coalition, Washington, DC, available at: www.arl.org/sparc/IR/IR_Final_Release_102.pdf

Crowley, G. H., & Gilreath, C. L. (2002). Probing user perceptions of service quality: using focus groups to enhance quantitative surveys. Performance Measurement and Metrics, 3(2), 78-84.

Crystal, A. & Greenberg, J. (2005), "Usability of a metadata creation application for resource authors", Library and Information Science Research, Vol. 27 No. 2, pp. 177-89.

Crystal, A. & Land, P. (2003), "Metadata and Search: Global Corporate Circle DCMI 2003 Workshop", available at: www.dublincore.org/groups/corporate/Seattle/ (accessed 14 January 2011).

Cullen, R., & Chawner, B. (2010). Institutional repositories: assessing their value to the academic community. Performance Measurement and Metrics,11(2), 131-147.

Cundiff, M. V. (2004). An introduction to the metadata encoding and transmission standard (METS). Library Hi Tech, 22(1), 52-64.

Currier, S., Barton, J., O'Beirne, R. and Ryan, B. (2004), "Quality assurance for digital learning object repositories: issues for the metadata creation process",ALT-J Research in Learning Technology, Vol. 12 No. 1, pp. 5-20.

Custard, M., & Sumner, T. (2005). Using machine learning to support quality judgments. D-Lib Magazine, 11(10), 1082-9873.

Dahl, O. J., & Nygaard, K. (1966). SIMULA - An ALGOL based simulation language. Communications of the ACM, 9(9),671-678.

Dalziel, J. (2003). Implementing learning design: The learning activity management system (LAMS).

Davies, R. (2007). Library and institutional portals: a case study. Electronic Library, The, 25(6), 641-647.

Davis, P. M., & Connolly, M. J. (2007). Institutional repositories: evaluating the reasons for non-use of Cornell University's installation of DSpace. D-lib Magazine, 13(3/4).

Day, M. (2001). Metadata in a nutshell.Information Europe, 6,2.

Day, M. W. (1997). Extending metadata for digital preservation. Ariadne, 9.Retrieved February 10, 2005, from http://www.ariadne.ac.uk/issue9/metadata/.

De Salas, K., & Ellis, L. (2006). The development and implementation of learning objects in a higher education setting. Interdisciplinary Journal of E-Learning and Learning Objects, 2(1), 1-22.

Dempsey, L., Heery, R., Hamilton, M., Hiom, D., Knight,J., Koch, T., Peereboom, M. and Powell, A. (1997). Specification for resource description methods. Part1.A review of metadata: a survey of current

resource description formats", available at: http://www.ukoln.ac.uk/metadata/DESIRE/overview (accessed 24May1998).

Deniman, D., Sumner, T., Davis, L., Bhushan, S., & Fox, J. (2006). Merging metadata and content-based retrieval. Journal of Digital Information, 4(3).

DeRidder, J. L. (2008). Googlizing a digital library. The Code4Lib Journal, (2).

DeRidder, J. L. (2011). Leveraging EAD for Low-Cost Access to Digitized Content at the University of Alabama Libraries. Journal of Library Innovation, 2(1), 45-60.

Devillers, R., Gervais, M., Bedard, Y. & Jeansoulin, R. (2002), "Spatial data quality: from metadata to quality indicators and contextual end-user manual", Proceedings of the OEEPE/ISPRS Joint Workshop on Spatial Data Quality Management, Instanbul, Turkey, 21-22 March.

Diamantopoulos, N., Sgouropoulou, C., Kastrantas K., & Manouselis, N. (2011). Developing a Metadata Application Profile for Sharing Agricultural Scientific and Scholarly Research

Dill, S., Eiron, N., Gibson, D., Gruhl, D., Guha, R., Jhingran, A., ... & Zien, J. Y. (2003). A case for automated large-scale semantic annotation. Web Semantics: Science, Services and Agents on the World Wide Web, 1(1), 115-132.

Ding, H., & Sølvberg, I. (2007). Rule-based metadata interoperation in heterogeneous digital libraries. Electronic Library, The, 25(2), 193-206.

Ding, Y., Jacob, E. K., Zhang, Z., Foo, S., Yan, E., George, N. L., & Guo, L. (2009). Perspectives on social tagging. Journal of the American Society for Information Science and Technology, 60(12), 2388-2401.

Dixon, P., Pickard, A., & Robson, H. (2002). Developing a criteria-based quality framework for measuring value. Performance Measurement and Metrics, 3(1), 5-9.

Dobratz, S., Schoger, A., & Strathmann, S. (2007). The nestor catalogue of criteria for trusted digital repository evaluation and certification. Journal of Digital Information, 8(2).

Dobyns, L. & Crawford-Mason, C. (1991). Quality or Else: The Revolution in World Business. Boston: Houghton Mifflin, 1991.

Dodero, J. M., Aedo, I., & Díaz, P. (2002). Participative knowledge production of learning objects for e-books. Electronic Library, The, 20(4), 296-305.

Donaldson, J. (1994), "Quality assessment in Scotland", in Craft, A. (Ed.), International Conference on Developments in Assuring Quality in Higher Education, Falmer Press, London.

Downes, S. (2001). Learning objects: resources for distance education worldwide. The International Review of Research in Open and Distance Learning, 2(1).

Downes, S. (2003, July 31). Design, Standards and Reusability. [viewed Jan 2004, verified 29 may 2004] http://www.downes.ca/cgi-bin/website/view.cgi?dbs=Article&key=1059622263&format=full

Downes, S. (2004). The learning marketplace. Meaning, metadata and content syndication in the learning object economy. Retrieved May 14, 2007 from http://www.downes.ca/files/book3.htm

Draganidis, F., Chamopoulou, P., & Mentzas, G. (2008). A semantic web architecture for integrating competence management and learning paths. Journal of knowledge Management, 12(6), 121-136.

Dron, J., Boyne, C. and Mitchell, R. (2002), "Evaluating assessment usingn-dimensional filtering", Proceedings of the AACE E-Learn Conference

Dushay, N., & Hillmann, D. (2003). Analyzing metadata for effective use and re-use. In Proceedings of the International Conference on Dublin Core and Metadata Applications: Supporting Communities of Discourse and Practice (pp. 1–10). Singapore: Dublin Core Metadata Initiative.

Duval E., Hodgins W., Sutton S., & Weibel SL. (2002). "Metadata Principles and Practicalities". D-Lib Magazine, 8. http://www.dlib.org/dlib/april02/weibel/04weibel.html. Accessed 18 January 2008.

Duval E., Smith N. & Coillie M. V. (2006). "Application Profiles for Learning". Proceedings of the 6th IEEE International Conference on Advanced Learning Technologies, ICALT 2006, Kerkrade, The Netherlands, IEEE Computer Society.

Duval, E., & Verbert, K. (2008). On the role of technical standards for learning technologies. Learning Technologies, IEEE Transactions on, 1(4), 229-234.

Duval, E., Hodgins, W., Sutton, S., & Weibel, S. L. (2002). Metadata principles and practicalities. D-lib Magazine, 8(4), 16.

Ebner, H., Manouselis, N., Palmer, M., Enoksson, F., Palavitsinis, N., Kastrantas, K., & Naeve, A. (2009). Learning Object Annotation for Agricultural Learning Repositories. In: Ninth IEEE International Conference on Advanced Learning Technologies, (pp.438-442), Riga. Latvia.

Efron, M. (2007). Metadata use in oai-compliant institutional repositories.Journal of Digital Information, 8(2).

Ehlers, U.D., Goertz, L., Hildebrandt, B., & Pawlowski, J.M. (2006). Quality in e-learning: use and dissemination of quality approaches in European e-learning: a study by the European Quality Observatory.

Elliott, K., & Sweeney, K. (2008). Quantifying the reuse of learning objects.Australasian Journal of Educational Technology, 24(2), 137-142.

El-Sherbini, M., & Klim, G. (2004). Metadata and cataloging practices. Electronic Library, The, 22(3), 238-248.

Eppler, M. (2003). Managing information quality: Increasing the value of information in knowledge-intensive products and processes. Berlin, Germany: Springer-Verlag.

EQO (2004), “The EQO model, v.1.2a”,European Quality Observatory, available at: www.eqo. info/files/EQO-Model-1.2a.pdf

Esichaikul, V., Lamnoi, S., & Bechter, C. (2011). Student Modelling in Adaptive E-Learning Systems. Knowledge Management & E-Learning: An International Journal (KM&EL), 3(3), 342-355.

ESS Quality Glossary 2010, Developed by Unit B1 "Quality; Classifications", Eurostat, 2010.

ETB (2002). “Recommended data model format to be used as a standard by national systems to include national/local resources in the EU Treasury Browser”, European Treasury Browser (ETB), D4.2, available at: www.eun.org/eun.org2/eun/en/etb/content.cfm?lang ¼en&ov ¼2348

Evans, J., & Lindsay, W. (2005). The management and control of quality (6th ed.). Mason, OH: Thomson.

Evans, M. F., & Thomas, S. (2007). Implementation of an integrated information management system at the National Library of Wales: a case study. Program: electronic library & information systems, 41(4), 325-337.

Even, A. & Kaiser, M. (2009). A Framework for Economics-Driven Assessment of Data Quality Decisions. In: Proceedings of the 15th Americas Conference on Information Systems (AMCIS)(Association for Information Systems, 2009)

Even, A. & Shankaranarayanan, G. (2007). Utility-driven assessment of data quality, The DATA BASE for Advances in Information Systems38(2) (2007) 75–93.

Fan, C. W., Huang, J. F., & Fang, R. J. (2008). Development of a learning system for a museum collection of print-related artifacts. Electronic Library, The, 26(2), 172-187.

Farance, F. (1999). Learning Technology Standards Committee Work Program and Process. Retrieved December 4, 2004, from: http://ltsc.ieee.org/meeting/199912/ doc/ltsc_wg--199912--farance.ppt

Farb, S. E., & Riggio, A. (2004). Medium or message? A new look at standards, structures, and schemata for managing electronic resources. Library Hi Tech, 22(2), 144-152.

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

Fischer, T., & Neuroth, H. (2000). SSG-FI–special subject gateways to high quality Internet resources for scientific users. Online Information Review, 24(1), 64-68.

Foti, S., & Ring, G. (2008). From ePortfolios to iPortfolios: The find, refine, design, and bind model. Canadian Journal of Learning and Technology/La revue canadienne de l’apprentissage et de la technologie, 34(3).

Foulonneau, M. (2007). Information redundancy across metadata collections.Information processing & management, 43(3), 740-751.

Francis, D. E., & Murphy, E. (2008). Instructional designers' conceptualisations of learning objects. Australasian Journal of Educational Technology, 24(5), 475-486.

Friesen, N. (2002), "Semantic interoperability, communities of practice and the Cancore learning object metadata profile",Alternative Paper Tracks, Proceedings of the 11th World Wide Web Conference (WWW 2002), Hawaii.

Friesen, N. (2004). International LOM Survey: Report (Draft). Retrieved from http://arizona.openrepository.com/arizona/bitstream/10150/106473/1/LOM_Survey_Report2.doc

Friesen, N. (2005). Interoperability and learning objects: An overview of e-learning standardization. Interdisciplinary Journal of E-Learning and Learning Objects, 1(1), 23-31.

Friesen, N., Roberts, A., & Fisher, S. (2002). CanCore: Metadata for learning objects. Canadian Journal of Learning and Technology/La revue canadienne de l'apprentissage et de la technologie, 28(3).

Fuhr, N., Hansen, P., Mabe, M., Micsik, A., & Sølvberg, I. (2001). Digital libraries: A generic classification and evaluation scheme. Lecture Notes in Computer Science, 2163(2001),187–199

Fuhr, N., Tsakonas, G., Aalberg, T. & Agosti, M. (2007), "Evaluation of digital libraries", International Journal of Digital Libraries, Vol. 8, pp. 21-38.

Fullerton, K., Greenberg, J., McClure, M., Rasmussen, E., & Stewart, D. (1999). A digital library for education: the PEN-DOR project. Electronic Library, The, 17(2), 75-82.

Gafni, R. (2009). Quality Metrics for PDA-based M-Learning Information Systems. Interdisciplinary Journal of E-Learning and Learning Objects, 5(1), 359-378.

Gallenseon, A., Heins, J. & Heins, T. (2002). Macromedia MX: Creating learning objects. [Macromedia White Paper]. Macromedia Inc. Retrieved 6/02/06 from http://download.macromedia.com/pub/elearning/objects/mx_creating_lo.pdf

García-Marco, F. J. (2011). Libraries in the digital ecology: reflections and trends. Electronic Library, The, 29(1), 105-120.

Garshol, L. M. (2004). Metadata? Thesauri? Taxonomies? Topic maps! Making sense of it all. Journal of information science, 30(4), 378-391.

Gatenby, J. (2002). Aiming at quality and coverage combined: blending physical and virtual union catalogues. Online information review, 26(5), 326-334.

Geisler, G., McArthur, D., & McClelland, M. (2002). Creating virtual collections in digital libraries: Benefits and implementation issues. Joint Conference on Digital Libraries, Portland, Oregon, 210-218.

Genoni, P. (2007). Towards a National Print Repository for Australia: Where from and Where to?. Australian Academic and Research Libraries, 38(2), 84.

GESTALT (1999), “D0401 courseware metadata design”,Getting Educational Systems Talking Across Leading-Edge Technologies (GESTALT), available at: www.fdgroup.co.uk/gestalt/

Gilliland-Swetland, A. J., Kafai, Y. B., & Landis, W. E. (2000). Application of Dublin Core metadata in the description of digital primary sources in elementary school classrooms. Journal of the American Society for Information Science,51(2), 193-201.

Gilliland-Swetland, A.J. (2000). Defining metadata. In M. Baca (Ed.), Introduction to metadata: Pathways to digital information. Los Angles, CA: Getty Information Institute. Available at: http://www.getty.edu/gri/standard/intrometadata/2_articles/index.htm.

Glover, E.J., Lawrence, S., Gordon, M.D., Birmingham, W.P. & Giles, C.L. (2001), “Web search – your way”,Communications of the ACM, Vol. 44 No. 12, pp. 97-102.

Glover, I., Xu, Z., & Hardaker, G. (2007). Online annotation–Research and practices. Computers & Education, 49(4), 1308-1320.

Goh, D. H. L., Chua, A., Khoo, D. A., Khoo, E. B. H., Mak, E. B. T., & Ng, M. W. M. (2006). A checklist for evaluating open source digital library software.Online Information Review, 30(4), 360-379.

Goh, D. H. L., Chua, A., Lee, C. S., & Razikin, K. (2009). Resource discovery through social tagging: a classification and content analytic approach. Online Information Review, 33(3), 568-583.

Golder, S. A., & Huberman, B. A. (2006). Usage patterns of collaborative tagging systems. Journal of information science, 32(2), 198-208.

Gonçalves, M. A., Fox, E. A., Watson, L. T., & Kipp, N. A. (2004). Streams, structures, spaces, scenarios, societies (5s): A formal model for digital libraries.ACM Transactions of the Information Systems, 22(2), 270–312.

Gonçalves, M. A., Moreira, B. L., Fox, E. A., & Watson, L. T. (2007). “What is a good digital library?”–A quality model for digital libraries. Information processing & management, 43(5), 1416-1437.

Gonzalez-Barbone, V., & Anido-Rifon, L. (2010). From SCORM to Common Cartridge: A step forward. Computers & Education, 54(1), 88-102.

González-Videgaray, M., Hernández-Zamora, G., & del-Río-Martínez, J. H. (2009). Learning objects in theory and practice: A vision from Mexican University teachers. Computers & Education, 53(4), 1330-1338.

Goodyear, P. (2005). Educational design and networked learning: Patterns, pattern languages and design practice. Australasian Journal of Educational Technology, 21(1), 82-101.

Govindasamy, T. (2001). Successful implementation of e-learning: Pedagogical considerations. The Internet and Higher Education, 4(3), 287-299.

Graham, R. A. (2001). Metadata harvesting. Library hi tech, 19(3), 290-295.

Granić, A., Mifsud, C., & Ćukušić, M. (2009). Design, implementation and validation of a Europe-wide pedagogical framework for e-Learning. Computers & Education, 53(4), 1052-1081.

Greenberg, J. & Robertson, W. (2002). Semantic web construction: an inquiry of authors' views on collaborative metadata generation, Proceedings of the International Conference on Dublin Core and Metadata for e-Communities 2002, 45–52

Greenberg, J. (2001). A quantitative categorical analysis of metadata elements in image-applicable metadata schemas. Journal of the American Society for Information Science and Technology, 52(11), 917-924.

Greenberg, J. (2002). Semantic web construction: An inquiry of authors' views on collaborative metadata generation. In Proceeding of the International Conference on Dublin Core and Metadata for e-Communities (pp. 45–52). Florence, Italy: Firenze University Press. Retrieved April 26, 2010, from http://dcpapers.dublincore.org/ojs/pubs/article/view/693/689

Greenberg, J. (2009). Theoretical considerations of lifecycle modeling: an analysis of the dryad repository demonstrating automatic metadata propagation, inheritance, and value system adoption. Cataloging & classification quarterly, 47(3-4), 380-402.

Greenberg, J., Pattuelli, M. C., Parsia, B., & Robertson, W. D. (2006). Author-generated Dublin Core metadata for web resources: a baseline study in an organization. Journal of Digital Information, 2(2).

Greenberg, J., Pattuelli, M.C., Parsia, B., & Davenport Robertson, W. (2001). "Author-generated Dublin Core Metadata for Web Resources: A Baseline Study in an Organization." Journal of Digital Information, Volume 2 Issue 2 Article No. 78, November 2001.

Greenberg, J., Spurgin, K., & Crystal, A. (2005). AMeGA (Automatic Metadata Generation Applications) Project. Final Report, University of North Carolina & Library of Congress. Retrieved from: http://www.loc.gov/catdir/bibcontrol/lc_amega_final_report.pdf

Greenberg, J., Sutton, S. and D, G. (2003), "Campbell Metadata: a fundamental component of the semantic web",Bulletin of the American Society for Information Science and Technology, Vol. 29 No. 4, pp. 16-18.

Greer, L. R. (2002). The learning matrix: Cataloging resources with rich metadata. Proceedings of the IEEE-CS Joint Conference on Digital Libraries(Portland, Oregon, JCDL 2002), 375

Griffiths, D., Blat, J., Garcia, R., Vogten, H., & Kwong, K. L. (2005). Learning design tools. In Learning Design (pp. 109-135). Springer Berlin Heidelberg.

Groenewald, R., & Breytenbach, A. (2011). The use of metadata and preservation methods for continuous access to digital data. Electronic Library, The, 29(2), 236-248.

Groenewald, R., & Breytenbach, A. (2011). The use of metadata and preservation methods for continuous access to digital data. Electronic Library, The, 29(2), 236-248.

Groenlund, A. (2004). “State of the Art in e-Gov Research – A Survey”, R.Traunmueller (Ed.), EGOV 2004, Lecture Notes on Computer Science 3183, 178-185, 2004.

Guenther, R. S. (2004). Using the Metadata Object Description Schema (MODS) for resource description: guidelines and applications. Library hi tech,22(1), 89-98.

Guinchard, C. (2002), “Dublin Core use in libraries: a survey”,OCLC Systems & Services, Vol. 18. No. 1, pp. 40-50.

Guy, M., Powell, A. & Day, M. (2004). Improving the quality of metadata in e-print archives. Ariadne, (no. 38). Retrieved 15 September, 2011 from http://www.ariadne.ac.uk/issue38/guy/

Haase, K. (2004). “Context for semantic metadata”. in Proceedings of the 12th ACM International Conference on Multimedia. 2004. New York, NY, United States: ACM

Hadjerrouit, S. (2010). Developing Web-Based Learning Resources in School Education: A User-Centered Approach. Interdisciplinary Journal of E-Learning and Learning Objects, 6(1), 115-135.

Hakala, J. (2004). Archiving the web: European experiences. Program: electronic library and information systems, 38(3), 176-183.

Halttunen, K., & Järvelin, K. (2005). Assessing learning outcomes in two information retrieval learning environments. Information processing & management, 41(4), 949-972.

Hamid, A. A. (2001). E-learning: is it the “e” or the learning that matters?. The Internet and Higher Education, 4(3), 311-316.

Han, H., Giles, C.L., Manavoglu, E., Zha, H., Zhang, Z. & Fox, E.A. (2003), “Automatic document metadata extraction using support vector machines”,Proceedings of the 3rd ACM/IEEE-CS Joint Conference on Digital Libraries. JCDL ’03, Houston, TX, USA, May 27-31,IEEE Computer Society, Piscataway, NJ, pp. 37-48

Han, J.Y. (2001). “A study on the strategy for oceanographic meta-database implementation”, Proceedings of the Korean Society for Information Management Conference, Vol. 8, pp. 117-24.

Han, X., Zhou, Q., & Yang, J. (2011). A Technical Mode for Sharing and Utilizing Open Educational Resources in Chinese Universities. Knowledge Management & E-Learning: An International Journal (KM&EL), 3(3), 356-374.

Han, Y. (2004). Digital content management: the search for a content management system. Library Hi Tech, 22(4), 355-365.

Hariri, N., & Norouzi, Y. (2011). Determining evaluation criteria for digital libraries' user interface: a review. Electronic Library, The, 29(5), 698-722.

Harman, K., & Koohang, A. (2005). Discussion board: A learning object. Interdisciplinary Journal of Knowledge and Learning Objects, 1,67-77. Retrieved August 19, 2006 from http://www.ijklo.org/Volume1/v1p067-077Harman.pdf

Harth, A. (2004). An integration site for semantic web metadata. Web Semantics: Science, Services and Agents on the World Wide Web, 1(2), 229-234.

Harvey, L. & Green, D. (1993). "Defining quality", Assessment and Evaluation in Higher Education 18, 9–34.

Harvey, R., & Thompson, D. (2010). Automating the appraisal of digital materials. Library Hi Tech, 28(2), 313-322.

Haslhofer, B., & Isaac. A. (2011). data.europeana.eu - The Europeana Linked Open Data Pilot. In Proceeding of the International Conference on Dublin Core and Metadata Applications (DC 2011)

Haßler, B., & Jackson, A. M. (2010). Bridging the bandwidth gap: Open educational resources and the digital divide. Learning Technologies, IEEE Transactions on, 3(2), 110-115.

Hawking, D., & Zobel, J. (2007). Does topic metadata help with Web search?.Journal of the American Society for Information Science and Technology, 58(5), 613-628.

He, D., Peng, Y., Mao, M., & Wu, D. (2010). Supporting information access in e-learning by integrating digital libraries and ontology. Online Information Review, 34(5), 704-728.

Heery, R. & Anderson, S. (2005). Digital repositories review. Bath, UK: UKOLN and Arts and Humanities Data Service. Retrieved 13 September, 2011 from http://www.jisc.ac.uk/uploaded_documents/digital-repositories-review-2005.pdf

Heinrich, B., & Klier, M. (2011). Assessing data currency—a probabilistic approach. Journal of Information Science, 37(1), 86-100.

HeldmanK., ManganoV., PMP:ProjectManagementProfessionalExamReviewGuide, Sybex,Wiley PublishingInc.,Indianapolis2009

Hemminger, B.M., Bolas, G., Carr, D., Jones, P., Schiff, D. and England, N. (2004), "Capturing content for virtual museums: from pieces to exhibitions",Proceedings of the 2004 Joint ACM/IEEE Conference on Digital libraries, Tuscon, AZ, June 7-11, p. 379.

Hense, A., & Quadt, F. (2011). Acquiring high quality research data. D-Lib Magazine, 17(1), 5.

Hernandez, N., Mothe, J., Ralalason, B., Ramamonjisoa, B., & Stolf, P. (2008). A Model to Represent the Facets of Learning Object. Interdisciplinary Journal of E-Learning and Learning Objects, 4(1), 65-82.

Hershkovitz, A., Azran, R., Hardof-Jaffe, S., & Nachmias, R. (2011). Types of online hierarchical repository structures. The Internet and Higher Education,14(2), 107-112.

Hewlett, J., & Walton, G. (2001). Assessing the quality of library and information services for United Kingdom health professionals and students: a comparison of the National Health Service and higher education approaches and the way forward. Performance Measurement and Metrics, 2(2), 81-95.

Hiddink, G. (2001). Solving reusability problems of online learning materials. Campus-Wide Information Systems, 18(4), 146-152.

Higginbotham, N., Albrecht, G., & Connor, L. (2001). Health social science: a transdisciplinary and complexity perspective| NOVA. The University of Newcastle's Digital Repository.

Higgins, S. (2008). The DCC curation lifecycle model. International Journal of Digital Curation, 3(1), 134-140.

Higgins, S., & Inglis, G. (2003). Implementing EAD: the experience of the NAHSTE project. Journal of the Society of Archivists, 24(2), 199-214.

Hill, L. L., Janee, G., Dolin, R., Frew, J., & Larsgaard, M. (1999). Collection metadata solutions for digital library applications. JASIS, 50(13), 1169-1181.

Hilligoss, B., & Rieh, S.Y. (2008). Developing a unifying framework of credibility assessment: Construct, heuristics, and interaction in context. Information Processing & Management, 44(4), 1467–1484.

Hillman, D., Dusshay, N. & Phipps, J. (2004), “Improving metadata quality: augmentation and recombination”,Proceedings of the International Conference on Dublin Core and Metadata Applications (DC-2004), Shanghai, China, 11-14 October.

Hockx-Yu, H. (2006). Digital preservation in the context of institutional repositories. Program: electronic library and information systems, 40(3), 232-243.

Hodgins, W. (Chair), (2008). Working group information, announcements & news. WG12: Learning Object Metadata. Retrieved from http://ltsc.ieee.org/wg12/

Hoermann, S., Hildebrandt, T., Rensing, C., & Steinmetz, R. (2005, June). ResourceCenter-A Digital Learning Object Repository with an Integrated Authoring Tool Set. In World Conference on Educational Multimedia, Hypermedia and Telecommunications (Vol. 2005, No. 1, pp. 3453-3460).

Holden, C., (2003). From Local Challenges to a Global Community: Learning Repositories and the Global Learning Repositories Summit. Academic ADL Co-Lab

Holley, R. (2004). Developing a digitisation framework for your organisation.Electronic Library, The, 22(6), 518-522.

Horstmann, W. (2009). Building Repository Networks with DRIVER: Joint Explorations with the IR Grid in China. In Symposium on Development of Library and Information Services 2009 (p. 116).

Horwood, L., Sullivan, S., Young, E., & Garner, J. (2004). OAI compliant institutional repositories and the role of library staff. Library Management, 25(4/5), 170-176.

Houghton, J. W. (2010). Economic Implications of Alternative Publishing Models: Self-archiving and Repositories. Liber Quarterly, 19(3/4), 275-292.

Howarth, L.C. (2003). Designing a common namespace for searching metadata-enabled knowledge repositories: An international perspective. Cataloging and Classification Quarterly, 37(1–2), 173–185.

Hsu, T. Y., Ke, H. R., & Yang, W. P. (2006). Unified knowledge-based content management for digital archives in museums. Electronic Library, The, 24(1), 38-50.

Hswe, P., Giarlo, M. J., Belden, M., Clair, K., Coughlin, D., & Klimczyk, L. (2012). Building a Community of Curatorial Practice at Penn State: A Case Study. Journal of Digital Information, 13(1).

Hughes, B. (2005). Metadata quality evaluation: Experience from the open language archives community. Digital Libraries: International Collaboration and Cross-Fertilization, 135–148.

Hummel, H., Manderveld, J., Tattersall, C. & Koper, R. (2004). Educational modelling language and learning design: New opportunities for instructional reusability and personalised learning. International Journal of Learning Technology, 1(1), 111-126.

Hunter, P., & Guy, M. (2004). Metadata for Harvesting: The Open Archives Initiative, and how to find things on the Web. Electronic Library, The, 22(2), 168-174.

IEEE LOM (2002). "Draft Standard for Learning Object Metadata", IEEE Learning Technology Standards Committee, IEEE 1484.12.1-2002, 15 July 2002.

IEEE LSTC. (2002). IEEE 1484.12.1-2002: Standard for Learning Object Metadata. Retrieved August 8, 2004 from http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf

IMS Global Learning Consortium, Inc. (2003). IMS learning design information model. Retrieved October 25, 2007, from http://www.imsglobal.org/learningdesign/ldv1p0/imsld_infov1p0.html

Injac, V. (1999). Gabriel: Gateway to Europe's national libraries. Glasnik Narodne biblioteke Srbije, 1(1), 231-238.

INSPIRE (2005), "Requirements for the definition of the INSPIRE implementing rules for metadata",Infrastructure for Spatial Information in Europe (INSPIRE), 24 September, available at: www.ec-gis.org/inspire/reports.cfm

International Organization for Standardization. (2005). Quality management systems—Fundamentals and vocabulary (ISO 9000: 2005).Geneva: ISO.

Isfandyari-Moghaddam, A., & Bayat, B. (2008). Digital libraries in the mirror of the literature: issues and considerations. Electronic Library, The, 26(6), 844-862.

Islam, M. S., Kunifuji, S., Miura, M., & Hayama, T. (2011). Adopting knowledge management in an e-learning system: Insights and views of KM and EL research scholars. Knowledge Management & E-Learning: An International Journal (KM&EL), 3(3), 375-398.

Ismail, J. (2001). The design of an e-learning system: Beyond the hype. the internet and higher education, 4(3), 329-336.

ISO DIS 11620 (2006), Information and Documentation – Library Performance Indicators, International Organization for Standardization, Geneva.

IVOA (2004), "Resource metadata for the virtual observatory, version 1.01",International Virtual Observatory Alliance (IVOA) Recommendation, available at: www.ivoa.net

Iwhiwhu, B. E., & Eyekpegha, E. O. (2009). Digitization of Nigerian university libraries: From technology challenge to effective information delivery. Electronic Library, The, 27(3), 529-536.

Jacobs, N., Thomas, A., & McGregor, A. (2008). Institutional Repositories in the UK: The JISC Approach. Library Trends, 57(2), 124-141.

Jacsó, P. (2010). Metadata mega mess in Google Scholar. Online Information Review, 34(1), 175-191.

Jacsó, P. (2010). Pragmatic issues in calculating and comparing the quantity and quality of research through rating and ranking of researchers based on peer reviews and bibliometric indicators from Web of Science, Scopus and Google Scholar. Online information review, 34(6), 972-982.

Jaén, X., Bohigas, X., & Novell, M. (2007). The need for virtual information managers in education. Computers & Education, 49(2), 254-268.

Jansen, B. J., Liu, Z., Weaver, C., Campbell, G., & Gregg, M. (2011). Real time search on the web: Queries, topics, and economic value. Information Processing & Management, 47(4), 491-506.

Jantz, R., & Giarlo, M. J. (2005). Digital preservation: Architecture and technology for trusted digital repositories. Microform & imaging review, 34(3), 135-147.

Jara, M., & Mellar, H. (2010). Quality enhancement for e-learning courses: The role of student feedback. Computers & Education, 54(3), 709-714.

Joho, H., & Jose, J. M. (2008). Effectiveness of additional representations for the search result presentation on the web. Information processing & management, 44(1), 226-241.

Jones, M. L., Gay, G. K., & Rieger, R. H. (1999). Project soup: comparing evaluations of digital collection efforts.

Jones, R. (2005). Designing adaptable learning resources with learning object patterns. Journal of Digital Information, 6(1).

Jones, R. A. (2008). A marathon not a sprint: Lessons learnt from the first decade of digitisation at the National Library of Wales. Program: electronic library and information systems, 42(2), 97-114.

Jones, R., & Andrew, T. (2005). Open access, open source and e-theses: the development of the Edinburgh Research Archive. Program: electronic library and information systems, 39(3), 198-212.

Jordan, M. (2006). The CARL metadata harvester and search service. Library hi tech, 24(2), 197-210.

Jörgensen, C. (1995). Image attributes: An investigation. Unpublished doctoral thesis, Syracuse University, NY.

Joyes, G., Gray, L., & Hartnell-Young, E. (2010). Effective practice with e-portfolios: How can the UK experience inform implementation. Australasian Journal of Educational Technology, 26(1), 15-27.

Jung, S., Herlocker, J. L., & Webster, J. (2007). Click data as implicit relevance feedback in web search. Information Processing & Management, 43(3), 791-807.

Juran, J.M. & Gryna, F.M. (1980). Quality Planning and Analysis, 2d ed. New York: McGraw-Hill, 1980.

Juran, J.M. (1989). Juran on Leadership for Quality: An Executive Handbook. New York: The Free Press, 1989.

Kahn R., & Wilensky, R. (2006). A Framework for Distributed Digital Object Services, Int'l J. Digital Libraries, vol. 6, pp. 115-123, 2006

Kastens, K., DeFelice, B., Devaul, H., DiLeonardo, C., Ginger, K., Larsen, S., (...), Tahirkheli, S. (2005). Questions & Challenges Arising in Building the Collection of a Digital Library for Education: Lessons from Five Years of DLESE. D-Lib Magazine, 11(11).

Kastens, K., Devaul, H., Ginger, K., Mogk, D., DeFelice, B., & DiLeonardo, C. (2005). Questions and challenges arising in building the collection of a digital library for education. D-Lib Magazine, 11(11), 1082-9873.

Kavulya, J. M. (2007). Digital libraries and development in Sub-Saharan Africa: A review of challenges and strategies. Electronic Library, The, 25(3), 299-315.

Kay, R. H., & Knaack, L. (2008). Exploring the impact of learning objects in middle school mathematics and science classrooms: A formative analysis.Canadian Journal of Learning and Technology/La revue canadienne de l'apprentissage et de la technologie, 34(1).

Kay, R. H., & Knaack, L. (2009). Assessing learning, quality and engagement in learning objects: the Learning Object Evaluation Scale for Students (LOES-S). Educational Technology Research and Development, 57(2), 147-168.

Kay, R., & Knaack, L. (2008). Investigating the use of learning objects for secondary school mathematics. Interdisciplinary Journal of E-Learning and Learning Objects, 4(1), 269-289.

Ke, H. R., & Hwang, M. J. (2000). The development of digital libraries in Taiwan.Electronic Library, The, 18(5), 336-347.

Kelly, B., Closier, A., & Hiom, D. (2005). 'Gateway Standardization: A Quality Assurance Framework for Metadata. Library Trends, 53(4), 637-650.

Kennedy, M. R. (2008). Nine questions to guide you in choosing a metadata schema. Journal of Digital Information, 9(1).

Khelifi, A., Talib, M. A., Farouk, M., & Hamam, H. (2009). Developing an initial open-source platform for the higher education sector—a case study: Alhosn University. Learning Technologies, IEEE Transactions on, 2(3), 239-248.

Kim, H. H., & Kim, Y. H. (2008). Usability study of digital institutional repositories. Electronic Library, The, 26(6), 863-881.

Kim, J. (2007). Motivating and impeding factors affecting faculty contribution to institutional repositories. Journal of Digital Information, 8(2).

Kim, K. S., & Sin, S. C. J. (2011). Selecting quality sources: Bridging the gap between the perception and use of information sources. Journal of Information Science, 37(2), 178-188.

Kim, Y. H., & Kim, H. H. (2008). Development and validation of evaluation indicators for a consortium of institutional repositories: A case study of dCollection. Journal of the American Society for Information Science and Technology, 59(8), 1282-1294.

Kirchberg, M., Schewe, K. D., Tretiakov, A., & Wang, B. R. (2007). A multi-level architecture for distributed object bases. Data & Knowledge Engineering,60(1), 150-184.

Kiryakov, A., Popov, B., Terziev, I., Manov, D., & Ognyanoff, D. (2004). Semantic annotation, indexing, and retrieval. Web Semantics: Science, Services and Agents on the World Wide Web, 2(1), 49-79.

Knight, S. (2005). Preservation metadata: National library of new zealand experience. Library trends, 54(1), 91-110.

Knowles, J. (2010). Collaboration nation: the building of the Welsh Repository Network. Program: electronic library and information systems, 44(2), 98-108.

Koch, T. (2000). Quality-controlled subject gateways: definitions, typologies, empirical overview. Online information review, 24(1), 24-34.

Koehler, W. (2004). Digital libraries, digital containers,"library patrons", and visions for the future. Electronic Library, The, 22(5), 401-407.

Koenig, J. J., & Mikeal, A. (2010). Creating complex repository collections, such as journals, with Manakin. Program: electronic library and information systems, 44(4), 393-402.

Koohang, A., Floyd, K., & Stewart, C. (2011). Design of an open source learning objects authoring tool–the LO creator. Interdisciplinary Journal of E-Learning and Learning Objects, 7(1), 111-123.

Koper, R. (2005). An introduction to learning design. In Learning design (pp. 3-20). Springer Berlin Heidelberg.

Koper, R., & Tattersall, C. (Eds.). (2005). Learning design: A handbook on modelling and delivering networked education and training. Springer.

Kopp, G., & Crichton, S. (2007). Embedded or linked learning objects: Implications for content development, course design and classroom use.Canadian Journal of Learning and Technology/La revue canadienne de l'apprentissage et de la technologie, 33(2).

Korres, M. P., & García-Barriocanal, E. (2008). Development of personalized learning objects for training adult educators of special groups. Journal of Knowledge Management, 12(6), 89-101.

Kounoudes, A. D., Artemi, P., & Zervas, M. (2010). Ktisis: Building an open access institutional and cultural repository. In Digital Heritage (pp. 504-512). Springer Berlin Heidelberg.

Koutsomitropoulos, D. A., Solomou, G. D., Papatheodorou, T. S., & Alexopoulos, A. D. (2010). The use of metadata for educational resources in digital repositories: Practices and perspectives. D-Lib Magazine, 16(1), 3.

Kovacevic, A., Devedzic, V., & Pocajt, V. (2010). Enhancing a core journal collection for digital libraries. Program: electronic library and information systems, 44(2), 132-148.

Kovacevic, A., Ivanovic, D., Milosavljevic, B., Konjovic, Z., & Surla, D. (2011). Automatic extraction of metadata from scientific publications for CRIS systems.Program: electronic library and information systems, 45(4), 376-396.

Kraan, W. (2003). "No one standard will suit all", The Centre for Educational Technology Interoperability Standards, Available http://www.cetis.ac.uk/content/20030513175232

Krauss, F., & Ally, M. (2005). A study of the design and evaluation of a learning object and implications for content development. Interdisciplinary Journal of E-Learning and Learning Objects, 1(1), 1-22.

Krishnamurthy, M. (2008). Open access, open source and digital libraries: A current trend in university libraries around the world. Program: electronic library & information systems, 42(1), 48-55.

Krishnamurthy, M., & Kemparaju, T. D. (2011). Institutional repositories in Indian universities and research institutes: a study. Program: electronic library & information systems, 45(2), 185-198.

Kruchten, P., Lago, P., & Van Vliet, H. (2006). Building up and reasoning about architectural knowledge. In Quality of Software Architectures (pp. 43-58). Springer Berlin Heidelberg.

Külcü, Ö., & Külcü, H. U. (2010). Contextual analysis of the e-records management requirements of Turkish Red Crescent Society. Electronic Library, The, 28(2), 314-333.

Kumar, V., Nesbit, J., & Han, K. (2005, July). Rating learning object quality with distributed bayesian belief networks: The why and the how. In Advanced Learning Technologies, 2005. ICALT 2005. Fifth IEEE International Conference on (pp. 685-687). IEEE.

Kunze, T., Brase, J., & Nejdl, W. (2002). Editing learning object metadata: Schema driven input of RDF metadata with the OLR3-Editor. Semantic

authoring, annotation & knowledge markup workshop. Proceedings of the 15th European Conf. on Artificial Intelligence, Lyon, France, July, 22-26.

Kurth, M., Ruddy, D., & Rupp, N. (2004). Repurposing MARC metadata: using digital project experience to develop a metadata management design. Library Hi Tech, 22(2), 153-165.

Kyrillidou, M., & Cook, C. (2008). The evolution of measurement and evaluation of libraries: a perspective from the Association of Research Libraries. Library Trends, 56(4), 888-909.

Lagoze, C. (2001). Keeping Dublin Core simple: cross-domain discovery or resource description?D-Lib Magazine, 7(1). Available from http://www.dlib.org/dlib/january01/lagoze/01lagoze.html. Accessed 24 June 2004.

Lagoze, C., & Van de Sompel, H. (2003). The making of the open archives initiative protocol for metadata harvesting. Library hi tech, 21(2), 118-128.

Lagoze, C., Lynch, C.A., & Daniel, R. (1996). The Warwick Framework: A container architecture for aggregating sets of metadata. Available at: http://cs-tr.cs.cornell.edu:80/Dients/Repository/2.0/Body/ncstrl.cornell%2fTR96-1593/html

L'Allier, J.J. (1997). Frame of Reference: NETg's Map to the Products, Their Structure and Core Beliefs. NetG. http://www.netg.com/research/whitepapers/frameref.asp

Lambert, F. (2010). Online community information seeking: The queries of three communities in Southwestern Ontario. Information processing & management,46(3), 343-361.

Lampert, C. K., & Chung, S. K. (2011). Strategic Planning for Sustaining User-Generated Content in Digital Collections. Journal of Library Innovation, 2(2), 74.

Leacock, T. L., & Nesbit, J. C. (2007). A framework for evaluating the quality of multimedia learning resources. Educational Technology & Society, 10(2), 44-59.

Lee, S. (2001), Digital Imaging: A Practical Handbook, Neal-Schuman, New York, NY.

Lee, Y.W., Strong, D.M., Kahn, B.K., & Wang, R.Y. (2002). AIMQ: A methodology for information quality assessment. Information & Management, 40(2), 133–146.

Lefoe, G., Philip, R., O'Reilly, M., & Parrish, D. (2009). Sharing quality resources for teaching and learning: A peer review model for the ALTC Exchange in Australia. Australasian Journal of Educational Technology, 25(1), 45-59.

LeFurgy, W. G. (2002). Levels of service for digital repositories. D-Lib Magazine, 8(5).

Lehmann, L., Hildebrandt, T., Rensing, C., & Steinmetz, R. (2007). Capturing, management and utilization of lifecycle information for learning resources. InCreating New Learning Experiences on a Global Scale (pp. 187-201). Springer Berlin Heidelberg.

Leinonen, T., Purma, J., Poldoja, H., & Toikkanen, T. (2010). Information architecture and design solutions scaffolding authoring of open educational resources. Learning Technologies, IEEE Transactions on, 3(2), 116-128.

Lesk, M. (1997). Practical Digital Libraries: Books, Bytes, and Bucks. San Francisco, CA: Morgan Kaufmann.

Levi, P. (2010). Digitising the past: The beginning of a new future at the Royal Tropical Institute of The Netherlands. Program: electronic library and information systems, 44(1), 39-47.

Li, L. (2006). Leveraging quality web-based library user services in the digital age. Library management, 27(6/7), 390-400.

Li, Q., Cheng, T., Wang, Y., & Bryant, S.H. (2010). PubChem as a public resource for drug discovery. Drug Discovery Today, 15(23–24), 1052–1057. doi: 10.1016/j.drudis.2010.10.003

Liao, C., To, P. L., Liu, C. C., Kuo, P. Y., & Chuang, S. H. (2011). Factors influencing the intended use of web portals. Online Information Review, 35(2), 237-254.

LibraryThing, LTFL: libraries using LibraryThing for libraries (2010). Available at: www.librarything.com/wiki/index.php/LTFL:Libraries_using_LibraryThing_for_Libraries (accessed 28 July 2010).

Liddy, E., Allen, E., Harwell, S., Corieri, S., Yilmazel, O., Ozgencil, N.,... (2002). Automatic metadata generation & evaluation. In Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 401–402). New York: ACM Press.

Liddy, E.D., Allen, E.E., & Finneran, C.M. (2003). MetaTest: Evaluation of metadata from generation to use. Paper presented at the May 27–31, 2003 Joint Conference on Digital Libraries (JCDL'03). Houston, TX, USA.

Lindner, R. (2001). CEN/ISSS/WS-LT Presentation. In Learning Technology Roundtable Panel Discussion.March 20, 2001, New York City, USA. Retrieved August 6, 2004 from http://jtc1sc36.org/doc/36N0085.pdf

Littlejohn, A., Falconer, I., & Mcgill, L. (2008). Characterising effective eLearning resources. Computers & Education, 50(3), 757-771.

Liu, J. H., Huang, B. S., & Chao, M. (2005, July). The design of learning object authoring tool based on SCORM. In Advanced Learning Technologies, 2005. ICALT 2005. Fifth IEEE International Conference on (pp. 778-782). IEEE.

Liu, S., & Zhou, Y. (2011). Developing an institutional repository using DigiTool. Electronic Library, The, 29(5), 589-608.

Liu, Y. Q. (2004). Best practices, standards and techniques for digitizing library materials: a snapshot of library digitization practices in the USA. Online Information Review, 28(5), 338-345.

Loewald, T., & DeRidder, J. (2010). Metadata In, Library Out. A Simple, Robust Digital Library System. Code4Lib Journal, 10.

Londhe, N. L., Desale, S. K., & Patil, S. K. (2011). Development of a digital library of manuscripts: A case study at the University of Pune, India. Program: electronic library and information systems, 45(2), 135-148.

Lopatin, L. (2006). Library digitization projects, issues and guidelines: A survey of the literature. Library hi tech, 24(2), 273-289.

Lopatin, L. (2010). Metadata practices in academic and non-academic libraries for digital projects: A survey. Cataloging & Classification Quarterly, 48(8), 716-742.

López, C. (2005). Los repositorios de objetos de aprendizaje como soporte a un entorno e-learning. [Learning objects repositories as platform of e-learning environment.]Dissertation. University of Salamanca. (Director: Francisco José García Peñalvo).

Losee, R. M. (2006). Browsing mixed structured and unstructured data.Information processing & management, 42(2), 440-452.

Lubas, R. L., Wolfe, R. H., & Fleischman, M. (2004). Creating metadata practices for MIT's OpenCourseWare Project. Library Hi Tech, 22(2), 138-143.

Lynch, C. (1995), "CNI white paper on networked information discovery and retrieval", available at: www.cni.org/projects/nidr/

Lynch, C. (2003). "Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age." ARL Bimonthly Report, No. 226, February 2003. http://www.arl.org/newsltr/226/ir.html (retrieved 18/7/2013)

Lyon, L. (2006), "E-science and the role of libraries",8th International Bielefeld Conference 2006, Bielefeld, 7-9 February 2006, available at: conference.ub.uni-bielefeld.de/2006/docs/docs_main.htm (accessed 30 April 2008).

MacDonald, G.F. (2000), "Digital visionary",Museum News, March/April.

Maceviciute, E., & Wilson, T. (2008). Evaluation of learning objects from the user's perspective: The case of the EURIDICE Service. Interdisciplinary journal of e-learning and learning objects, 4(1), 83-95.

MacKinnon, G., & Aylward, M. L. (2009). Models for Building Knowledge in a Technology-Rich Setting: Teacher Education. Canadian Journal of Learning and Technology/La revue canadienne de l'apprentissage et de la technologie, 35(1).

MacLean, P., & Scott, B. (2011). Competencies for learning design: A review of the literature and a proposed framework. British Journal of Educational Technology, 42(4), 557-572.

Madnick, S., & Zhu, H. (2006). Improving data quality through effective use of data semantics. Data & Knowledge Engineering, 59(2), 460-475.

Making of America II. (1998). Testbed project white paper, version 2.0. Available at: http://sunsite.berkeley.edu/moa2/wp-v2.html.

Makris, K., Skevakis, G., Kalokyri, V., Gioldasis, N., Kazasis, F. G., & Christodoulakis, S. (2011). Bringing Environmental Culture Content into the Europeana.eu Portal: The Natural Europe Digital Libraries Federation Infrastructure. Metadata and Semantic Research. Communications in Computer and Information Science, 240, 400-411.

Malaxa, V., & Douglas, I. (2005). A framework for metadata creation tools. Interdisciplinary Journal of E-Learning and Learning Objects, 1(1), 151-162.

Maldonado, U. P. T., Khan, G. F., Moon, J., & Rho, J. J. (2011). E-learning motivation and educational portal acceptance in developing countries. Online Information Review, 35(1), 66-85.

Malliari, A., Moreleli-Cacouris, M., & Kapsalis, K. (2010). Usage patterns in a Greek academic library catalogue: a follow-up study. Performance Measurement and Metrics, 11(1), 47-55.

Manouselis, N., & Sampson, D. G. (2004a). Multiple Dimensions of User Satisfaction as Quality Criteria for Web Portals. In ICWI (pp. 535-542).

Manouselis, N., & Sampson, D. (2004b). Recommendation of quality approaches for the European Quality Observatory. In Advanced Learning Technologies, 2004. Proceedings. IEEE International Conference on (pp. 1082-1083). IEEE.

Manouselis, N., & Costopoulou, C. (2006). Quality in metadata: a schema for e-commerce. Online Information Review, 30(3), 217-237.

Manouselis, N., Kastrantas, K., Sanchez-Alonso, S., Cáceres, J., Ebner, H., & Palmer, M. (2009). Architecture of the organic. edunet web portal. International Journal of Web Portals (IJWP), 1(1), 71-91.

Marcial, L. H., & Hemminger, B. M. (2010). Scientific data repositories on the Web: An initial survey. Journal of the American Society for Information Science and Technology, 61(10), 2029-2048.

Margaritopoulos, M., Margaritopoulos, T., Kotini, I., & Manitsaris, A. (2008). Automatic metadata generation by utilising pre-existing metadata of related resources. International Journal of Metadata, Semantics and Ontologies, 3(4), 292–304.

Margaritopoulos, M., Margaritopoulos, T., Mavridis, I., & Manitsaris, A. (2012). Quantifying and measuring metadata completeness. Journal of the American Society for Information Science and Technology, 63(4), 724-737.

Margaritopoulos, T., Margaritopoulos, M., Mavridis, I., & Manitsaris, A. (2008). A conceptual framework for metadata quality assessment. In Proceedings of the International Conference on Dublin Core and Metadata Applications, Berlin, Germany, 2008

Marieke, G., Powell, A., and Day, M. (2004). “Improving the Quality of Metadata in Eprint Archives”, Ariadne, Issue 38. Online: http://www.ariadne.ac.uk/issue38/guy/

Markey, K. (1984). Subject Searching in Library Catalogs, OCLC, Dublin, OH.

Markland, M. (2005). Does the student's love of the search engine mean that high quality online academic resources are being missed?. Performance measurement and metrics, 6(1), 19-31.

Marschak, J. (1971). Economics of information systems. Journal of the American Statistical Association, 66(333), 192–219.

Marshall, C. C. (1998). Making metadata: A study of metadata creation for a mixed physical-digital collection. Proceedings of the 3rd ACM Conference on Digital Libraries, (Pittsburgh, Pennsylvania, June, 1998) 162-171.

Marshall, C.C., & Shipman, F.M. (2003). Which semantic web? In Proceedings of the fourteenth ACM conference on Hypertext and hypermedia. 2003. Nottingham, UK: ACM Press

Marshall, S. (2012). Improving the quality of e-learning: lessons from the eMM. Journal of Computer Assisted Learning, 28(1), 65-78.

Martinez-Ortiz, I., Sierra, J. L., & Fernandez-Manjon, B. (2009). Authoring and reengineering of IMS learning design units of learning. Learning Technologies, IEEE Transactions on, 2(3), 189-202.

Marty, P. and Twidale, M. (2000). “Unexpected Help with Your Web-based Collections: Encouraging Data Quality Feedback from your Museum Visitors”, Museums and the Web 2002 Papers. Online: http://www.archimuse.com/mw2000/papers/marty/marty.html

Masoumi, D., & Lindström, B. (2012). Quality in e-learning: a framework for promoting and assuring quality in virtual institutions. Journal of Computer Assisted Learning, 28(1), 27-41.

Massart, D., Shulman, E., Nicholas, N., Ward, N., & Bergeron, F. (2010). Taming the metadata beast: ILOX. D-lib Magazine, 16(11), 1.

McCallum, S. H. (2004). An introduction to the metadata object description schema (MODS). Library Hi Tech, 22(1), 82-88.

McClave, J., & Benson, P. (1992). A first course in business statistics (5th ed.). New York, NY: Macmillan

McClelland, M., McArthur, D., Giersch, S., & Geisler, G. (2002). Challenges for service providers when importing metadata in digital libraries. D-Lib Magazine, 8(4), 1082-9873.

McCord, S. K., & Nofsinger, M. M. (2002). Continuous assessment at Washington State University Libraries: a case study. Performance Measurement and Metrics, 3(2), 68-73.

McCutcheon, S., Kreyche, M., Maurer, M. B., & Nickerson, J. (2008). Morphing metadata: maximizing access to electronic theses and dissertations. Library Hi Tech, 26(1), 41-57.

McDonald, J., & Bolland, K. (2003). Collaboration and resource sharing in and from the South Pacific: The National Library of New Zealand. Interlending & document supply, 31(2), 117-122.

McDonnell, M., & Shiri, A. (2011). Social search: A taxonomy of, and a user-centred approach to, social web search. Program: electronic library and information systems, 45(1), 6-28.

McDowell, C. (2007). Evaluating institutional repository deployment in American academe since early 2005: Repositories by the numbers, part 2. D-lib Magazine, 13(9), 3.

McGilvray, D., 2008. Executing Data Quality Projects: Ten Steps to Quality Data and Trusted Information, Morgan Kaufmann.

McGorry, S. Y. (2003). Measuring quality in online programs. The Internet and Higher Education, 6(2), 159-177.

McGreal, R. (2002). eduSource: A pan–Canadian learning object repository. In World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education (Vol. 2002, No. 1, pp. 32-40).

McGreal, R. (2004). Learning objects: A practical definition. International Journal of Instructional Technology and Distance Learning, 1(9). Retrieved September 15, 2011 from http://www.itdl.org/Journal/Sep_04/article02.htm

McMartin, F. (2006). MERLOT: a model for user involvement in digital library design and implementation. Journal of Digital Information, 5(3).

Medawar, K. (1995). "Database Quality: A Literature Review of the Past and a Plan for the Future", Program, vol. 29, no. 3, 257-272.

MedCIRCLE Workshop (2002), "Towards a collaborative, open, semantic web of trust for health information on the web: interoperability of health information gateways", MedCIRCLE: Collaboration for Internet Rating, Certification, Labeling and Evaluation of Health Information, Brussels, 12 September.

Mehrjerdi, Y. Z., Toranlo, H. S., & Jamali, R. (2009). Measuring academic libraries service quality in fuzzy environment. Performance Measurement and Metrics, 10(2), 94-115.

Mei, Q. (2004). "A knowledge processing oriented life cycle study from a digital museum system", Proceedings of the 42nd Annual Southeast Regional Conference, Huntsville, AL, April 2-3, pp. 116-21

Meier, C., Seufert, S., & Euler, D. (2012). Quality assessment and development in the course of the EFMD CEL programme accreditation. Journal of Computer Assisted Learning, 28(1), 52-64.

Melike, S., & Wade, V. (2011). Automatic Metadata Mining from Multilingual Enterprise Content. Web Semantics: Science, Services and Agents on the World Wide Web, 11.

Meng-xing, H., Chun-xiao, X., & Yong, Z. (2010). Supply chain management model for digital libraries. Electronic Library, The, 28(1), 29-37.

Mercun, T., & Žumer, M. (2008). New generation of catalogues for the new generation of users: a comparison of six library catalogues. Program: electronic library and information systems, 42(3), 243-261.

Merk, C., Scholze, F., & Windisch, N. (2009). Item-level usage statistics: a review of current practices and recommendations for normalization and exchange. Library Hi Tech, 27(1), 151-162.

Merriam-Webster's Online Dictionary(2009), "Quality entry", available at: http://mw1.merriamwebster.com/dictionary/quality (accessed May 28, 2009).

Merrill, M. D. (2000). Knowledge objects and mental models. In D. A. Wiley (Ed.), The instructional use of learning objects. Retrieved August 21, 2007 from http://reusability.org/read/chapters/merrill.doc

Mikhail, Y., Adly, N., & Nagi, M. (2012). DAR: A Modern Institutional Repository with a Scalability Twist. Journal of Digital Information, 13(1).

Miller, P. (1996). "Metadata for the masses",Ariadne, Vol. 5.

Milosevic, D., Brkovic, M., Debevc, M., & Krneta, R. (2007). Adaptive learning by using scos metadata. Interdisciplinary Journal of E-Learning and Learning Objects, 3(1), 163-174.

Mittal, R., & Mahesh, G. (2008). Digital libraries and repositories in India: an evaluative study. Program: electronic library and information systems, 42(3), 286-302.

Moen, W. E., Stewart, E. L. & McClure, C. R. (1997). The Role of Content Analysis in Evaluating Metadata for the US Government Information Locator Service (GILS): results from an exploratory study http://www.unt.edu/wmoen/publications/GILSMDContentAnalysis.htm

Moen, W.E., & McClure, C.R. (1997). An evaluation of the federal government's implementation of the Government Information Locator Service (GILS): Final report. Retrieved September 18, 2007 from http://www.gpoaccess.gov/gils/gils-eval.pdf

Moen, W.E., Stewart, E., & McClure, C.R. (1997). The role of content analysis in evaluating metadata for the U.S. Government Information Locator Service (GILS): Results from an exploratory study. Retrieved September 18, from http://www.unt.edu/wmoen/publications/GILSMDContentAnalysis.htm

Moen, W.E., Stewart, E., & McClure, C.R. (1998). Assessing metadata quality: Findings and methodological considerations from an evaluation of the U.S. Government Information Locator Service (GILS). In Proceedings of the IEEE International Forum on Research and Technology. Advances in Digital Libraries (pp. 246–255). Washington DC: IEEE Computer Society.

Moghaddam, G. G., & Moballeghi, M. (2008). Total quality management in library and information sectors. Electronic Library, The, 26(6), 912-922.

Mogharreban, N., & Guggenheim, D. (2008). Learning Pod: A New Paradigm for Reusability of Learning Objects. Interdisciplinary Journal of E-Learning and Learning Objects, 4(1), 303-315.

Mohamed, K. A. (2006). The impact of metadata in web resources discovering.Online Information Review, 30(2), 155-167.

Mohan, P. (2004). Building an online course based on the e-Learning standards: Guidelines, issues, and challenges. Canadian Journal of Learning and Technology/La revue canadienne de l'apprentissage et de la technologie,30(3).

Monge, S., Ovelar, R., & Azpeitia, I. (2008). Repository 2.0: Social dynamics to support community building in learning object repositories. Interdisciplinary Journal of E-Learning and Learning Objects, 4(1), 191-204.

Moraga, Á., Calero, C., & Piattini, M. (2006). Comparing different quality models for portals. Online Information Review, 30(5), 555-568.

Morgan, J., & Bawden, D. (2006). Teaching knowledge organization: educator, employer and professional association perspectives. Journal of Information Science, 32(2), 108-115.

Morris, E. (2005). Object oriented learning objects. Australasian Journal of Educational Technology, 21(1), 40-59.

Moyle, M. (2008). Improving access to European e-theses: the DART-Europe Programme. Liber Quarterly, 18(3/4), 413-423.

Muhonen, A., Saarti, J., & Vattulainen, P. (2010). Managing the life cycles of the document and library collections in Finnish academic libraries: Two case studies: Aalto and UEF. Library Management, 31(8/9), 669-677.

Mustaro, P., & Silveira, I. (2006). Learning objects: Adaptive retrieval through learning styles. Interdisciplinary Journal of E-Learning and Learning Objects, 2(1), 35-46.

Muzio, J. A., Heins, T., & Mundell, R. (2002). Experiences with reusable E-learning objects: From theory to practice. The Internet and Higher Education,5(1), 21-34.

Nagata, H., Satoh, Y., Gerrard, S., & Kytömäki, P. (2004). The dimensions that construct the evaluation of service quality in academic libraries. Performance Measurement and Metrics, 5(2), 53-65.

Najjar, J. & Duval, E. (2006), "Actual use of learning objects and metadata: an empirical analysis", IEEE Technical Committee on Digital Libraries (TCDL) Bulletin, Vol. 2 No. 2, available at: www.ieee-tcdl.org/Bulletin/v2n2/najjar/najjar.html (accessed 1 February 2010).

Najjar, J., Ternier, S., & Duval, E. (2003). The Actual Use of Metadata in ARIADNE: An Empirical Analysis. ARIADNE 3rd Conference, 2003. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.93.3666&rep=rep1&type=pdf

Nakabayashi, K., Morimoto, Y., & Hada, Y. (2010). Design and implementation of an extensible learner-adaptive

environment. Knowledge Management & E-Learning: An International Journal (KM&EL), 2(3), 246-259.

Namuth, D., Fritz, S., King, J., & Boren, A. (2005). Principles of sustainable learning object libraries. Interdisciplinary Journal of E-Learning and Learning Objects, 1(1), 181-196.

Nash, S. (2005). Learning objects, learning object repositories, and learning theory: Preliminary best practices for online courses. Interdisciplinary Journal of E-Learning and Learning Objects, 1(1), 217-228.

National Archives, US national archives records on footnote(2010). Available at: http://go.footnote.com/nara/ (accessed 28 July 2010).

NDN (2004), "Quality metadata", paper presented at the National Data Network (NDN) Metadata Workshop, November 8, available at: www.nationaldatanetwork.org/

Nesbit, J., Belfer, K., & Leacock, T. (2003). Learning object review instrument (LORI). InE-Learning Research and assessment network.

Nesbit, J., Belfer, K., & Vargo, J. (2002). A convergent participation model for evaluation of learning objects. Canadian Journal of Learning and Technology/La revue canadienne de l'apprentissage et de la technologie, 28(3).

Névéol, A., Rogozan, A., & Darmoni, S. (2006). Automatic indexing of online health resources for a French quality controlled gateway. Information processing & management, 42(3), 695-709.

Nikisch, J. A., & Górny, M. (2005). Regional digital libraries in Poland.Electronic Library, The, 23(4), 474-479.

Nikolaidou, M., Anagnostopoulos, D., & Hatzopoulos, M. (2005). Development of a medical digital library managing multiple collections. Electronic Library, The, 23(2), 221-236.

NISO (2004), Understanding Metadata, National Information Standards Organisation, NISO Press, Ann Arbor, MI.

Noh, Y. (2011). A study on metadata elements for web-based reference resources system developed through usability testing. Library Hi Tech, 29(2), 242-265.

Novak, E., Razzouk, R., & Johnson, T. E. (2012). The educational use of social annotation tools in higher education: A literature review. The Internet and Higher Education, 15(1), 39-49.

Ochoa, X. & Duval, E. (2009b), "Quantitative analysis of learning object repositories",IEEE Transactions on Learning Technologies, Vol. 2 No. 3, pp. 226-38.

Ochoa, X. (2011). Modeling the macro-behavior of learning object repositories. Interdisciplinary Journal of E-Learning and Learning Objects, 7(1), 25-35.

Ochoa, X., & Duval, E. (2008). Relevance Ranking Metrics for learning objects.Learning Technologies, IEEE Transactions on, 1(1), 34-48.

Ochoa, X., & Duval, E. (2009). Automatic evaluation of metadata quality in digital libraries. International Journal on Digital Libraries, 10(2–3), 67–91.

Ochoa, X., & Duval, E. (2009). Quantitative analysis of learning object repositories. Learning Technologies, IEEE Transactions on, 2(3), 226-238.

Ochoa, X., Cardinaels, K., Meire, M., & Duval, E. (2005, June). Frameworks for the automatic indexation of learning management systems content into learning object repositories. Paper presented at the ED-MEDIA 2005 World Conference on Educational Multimedia, Hypermedia and Telecommunications, Montreal, Canada.

Ochoa, X., Klerkx, J., Vandeputte, B., & Duval E. (2011). On the use of Learning Object Metadata: the GLOBE experience. Lecture Notes in Computer Science, 6964, 271-284.

O'Connor, S. (2005). The economics of repository libraries in the context of the future conventional libraries. Library Management, 26(1/2), 18-25.

Oduwole, A. A., & Sowole, A. O. (2006). Utilisation and impact of The Essential Electronic Agricultural Database (TEEAL) on library services in a Nigerian university of agriculture. Program: electronic library and information systems,40(2), 157-167.

Ojokoh, B. A., Adewale, O. S., & Falaki, S. O. (2009). Automated document metadata extraction. Journal of Information Science, 35(5), 563-570.

Olimpo, G., Chioccariello, A., Tavella, M., & Trentin, G. (1990). On the concept of reusability in education design. Learning technology in the European communities. Proceedings of the Delta Conference on Research and Development – The Hague,535-548.

Ossiannilsson, E., & Landgren, L. (2012). Quality in e-learning–a conceptual framework based on experiences from three international benchmarking projects. Journal of Computer assisted learning, 28(1), 42-51.

Osswald, A. (2008). E-science and information services: a missing link in the context of digital libraries. Online Information Review, 32(4), 516-523.

Otto, B., Hüner, K.M. & Österle, H. (2009). Identification of business oriented data quality metrics. In: Proceedings of the 14th International Conference on Information Quality (ICIQ) (HPI Potsdam, 2009).

Ouyang, Y., & Zhu, M. (2008). eLORM: learning object relationship mining-based repository. Online Information Review, 32(2), 254-265.

Özel, S. A., Altingövde, I. S., Ulusoy, Ö., Özsoyoğlu, G., & Özsoyoğlu, Z. M. (2004). Metadata-based modeling of information resources on the Web. Journal of the American Society for Information Science and Technology, 55(2), 97-110.

Pages, C., Sicilia, M.A., Garcıa, E., Martınez, J.J. & Gutierrez, J.M. (2003), “On the evaluation of completeness of learning object metadata in open repositories”,Proceedings of the Second International Conference on

Multimedia and Information & Communication Technologies in Education (m-ICTE 2003), pp. 1760-1764.

Palavitsinis, N., & Manouselis, N. (2009). “A Survey of Knowledge Organization Systems in Environmental Sciences”, in I.N. Athanasiadis, P.A. Mitkas, A.E. Rizzoli & J. Marx-Gómez (eds.), Information Technologies in Environmental Engineering,Proceedings of the 4th International ICSC Symposium, Springer Berlin Heidelberg.

Palavitsinis, N., & Manouselis, N. (2013). “Agricultural Knowledge Organisation Systems: An Analysis of an Indicative Sample” in Sicilia M.-A. (Ed.), Handbook of Metadata, Semantics and Ontologies, World Scientific Publishing Co. (in press).

Palavitsinis, N., Ebner, H., Manouselis, N. & Sanchez-Alonso, S. (2010). “Using e-Learning Technologies and Standardsto Make Educational Content Available: The Organic.Edunet Approach”, in Proc. of the 3rd Information and Communication Technologies in Agriculture, Food, Forestry and Environment Conference (ITAFFE'10), Samsun, Turkey, June 2010

Palavitsinis, N., Kastrantas, K., & Manouselis, N., (2009a). Interoperable metadata for a federation of learning repositories on organic agriculture and agroecology, in Proc. of the Joint International Agricultural Conference 2009 (JIAC 2009), Wageningen, The Netherlands, July 2009

Palavitsinis, N., Ebner, H., Manouselis, N., Sanchez S. & Naeve, A., (2009b). "Evaluating Metadata Application Profiles Based on Usage Data: The Case of a Metadata Application Profile for Agricultural Learning Resources", in Proc. of the International Conference on Digital Libraries and the Semantic Web (ICSD 2009), Trento, Italy, September 2009

Palavitsinis, N., Manouselis, N. & Sanchez, S., (2009c). Evaluation of a Metadata Application Profile for Learning Resources on Organic Agriculture, in Proc. of 3rd International Conference on Metadata and Semantics Research (MTSR09), Milan, Italy, October 2009

Palavitsinis, N., Manouselis, N. & Sanchez, S., (2011). Metadata quality in learning repositories: Issues and considerations, in Proc. of the World Conference on Educational Multimedia, Hypermedia & Telecommunications (ED-MEDIA 2011), Lisbon, Portugal

Palavitsinis, N., Manouselis, N. & Sanchez, S., (in press). Metadata Quality in Learning Object Repositories: A Case Study. The Electronic Library.

Palavitsinis, N., Manouselis, N. & Sanchez, S., (in press-b). Metadata Quality in Digital Repositories: Empirical Results from the Cross-Domain Transfer of a Quality Assurance Process. Journal of the American Society for Information Science and Technology

Palavitsinis, N., Manouselis, N. & Karagiannidis, C. (2013). Applying the Nominal Group Technique for Metadata Training of Domain Experts, in Proc. of 7th International Conference on Metadata and Semantics Research (MTSR13), Thessaloniki, Greece, November 2013

Palmer, C. L., Teffeau, L. C., & Newton, M. P. (2008). Strategies for institutional repository development: a case study of three evolving initiatives.Library Trends, 57(2), 142-167.

Papotti, P., & Torlone, R. (2009). Schema exchange: Generic mappings for transforming data and metadata. Data & Knowledge Engineering, 68(7), 665-682.

Paradigm Project (2005-2007b), "Guidance for creators of personal papers", available at: www. paradigm.ac.uk/guidanceforcreators/guidance-for-creators-of-personal-digital-archives.pdf

Park, E. G. (2007). Building interoperable Canadian architecture collections: initial metadata assessment. Electronic Library, The, 25(2), 207-218.

Park, E. G., & Richard, M. (2011). Metadata assessment in e-theses and dissertations of Canadian institutional repositories. Electronic Library, The, 29(3), 394-407.

Park, E. G., Zou, Q., & McKnight, D. (2007). Electronic thesis initiative: pilot project of McGill University, Montreal. Program: electronic library and information systems, 41(1), 81-91.

Park, J. R., & Tosaka, Y. (2010). Metadata quality control in digital repositories and collections: Criteria, semantics, and mechanisms. Cataloging & Classification Quarterly, 48(8), 696-715.

Park, J.R. (2006). Semantic interoperability and metadata quality: An analysis of metadata item records of digital image collections. Knowledge Organization, 33(1), 20–34.

Park, J.R. (2009). 'Metadata Quality in Digital Repositories: A Survey of the Current State of the Art', Cataloging & Classification Quarterly, 47: 3, 213 — 228

Parsons, M. A., Godøy, Ø., LeDrew, E., De Bruin, T. F., Danis, B., Tomlinson, S., & Carlson, D. (2011). A conceptual framework for managing very diverse data for complex, interdisciplinary science. Journal of Information Science,37(6), 555-569.

Parsons, M.A., & Duerr, R. (2005). Designating user communities for scientific data: challenges and solutions.Data Science Journal2005; 4: 31–38.

Parsons, M.A., Brodzik, M.J., & Rutter, N.J. (2004). Data management for the cold land processes experiment: improving hydrological science. Hydrological Processes2004; 18: 3637–3653.

Paskin, N. (1999). The digital object identifier system: digital technology meets content management. Interlending & document supply, 27(1), 13-16.

Patra, C. (2008). Digital repository in ceramics: a metadata study. Electronic Library, The, 26(4), 561-581.

Patton, M., Reynolds, D., Choudhury, G. S., & DiLauro, T. (2004). Toward a Metadata Generation Framework. D-Lib Magazine, 10(11), 1082-9873.

Pawlowski J.M. (2007). The quality adaptation model: adaptation and adoption of the quality standard ISO/IEC 19796-1 for learning, education, and training.Educational Technology & Society10, 3–16.

Payette, S., Blanchi, C., Lagoze, C., & Overly, E. (1999). Interoperability for digital objects and repositories. D-Lib Magazine, 5(5), 1082-9873.

Pearce, J., Pearson, D., Williams, M., & Yeadon, S. (2008). The Australian METS Profile–A journey about metadata. National Library of Australia Staff Papers.

Peng, F., & McCallum, A. (2004). "Accurate information extraction from research papers using conditional random fields",Proceedings of Human Language Technology Conference and North American Chapter of the Association for Computational Linguistics. HLTNAACL, Boston, MA, 2-7 May, pp. 329-36

Persico, D., Sarti, L., & Viarengo, V. (1992). Browsing a databaseof multimedia learning material. Interactive Learning International, 8, 213-235.

Peters, D., & Lossau, N. (2011). DRIVER: building a sustainable infrastructure for global repositories. Electronic Library, The, 29(2), 249-260.

Petr, K. (2007). Quality measurement of Croatian public and academic libraries: a methodology. Performance Measurement and Metrics, 8(3), 170-179.

Peyala, V. (2011). Impact of using information technology in central university libraries in India: Results of a survey. Program: electronic library and information systems, 45(3), 308-322.

Phillips, M.E. (2005). "Selective archiving of web resources: a study of acquisition costs at the National Library of Australia",RLG DigiNews, Vol. 9 No. 13, available at: http://www.nla.gov.au/openpublish/index.php/nlasp/article/view/1229/1514

Pickton, M., Morris, D., Meece, S., Coles, S., & Hitchcock, S. (2011). Preserving repository content: practical tools for repository managers. Journal of Digital Information, 12(2).

Pietruch-Reizes, D. (2010). A new look at the university libraries in context: European Research Area. Library Management, 31(1/2), 36-41.

Pisanski, J., & Žumer, M. (2005). National library web sites in Europe: an analysis. Program: electronic library and information systems, 39(3), 213-226.

Pisanski, J., & Žumer, M. (2007). Functional requirements for bibliographic records: an investigation of two prototypes. Program: electronic library and information systems, 41(4), 400-417.

Plodzien, J., Stemposz, E., & Stasiecka, A. (2006). An approach to the quality and reusability of metadata specifications for e-learning objects. Online Information Review, 30(3), 238-251.

Plum, T., Franklin, B., Kyrillidou, M., Roebuck, G., & Davis, M. (2010). Measuring the impact of networked electronic resources: developing an

assessment infrastructure for libraries, state, and other types of consortia.Performance Measurement and Metrics, 11(2), 184-198.

Poll, R. (2007). Benchmarking with quality indicators: national projects.Performance Measurement and Metrics, 8(1), 41-53.

Poll, R. (2008). Ten years after: Measuring Quality revised.Performance Measurement and Metrics, 9(1), 26-37.

Poll, R., & te Boekhorst, P. (2007). Measuring Quality: Performance Measurement in Libraries 2nd revised edition. IFLA Publication, 127.

Polsani, P. R. (2006). Use and abuse of reusable learning objects. Journal of Digital information, 3(4).

Polydoratou, P. (2007). Use of digital repositories by chemistry researchers: results of a survey. Program: electronic library and information systems, 41(4), 386-399.

Pond, W. K. (2001). Twenty-first century education and training: Implications for quality assurance. The Internet and higher education, 4(3), 185-192.

Pors, N. O., Dixon, P., & Robson, H. (2004). The employment of quality measures in libraries: cultural differences, institutional imperatives and managerial profiles. Performance Measurement and Metrics, 5(1), 20-27.

Prasad, A. R. D., & Guha, N. (2008). Concept naming vs concept categorisation: a faceted approach to semantic annotation. Online information review, 32(4), 500-510.

Preston, G., Phillips, R., Gosper, M., McNeill, M., Woo, K., & Green, D. (2010). Web-based lecture technologies: Highlighting the changing nature of teaching and learning. Australasian Journal of Educational Technology, 26(6), 717-728.

Prudlo, M. (2005). E-archiving: An overview of some repository management software tools. UKOLN.

Pynoo, B., Tondeur, J., Van Braak, J., Duyck, W., Sijnave, B., & Duyck, P. (2012). Teachers' acceptance and use of an educational portal. Computers & Education, 58(4), 1308-1317.

Qutab, S., & Mahmood, K. (2009). Library web sites in Pakistan: An analysis of content. Program: electronic library and information systems, 43(4), 430-445.

Rada, R. (1995a). Developing educational hypermedia: coordination and reuse. Norwood, NJ: Ablex.

Rada, R. (1995b). Hypertext, multimedia, and hypermedia. The New Review of Hypermedia and Multimedia, 1, 1-21.

Rada, R. (2001). Levels of reuse in educational info rmation systems. Campus-Wide Information Systems, 18(3), 103-109.

Raddick, J., Lintott, C.J., Schawinski, K., Thomas, D., Nichol, R.C., Andreescu, D., Bamford, S., Land, K.R., Murray, P., Slosar, A., Szalay, A.S., & Vandenberg, J. (2007). Galaxy zoo: An experiment in public science participation. Bulletin of the American Astronomical Society, 39, 892.

Raitt, D. (1999). “Some European developments in digital libraries”, paper presented at the 11th International Conference on New Information Technology, Taipei, August 18-20

Ranganathan, S.R. (1931). The Five Laws of Library Science, Madras Library Association, Madras.

Recker, M., Dorward, J., Dawson, D., Mao, X., Liu, Y., Palmer, B., ... & Park, J. (2005). Teaching, designing, and sharing: A context for learning objects. Interdisciplinary Journal of E-Learning and Learning Objects, 1(1), 197-216.

Recker, M.M., & Wiley, D.A. (2001). “A non-authoritative educational metadata ontology for filtering and recommending learning objects”, Journal of Interactive Learning Environments, Swets and Zeitlinger, Lisse.

Reerink, H. (2003). Practice, organisation and quality control of digitization projects. Liber Quarterly, 13(2).

Rehak, D. R., Nicholas, N., & Ward, N. (2009). Service-Oriented Models for Educational Resource Federations. D-Lib Magazine, 15(11/12), 1082-9873.

Reilly, B.F. (2003). Developing Print Repositories: Models for Shared Preservation and Access, Council on Library and Information Resources, Washington DC, available at: www.clir.org/pubs/reports/pub117/pub117.pdf (accessed 27 March 2007).

Renda, M. E., & Straccia, U. (2005). A personalized collaborative digital library environment: a model and an application. Information processing & management, 41(1), 5-21.

Research Libraries Group (1999). REACH (Record Export for Art and Cultural Heritage) Project Summary Report, available at: www.rlg.org/reach.html

Resnick, T., Ugaz, A., Burford, N., & Carrigan, E. (2008). E-resources: transforming access services for the digital age. Library Hi Tech, 26(1), 141-154.

Reznik-Zellen, R. C., & Adamick, J. (2010). Representation and Recognition of Subject Repositories. D-Lib Magazine, 16(9/10).

Reznik-Zellen, R. C., & Adamick, J. (2010). Trends in Large-Scale Subject Repositories. D-Lib Magazine, 16(11/12).

Rieh, S. Y. (2002). Judgment of information quality and cognitive authority in the Web. Journal of the American Society for Information Science and Technology,53(2), 145-161.

Rieh, S. Y., Jean, B. S., Yakel, E., Markey, K., & Kim, J. (2008). Perceptions and experiences of staff in the planning and implementation of institutional repositories. Library Trends, 57(2), 168-190.

Ritzhaupt, A. (2010). Learning Object Systems and Strategy: A description and discussion. Interdisciplinary Journal of E-Learning and Learning Objects, 6(1), 217-238.

Roberts, A. (2003). “Metaphysics of metadata”, presentation, available at: www.edusource.ca/english/resources/MetadataNB.ppt

Robertson, J.R. (2005). “Metadata quality: implications for library and information science professionals”, Library Review, Vol. 54 No. 5, pp. 295-300.

Robinson, M. L., & Wusteman, J. (2007). Putting Google Scholar to the test: a preliminary study. Program: electronic library and information systems, 41(1), 71-80.

Rodriguez, M. A., Bollen, J., & Van de Sompel, H. (2006). The convergence of digital libraries and the peer-review process. Journal of Information Science,32(2), 149-159.

Roknuzzaman, M., & Umemoto, K. (2009). How library practitioners view knowledge management in libraries: A qualitative study. Library Management, 30(8/9), 643-656.

Rorissa, A. (2008). User-generated descriptions of individual images versus labels of groups of images: A comparison using basic level theory. Information Processing & Management, 44(5), 1741-1753.

Rorissa, A. (2010). A comparative study of Flickr tags and index terms in a general image collection. Journal of the American Society for Information Science and Technology, 61(11), 2230-2242.

Rosemann, M. (2000). Using reference models within the enterprise resource planning lifecycle. Australian Accounting Review, 10(22), 19-30.

Roszkowski, M., & Lukas, C. (1998). A distributed architecture for resource discovery using metadata. D-Lib magazine, 6.

Rothenberg, J. (1996). "Metadata to Support Data Quality and Longevity". 1st IEEE Metadata Conference, Silver Spring, Maryland http://www.computer.org/conferences/meta96/rothenberg_Paper/ieee. data-quality.html

Rothenberg, J., & Rand (1997). “A Discussion of Data Quality for Verification, Validation, and Certification (VV&C) of Data to be Used in Modeling”, Rand Project Memorandum PM-709-DMSO, Rand. See also Data Quality Templates: http://vva.dmso.mil/Templates/Dataquality/default.htm

Rotman, D., Procita, K., Hansen, D., Sims Parr, C., & Preece, J. (2012). Supporting content curation communities: The case of the Encyclopedia of Life.Journal of the American Society for Information Science and Technology, 63(6), 1092-1107.

Rowley, J. (2001). Knowledge management in pursuit of learning: the learning with knowledge cycle. Journal of Information Science, 27(4), 227-237.

Rull, G., Farré, C., Teniente, E., & Urpí, T. (2008). Validation of mappings between schemas. Data & Knowledge Engineering, 66(3), 414-437.

Saarti, J., Juntunen, A., & Taskinen, A. (2010). Multi-faceted measuring of the quality as a tool for quality improvement in the Kuopio University Library, Finland. Performance measurement and metrics, 11(1), 39-46.

Sadykova, G., & Meskill, C. (2009). The Language of Digital Learning Objects: A Cross-disciplinary Study. Language, 5(2).

Saeed, R.S. (2006), "How digital libraries can support e-learning",The Electronic Library, Vol. 24, No. 3, pp. 389-401.

Sánchez-Alonso, S. (2009). Enhancing availability of learning resources on organic agriculture and agroecology. The Electronic Library 27(5): 792-813.

Sanz-Rodríguez, J., Dodero, J. M., & Sánchez-Alonso, S. (2010). Ranking Learning Objects through Integration of Different Quality Indicators. Learning Technologies, IEEE Transactions on, 3(4), 358-363.

Sarti, L., & Marcke, K.V. (1995). Reuse in intelligent courseware authoring. In N. Major, T. Murray, & C. Bloom (Eds.), AI-ED 95 Workshop on Authoring Shelles for Intelligent Tutoring Systems – 7th World Conference on Artificial Intelligence in Education. Washington (U.S.A) – August 16-19, 1995(pp. 82-87). Washington: AACE.

Satoh, Y., Nagata, H., Kytömäki, P., & Gerrard, S. (2005). Evaluation of the university library service quality: analysis through focus group interviews.Performance Measurement and Metrics, 6(3), 183-193.

Saumure, K., & Shiri, A. (2008). Knowledge organization trends in library and information studies: a preliminary comparison of the pre-and post-web eras.Journal of Information Science, 34(5), 651-666.

Saye, J.D. (2001). "The organization of electronic resources in the library and information science curriculum",OCLC Systems and Services, Vol. 17 No. 2, pp. 71-8.

Schibeci, R., Lake, D., Phillips, R., Lowe, K., Cummings, R., & Miller, E. (2008). Evaluating the use of learning objects in Australian and New Zealand schools. Computers & Education, 50(1), 271-283.

Schoner, V., Buzza, D., Harrigan, K., & Strampel, K. (2005). Learning objects in use:'Lite'assessment for field studies. Journal of Online Learning and Teaching, 1(1), 1-18.

Schreiber, G., Amin, A., Aroyo, L., van Assem, M., de Boer, V., Hardman, L., ... & Wielinga, B. (2008). Semantic annotation and search of cultural-heritage collections: The MultimediaN E-Culture demonstrator. Web Semantics: Science, Services and Agents on the World Wide Web, 6(4), 243-249.

Schwartz, C. (2001). Sorting out the Web: Approaches to Subject Access, Ablex, Westport, CT.

Schwartz, C. (2008). Thesauri and facets and tags, oh my! A look at three decades in subject analysis. Library Trends, 56(4), 830-842.

Schweik, C.M., Stepanov, A., & Grove, J.M. (2005). "The open research system: a web-based metadata and data repository for collaborative research",Computers and Electronics in Agriculture, Vol. 47 No. 3, pp. 221-42.

Scigliano, J. A., & Dringus, L. P. (2000). A lifecycle model for online learning management: 21 critical metrics for the 21st century. The Internet and higher education, 3(1), 99-115.

Secker, J. (2005). DELIVERing library resources to the virtual learning environment. Program: electronic library and information systems, 39(1), 39-49.

Secker, J. (2008). Social software and libraries: a literature review from the LASSIE project. Program: electronic library and information systems, 42(3), 215-231.

Segura, A., Vidal-Castro, C., Menéndez-Domínguez, V., Campos, P. G., & Prieto, M. (2011). Using data mining techniques for exploring learning object repositories. Electronic Library, The, 29(2), 162-180.

Sellitto, C., & Burgess, S. (2005). Towards a weighted average framework for evaluating the quality of web-located health information. Journal of Information Science, 31(4), 260-272.

Semple, N. (2004). Developing a digital preservation strategy at Edinburgh University Library. Vine, 34(1), 33-37.

Sen, A. (2004). "Metadata management: past, present and future", Decision Support Systems, Vol. 37, pp. 151-73.

Shabajee, P., Bollen, J., Luce, R., & Weig, E. (2002). Primary Multimedia Objects and'Educational Metadata'A Fundamental Dilemma for Developers of Multimedia Archives; Evaluation of Digital Library Impact and User Communities by Analysis of Usage Patterns; The KYVL Kentuckiana Digital Library Project: Background and Current Status; DPDx Collection. D-Lib Magazine, 8(6), n6.

Shankaranarayanan, G. & Even, A. (2006). "The metadata enigma",Communications of the ACM, Vol. 49 No. 2, pp. 88-94.

Shankaranarayanan, G., Ziad, M. & Wang, R.Y. (2003). "Managing data quality in dynamic decision-making environments: an information product approach",Journal of Database Management, Vol. 14 No. 4, pp. 14-32.

Sharifabadi, S. R. (2006). How digital libraries can support e-learning. Electronic Library, The, 24(3), 389-401.

Shea, P., McCall, S., & Ozdogru, A. (2006). Adoption of the multimedia educational resource for learning and online teaching (MERLOT) among higher education faculty: Evidence from the State University of New York Learning Network. MERLOT Journal of Online Learning and Teaching, 2(3).

Shenton, A. K. (2011). Two pupil-centred approaches to the assessment of school libraries. Performance Measurement and Metrics, 12(1), 38-49.

Shepherd, E. (2010). In-service training for academic librarians: a pilot programme for staff. Electronic Library, The, 28(4), 507-524.

Shin, E. J. (2006). Implementing a collaborative digital repository: the dCollection experience in South Korea. Interlending & document supply, 34(4), 160-163.

Shindo, K., Koshizuka, N. & Sakamura, K. (2003). “Large-scale ubiquitous information system for digital museum”,Applied Informatics, March, pp. 172-8.

Shiri, A. (2008). Metadata-enhanced visual interfaces to digital libraries. Journal of Information Science, 34(6), 763-775.

Shiri, A. (2009). An examination of social tagging interface features and functionalities: An analytical comparison. Online Information Review, 33(5), 901-919.

Shoeb, M. Z. H., & Ahmed, S. Z. (2009). Individual differences in service quality assessment: A study of a private university library system in Bangladesh.Performance Measurement and Metrics, 10(3), 193-211.

Shon, J. and Musen, M.A. (1999), “The low availability of metadata elements for evaluating the quality of medical information on the world wide web”,Proceedings of the 1999 American Medical Informatics Association (AMIA’99) Symposium, 6-10 November.

Shreeves, S. L., & Cragin, M. H. (2008). Introduction: Institutional repositories: Current state and future. Library Trends, 57(2), 89-97.

Shreeves, S. L., & Kirkham, C. M. (2006). Experiences of educators using a portal of aggregated metadata. Journal of Digital Information, 5(3).

Shreeves, S. L., Kaczmarek, J. S., & Cole, T. W. (2003). Harvesting cultural heritage metadata using the OAI protocol. Library hi tech, 21(2), 159-169.

Sicilia, M. A. (2006), “Metadata, semantics and ontology: providing meaning to information resources”, International Journal of Metadata, Semantics and Ontologies, Vol. 1 No. 1, pp. 83-6.

Sicilia, M. A., García, E., Pagés, C., Martínez, J.J., & Gutiérrez, J. (2005). Complete metadata records in learning object repositories: Some evidence and requirements. International Journal of Learning Technology, 1(4), 411–424.

Sicilia, M. A., Garcia-Barriocanal, E., Sanchez-Alonso, S., & Soto, J. (2005, July). A semantic lifecycle approach to learning object repositories. In Telecommunications, 2005. advanced industrial conference on telecommunications/service assurance with partial and intermittent resources conference/e-learning on telecommunications workshop. aict/sapir/elete 2005. proceedings (pp. 466-471). IEEE.

Silius, K., Miilumaki, T., Huhtamaki, J., Tebest, T., Merilainen, J., & Pohjolainen, S. (2010). Students' motivations for social media enhanced studying and learning. Knowledge Management & E-Learning: An International Journal (KM&EL), 2(1), 51-67.

Silva, A. J., Gonçalves, M. A., Laender, A. H., Modesto, M. A., Cristo, M., & Ziviani, N. (2009). Finding what is missing from a digital library: A case study in the Computer Science field. Information Processing & Management, 45(3), 380-391.

Simeoni, F., Yakici, M., Neely, S., & Crestani, F. (2008). Metadata harvesting for content-based distributed information retrieval. Journal of the American Society for information science and technology, 59(1), 12-24.

Simpson, P., & Hey, J. (2006). Repositories for research: Southampton's evolving role in the knowledge cycle. Program: electronic library and information systems, 40(3), 224-231.

Singh, G., Mittal, R., & Ahmad, M. (2007). A bibliometric study of literature on digital libraries. Electronic Library, The, 25(3), 342-348.

Singh, H. (2000). Learning Content Management Systems: new technologies for new learning approaches.

Sinikara, K. (2006). Evaluation as a Tool for Developing the Quality of Academic Libraries. Case Study at a Large and Traditional Research Oriented Scandinavian University. Liber Quarterly, 16(3/4).

Slaton, A. & Abbate, J. (2001) The Hidden Lives of Standards: Technical Prescriptions and the Transformation of Work in America. In M. T. Allen & G. Hecht (Eds.), Technologies of Power (pp 95-144). Cambridge, MA: MIT Press.

Smith, M., Barton, M., Bass, M., Branschofsky, M., McClellan, G., Stuve, D., ... & Walker, J. H. (2003). DSpace: An open source dynamic digital repository.

Sokvitne, L. (2000). An evaluation of the effectiveness of current Dublin Core metadata for retrieval.Presented at the VALA Conference, 2000. Available from http://www.vala.org.au/vala2000/2000pdf/Sokvitne.PDF. Accessed 19 May 2004.

Sosteric, M. & Hesemeier, S. (2004). A first step toward a theory of learning objects. In R. McGreal (Ed.), Online education using learning objects. London: Routledge/Falmer.

South, J. B., & Monson, D. W. (2000). A university-wide system for creating, capturing, and delivering learning objects. In D. A. Wiley (Ed.), The instructional use of learning objects. Retrieved August 26th, 2005, from http://reusability.org/read/chapters/south.doc

Sreenivasulu, V. (2000). The role of a digital librarian in the management of digital information systems (DIS). Electronic Library, The, 18(1), 12-20.

Stefaner, M., Dalla Vecchia, E., Condotta, M., Wolpers, M., Specht, M., Apelt, S., & Duval, E. (2007). MACE–enriching architectural learning objects for experience multiplication. In Creating New Learning Experiences on a Global Scale (pp. 322-336). Springer Berlin Heidelberg.

Steinacker, A., Ghavam, A. & Steinmetz, R. (2001), “Metadata standards for Web-based resources”,IEEE Multimedia, January-March, pp. 70-6.

Stephens, R.T. (2004), “Utilizing metadata as a knowledge communication tool”,International Professional Communication Conference, Twin City, MN, 29 September – 1 October 2004, pp. 55-60.

Streatfield, D., & Markless, S. (2009). What is impact assessment and why is it important?. Performance Measurement and Metrics, 10(2), 134-141.

Strong, D., Lee, Y., & Wang, R. (1997). Data quality in context. Communications of the ACM, 40(5), 103–110

Sturdy, D. (2001), "Squirrels and nuts: metadata and knowledge management", Business Information Review, Vol. 18 No. 4, pp. 34-42.

Stvilia, B., & Gasser, L. (2008). Value based metadata quality assessment. Library and Information Science Research, 30(1), 67–74.

Stvilia, B., & Jörgensen, C. (2010). Member activities and quality of tags in a collection of historical photographs in Flickr. Journal of the American Society for Information Science and Technology, 61(12), 2477-2489.

Stvilia, B., Gasser, L., Twidale, M. B., & Smith, L. C. (2007). A framework for information quality assessment. Journal of the American Society for Information Science and Technology, 58(12), 1720-1733.

Stvilia, B., Gasser, L., Twidale, M., Shreeves, S., & Cole, T. (2004). Metadata quality for federated collections. In Proceedings of ICIQ04 - 9th International Conference on Information Quality, (pp. 111-125). Boston, MA.

Su, A., Yang, S. J., Hwang, W. Y., & Zhang, J. (2010). A Web 2.0-based collaborative annotation system for enhancing knowledge sharing in collaborative learning environments. Computers & Education, 55(2), 752-766.

Subirats, I., Onyancha, I., Salokhe, G., Kaloyanova, S., Anibaldi, S., & Keizer, J. (2008). Towards an architecture for open archive networks in agricultural sciences and technology. Online Information Review, 32(4), 478-487.

SuKantarat, W. (2008). Digital initiatives and metadata use in Thailand.Program: electronic library and information systems, 42(2), 150-162.

Suleman, H., Fox, E. A., Kelapure, R., Krowne, A., & Luo, M. (2003). Building digital libraries from simple building blocks. Online Information Review, 27(5), 301-310.

Sullivan, B., Wooda, C.L., Iliffa, M.J., Bonneya, R.E., Finka, D., & Kelling, S. (2009). eBird: A citizen-based bird observation network in the biological sciences. Biological Conservation, 142(10), 2282– 2292.

Sun, L. (2008). A metadata manager's role in collaborative projects: The Rutgers University Libraries experience. Electronic Library, The, 26(6), 777-789.

Sun, L. (2011). Batch loading in metadata creation: a case study. Electronic Library, The, 29(4), 538-549.

Sung, Y. T., Chang, K. E., & Yu, W. C. (2011). Evaluating the reliability and impact of a quality assurance system for E-learning courseware. Computers & Education, 57(2), 1615-1627.

Supekar, K. (2005), “A peer-review approach for ontology evaluation”,Proceedings of the Eighth International Prote ´ge´ Conference, Madrid, Spain, 18-21 July.

Supekar, K., Patel, C. & Lee, Y. (2004), “Characterizing quality of knowledge on semantic web”, in Barr, V. and Markov, Z. (Eds), Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference (FLAIRS’04), AAAI Press, Miami Beach, FL

Surry, D. W., Grubb, A. G., Ensminger, D. C., & Ouimette, J. (2009). Implementation of web-based learning in colleges of education: Barriers and enablers. Canadian Journal of Learning and Technology/La revue canadienne de l’apprentissage et de la technologie, 35(3).

Suseela, V. J. (2011). Application of usage statistics for assessing the use of e-journals in University of Hyderabad: A case study. Electronic Library, The,29(6), 751-761.

Sutradhar, B. (2006). Design and development of an institutional repository at the Indian Institute of Technology Kharagpur. Program: electronic library and information systems, 40(3), 244-255.

Sutton, S.A. (1999), “Conceptual design and deployment of a metadata framework for educational resources on the internet”,Journal of the American Society for Information Science, Vol. 50 No. 13, pp. 1182-92.

Swain, D. K. (2010). Students' keenness on use of e-resources. Electronic Library, The, 28(4), 580-591.

Swan, A. (2007). ‘The business of digital repositories’, in A DRIVER’S Guide to European Repositories(edited by Weenink, K., L. Waaijers and K. van Godtsenhoven). Amsterdam: Amsterdam University Press, http://www.keyperspectives.co.uk/openaccessarchive/reports/The_business_of_digital_repositories.pdf

Taibi, D., Gentile, M., & Seta, L. (2005). A semantic search engine for learning resources. Paper presented at the Third International Conference on Multimedia and Information & Communication Technologies in Education.

Takasu, A. (2003). “Bibliographic attribute extraction from erroneous references based on a statistical model”, Proceedings of the 3rd ACM/IEEE-CS Joint Conference on DigitalLlibraries. JCDL ’03, Houston, TX, USA, May 27-31, IEEE Computer Society, New York, NY, pp. 49-60.

Tambouris, E., Manouselis, N., & Costopoulou, C. (2007). Metadata for digital collections of e-government resources. Electronic Library, The, 25(2), 176-192.

Tammaro, A. M. (2005). Recognition and quality assurance in LIS: New approaches for lifelong learning in Europe. Performance measurement and metrics, 6(2), 67-79.

Tammaro, A. M. (2008). User perceptions of digital libraries: a case study in Italy. Performance Measurement and Metrics, 9(2), 130-137.

Tang, J., & Zhang, J. (2010). Modeling the evolution of associated data. Data & Knowledge Engineering, 69(9), 965-978.

Tansley, R., Bass, M., Stuve, D., Branschofsky, M., Chudnov, D., McClellan, G., & Smith, M. (2003, May). The DSpace institutional digital repository system: current functionality. In Proceedings of the 3rd ACM/IEEE-CS joint conference on Digital libraries (pp. 87-97). IEEE Computer Society.

Tarrant, D., O'Steen, B., Brody, T., Hitchcock, S., Jefferies, N., & Carr, L. (2009). Using OAI-ORE to transform digital repositories into interoperable storage and services applications. Code4Lib Journal, 6.

Tate, M., & Hoshek, D. (2009). A Model for the Effective Management of Re-Usable Learning Objects (RLOs): Lessons from a Case Study. Interdisciplinary Journal of E-Learning and Learning Objects, 5(1), 51-72.

Taylor, A. (2004). The organization of information (library and information science text series) (2nd ed.). Westport, CT: Libraries Unlimited.

Taylor, C. (2003). An Introduction to Metadata, University of Queensland Library, available at: www.library.uq.edu.au/iad/ctmeta4.html

Taylor, R. (1986). Value-added processes in information systems. Norwood, NJ: Ablex Publishing.

Tedd, L. A. (2006). ITProgram/IT: a record of the first 40 years of electronic library and information systems. Program: electronic library & information systems, 40(1), 11-26.

Tennant, R. (2004). A bibliographic metadata infrastructure for the twenty-first century. Library Hi Tech, 22(2), 175-181.

Tennant, R. (2004). Metadata's bitter harvest. Library Journal, 129(12), 32, 15 July 2004.

Tennant, R. (2005), "Lipstick on a pig", Library Journal, Vol. 130 No. 7, p. 34.

The JORUM Team. (2006). Repository Systems Watch.Retrieved July 15, 2007 from http://www.jorum.ac.uk/docs/pdf/Repository_Watch_final_05012006.pdf

Thibodeau, K. (2007). If you build it, will it fly? Criteria for success in a digital repository. Journal of digital information, 8(2).

Thomas, C., & McDonald, R. H. (2007). Measuring and comparing participation patterns in digital repositories. D-lib Magazine, 13(9/10), 1082-9873.

Thomas, M., Caudle, D. M., & Schmitz, C. M. (2009). To tag or not to tag?.Library Hi Tech, 27(3), 411-434.

Thompson, B., Kyrillidou, M., & Cook, C. (2008). How you can evaluate the integrity of your library service quality assessment data: intercontinental LibQUAL+® analyses used as concrete heuristic examples. Performance Measurement and Metrics, 9(3), 202-215.

Thompson, T. L., & MacDonald, C. J. (2005). Community building, emergent design and expecting the unexpected: Creating a quality eLearning experience.The Internet and Higher Education, 8(3), 233-249.

Thomson, J., Adams, D., Cowley, P.J. & Walker, K. (2003), "Metadata's role in a scientific archive",Computer, Vol. 36 No. 12, pp. 27-34.

Todd, C. (2003). Metadata mayhem: cataloguing electronic resources in the National Library of New Zealand. Electronic Library, The, 21(3), 214-222.

Tolley, H., & Shulruf, B. (2009). From data to knowledge: The interaction between data management systems in educational institutions and the delivery of quality education. Computers & Education, 53(4), 1199-1206.

Towle, B., & Halm, M. (2005). Designing adaptive learning environments with learning design. In Learning Design (pp. 215-226). Springer Berlin Heidelberg.

Tozer, G. (1999). Metadata Management for Information Control and Business Success (Boston: Artech House)

Tran, L. A. (2009). Evaluation of community web sites: A case study of the Community Social Planning Council of Toronto web site. Online Information Review, 33(1), 96-116.

Tran, L. A. (2010). A quantitative assessment of relationships between user characteristics and frequency of access to community information networks in New Zealand public libraries. Online Information Review, 34(3), 361-376.

Treloar, A., Groenewegen, D., & Harboe-Ree, C. (2007). The data curation continuum: Managing data objects in institutional repositories. D-Lib Magazine, 13(9), 4.

Trigg, R.H., Blomberg, J. & Suchman, L. (1999). "Moving document collections online: the evolution of a shared repository", in Bødker, S., Kyng, M. and Schmidt, K. (Eds) Proceedings of the 6th European Conference on Computer Supported Cooperative Work, Kluwer Academic Publishers, Copenhagen, Denmark, pp. 331-50.

Tsai, C. F., McGarry, K., & Tait, J. (2006). Qualitative evaluation of automatic assignment of keywords to images. Information processing & management,42(1), 136-154.

Tsai, L. C., Hwang, S. L., & Tang, K. H. (2011). Analysis of keyword-based tagging behaviors of experts and novices. Online Information Review, 35(2), 272-290.

Tsakonas, G., & Papatheodorou, C. (2008). Exploring usefulness and usability in the evaluation of open access digital libraries. Information Processing & Management, 44(3), 1234-1250.

Tzikopoulos, A., Manouselis, N., Yialouris, C. P., & Sideridis, A. B. (2007, July). Using educational metadata in a learning repository that supports lifelong learning needs of rural SMEs. In EFITA/WCCA 2007 Conference on Environmental and rural sustainability through ICT, Glaskow, UK (July 2007).

Uddin, M. N., & Janecek, P. (2007). The implementation of faceted classification in web site searching and browsing. Online information review,31(2), 218-233.

University of California Libraries, Bibliographic Services Task Force (2005). Rethinking How We Provide Bibliographic Services for the University of California, Final report, December, available at: http://libraries.universityofcalifornia.edu/sopag/BSTF/Final.pdf

Uren, V., Cimiano, P., Iria, J., Handschuh, S., Vargas-Vera, M., Motta, E., & Ciravegna, F. (2006). Semantic annotation for knowledge management: Requirements and a survey of the state of the art. Web Semantics: science, services and agents on the World Wide Web, 4(1), 14-28.

Valsamidis, S., Kontogiannis, S., Kazanidis, I., & Karakos, A. (2011). E-Learning Platform Usage Analysis. Interdisciplinary Journal of E-Learning and Learning Objects, 7(1), 185-204.

Van Hooland, S., Rodríguez, E. M., & Boydens, I. (2011). Between commodification and engagement: On the double-edged impact of user-generated metadata within the cultural heritage sector. Library trends, 59(4), 707-720.

Van Westrienen, G., & Lynch, C. A. (2005). Academic institutional repositories. D-lib Magazine, 11(9), 1082-9873.

Varlamis, I., & Apostolakis, I. (2006). The present and future of standards for e-learning technologies. Interdisciplinary Journal of E-Learning and Learning Objects, 2(1), 59-76.

VOA3R, (2013). D5.4. Report on Content Integration/Population (Phase 3), May 2013.

Veiga e Silva, L., Gonçalves, M. A., & Laender, A. H. (2007). Evaluating a digital library self-archiving service: The BDBComp user case study. Information Processing & Management, 43(4), 1103-1120.

Venturi, G., & Bessis, N. (2006, October). User-centred evaluation of an e-learning repository. In Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles (pp. 203-211). ACM.

Vinagre, M. H., Pinto, L. G., & Ochôa, P. (2011). Revisiting digital libraries quality: a multiple-item scale approach. Performance Measurement and Metrics, 12(3), 214-236.

Visual Resources Association (2003). Cataloging Cultural Objects: A Guide to Describing Cultural Works and Their Images, available at: www.vraweb.org/ccoweb

Vlasceanu, L., Gru¨nberg, L. & Pa^rlea, D. (Eds) (2007). Quality Assurance and Accreditation: A Glossary of Basic Terms and Definitions, 2nd ed., UNESCO European Centre for Higher Education (CEPES), Bucharest, Romania, available at: www.cepes.ro/publications/pdf/ Glossary_2nd.pdf (accessed 1 February 2010).

Vogias, K., Hatzakis, I., Manouselis, N., Szegedi, P. (2013). Extraction and Visualization of Metadata Analytics for Multimedia Learning

Repositories: the case of Terena TF-media net-work. Proceedings of the LACRO 2013 Workshop.

Völkel, M., & Haller, H. (2009). Conceptual data structures for personal knowledge management. Online Information Review, 33(2), 298-315.

Vuorikari, R., Manouselis, N., & Duval, E. (2008). Using metadata for storing, sharing and reusing evaluations for social recommendations: the case of learning resources. Social information retrieval systems: Emerging technologies and applications for searching the web effectively, 87-107.

Vuorikari, R., & Ochoa, X. (2009). Exploratory analysis of the main characteristics of tags and tagging of educational resources in a multi-lingual context. Journal of Digital Information, 10(2).

Vuorikari, R., Poldoja, H., & Koper, R. (2010). Comparison of tagging in an educational context: Any chances of interplay?. International Journal of Technology Enhanced Learning, 2(1), 111-131.

Waaijers, L. (2005). From libraries to 'libratories'. First Monday, 10(12).

Waaijers, L., & van der Graaf, M. (2011). Quality of Research Data, an Operational Approach. D-Lib Magazine, 17(1/2).

Waddington, S., Green, R., & Awre, C. (2012). CLIF: Moving repositories upstream in the content lifecycle. Journal of Digital Information, 13(1).

Waibel, G., LeVan, R., & Washburn, B. (2010). Museum Data Exchange: Learning How to Share. D-Lib Magazine, 16(3), 3.

Walters, D.J. (1998). "What are digital libraries?",CLIR Issues, No. 4, July/August, available at: www.clir.org/pubs/issues04.html

Walters, W. H. (2003). Expertise and evidence in the assessment of library service quality. Performance Measurement and Metrics, 4(3), 98-102.

Wand, Y., & Wang, R. (1996). Anchoring data quality dimensions in ontological foundations. Communications of the ACM, 39(11), 86–95.

Wang, H., Xie, M., & Goh, T. N. (1999). Service quality of Internet search engines. Journal of Information Science, 25(6), 499-507.

Wang, M. Y., & Hwang, M. J. (2004). The e-learning library: only a warehouse of learning resources?. Electronic Library, The, 22(5), 408-415.

Wang, R.Y. (1998). A Product Perspective on Total Data Quality Management. Communications of the ACM, 41, 2 (February 1998)

Wang, R.Y., & Strong, D.M. (1996). Beyond accuracy: What data quality means to data consumers. Journal of Management Information Systems, 12(4), 5–33.

Wang, Y. M., & Wang, Y. S. (2009). Examining the dimensionality and measurement of user-perceived knowledge and information quality in the KMS context. Journal of Information Science, 35(1), 94-109.

Wang, Z. (2003). "The digital library and librarians quality",SCI/TECH Information Development & Economy, Vol. 13 No. 1, pp. 32-4.

Warwick, C., Terras, M., Galina, I., Huntington, P., & Pappa, N. (2008). Library and information resources and users of digital resources in the

humanities.Program: electronic library and information systems, 42(1), 5-27.

Wayne, L. (2004). “Quality metadata”, paper presented at the 2004 ESRI Federal User Conference, Washington DC, January 22, available at: wwwfgdcgov/library/presentations/2004-presentations

Weaver, D., Spratt, C., & Nair, C. S. (2008). Academic and student use of a learning management system: Implications for quality. Australasian Journal of Educational Technology, 24(1), 30-41.

Weaver, M. (2007). Contextual metadata: faceted schemas in virtual library communities. Library Hi Tech, 25(4), 579-594.

Wei, C. P., Jen-Hwa Hu, P., & Chen, H. H. (2002). Design and evaluation of a knowledge management system. Software, IEEE, 19(3), 56-59.

Weibel, S. (1995). Metadata: the foundations of resource description. D-Lib Magazine, July 1995, at http://www.dlib.org/dlib/July95/07weibel.html.

Weibel, S. (1997). The evolving metadata architecture for the World Wide Web: bringing together the semantics, structure, and syntax of resource description. Proceedings of the International Symposium on Research, Development and Practice in Digital Libraries 1997 (ISDL97), Tsukuba, Japan, 18-21 November. ULIS, pp. 16-22.

Weibel, S. and Miller, E. (1997). “Image description on the internet: a summary of the CNI/OCLC image metadata workshop”, D-Lib Magazine, Vol. 3, January, available at: www.dlib.org/dlib/january97/oclc/01weibel.html

Weibel, S., & Wagner, H. (2005). The Dublin Core Metadata Registry: Requirements, Implementation, and Experience. JoDI: Journal of Digital Information, 6(2), 3.

Weinheimer, J. (2000). "How to Keep the Practice of Librarianship Relevant in the Age of the Internet". Vine, Special issue on Metadata, part 1, 116: 14-27

Westbrook, R. N., Johnson, D., Carter, K., & Lockwood, A. (2012). Metadata clean sweep: A digital library audit project. D-Lib Magazine, 18(5), 2.

Westell, M. (2006). ‘Institutional repositories: proposed indicators of success’, Library Hi Tech, 24(2), p. 211–226.

Westrum, A. L., Rekkavik, A., & Tallerås, K. (2012). Improving the presentation of library data using FRBR and Linked data. Code4Lib Journal, 16.

White, A. (2005). The development of digital resources by library and information professionals and historians: Two case studies from Northern Ireland. Program: electronic library and information systems, 39(4), 324-336.

Wiley, D. A. (2000). Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy. In D. A. Wiley (Ed.), The instructional use of learning objects.Retrieved August 21, 2007, from http://reusability.org/read/chapters/wiley.doc

Wiley, D.A. (2003). Learning objects. In A. Kovalchick, & K. Dawson (Eds.), Educational technology: An encyclopedia. Santa Barbara: ABC-CLIO. Retrieved July 17, 2004 from http://wiley.ed.usu.edu/docs/encyc.pdf

Willer, M., Buzina, T., Holub, K., Zajec, J., Milinovic, M., & Topolščak, N. (2008). Selective archiving of web resources: a study of processing costs.Program: electronic library and information systems, 42(4), 341-364.

William, Y. (2001),Digital Libraries, Massachusetts Institute of Technology, London.

Willoughby, T., Anderson, S., Wood, E., Mueller, J., & Ross, C. (2009). Fast searching for information on the Internet to use in a learning context: The impact of domain knowledge. Computers & Education, 52(3), 640-648.

Wilson, A.J. (2007). Toward releasing the metadata bottleneck: A baseline evaluation of contributor-supplied metadata. Library Resources & Technical Service, 51(1), 16–28.

Wilson, F., & Town, J. S. (2006). Benchmarking and library quality maturity.Performance measurement and metrics, 7(2), 75-82.

Wilson, G., & Stacey, E. (2004). Online interaction impacts on learning: Teaching the teachers to teach online. Australian Journal of Educational Technology, 20(1), 33-48.

Witt, M. (2008). Institutional repositories and research data curation in a distributed environment. Library Trends, 57(2), 191-201.

Witten, I. H., & Bainbridge, D. (2005). Creating digital library collections with Greenstone. Library hi tech, 23(4), 541-560.

Witten, I. H., Loots, M., Trujillo, M. F., & Bainbridge, D. (2001). The promise of digital libraries in developing countries. Communications of the ACM, 44(5), 82-85.

Wolpers, M., Memmel, M., & Stefaner, M. (2010). Supporting architecture education using the MACE system. International Journal of Technology Enhanced Learning, 2(1), 132-144.

Wong, I. S., & Li, Y. O. (2009). Creating a virtual union catalog for Hong Kong dissertations and theses collections. Electronic Library, The, 27(2), 331-341.

Woudstra, L., van den Hooff, B., & Schouten, A. P. (2012). Dimensions of quality and accessibility: Selection of human information sources from a social capital perspective. Information Processing & Management, 48(4), 618-630.

Wu, H. Y., & Lin, H. Y. (2012). A hybrid approach to develop an analytical model for enhancing the service quality of e-learning. Computers & Education,58(4), 1318-1338.

Wu, Y. D., & Liu, M. (2001). Content management and the future of academic libraries. Electronic Library, The, 19(6), 432-440.

Wynne, S. C., & Hanscom, M. J. (2011). The effect of next-generation catalogs on catalogers and cataloging functions in academic libraries. Cataloging & Classification Quarterly, 49(3).

Xia, J., & Sun, L. (2007). 'Factors to assess self-archiving in institutional repositories', Serials Review, 33(2), pp. 73–80, doi: 10.1016/j.serrev.2006.09.002.

Xiao, T. (2003). "Studying on the concept of digital library",Information Research, No. 3, pp. 10-12.

Xie, H. I. (2008). Users' evaluation of digital libraries (DLs): Their uses, their criteria, and their assessment. Information processing & management, 44(3), 1346-1373.

Yahya, Y., & Yusoff, M. (2008). Towards a comprehensive learning object metadata: Incorporation of context to stipulate meaningful learning and enhance learning object reusability. Interdisciplinary Journal of E-Learning and Learning Objects, 4(1), 13-48.

Yakel, E., & Tibbo, H. (2010). Standardized survey tools for assessment in archives and special collections. Performance Measurement and Metrics, 11(2), 211-222.

Yamada, T. (2004). "Role of metadata in quality assurance of multi-country statistical data in the case of UNIDO industrial statistics",Proceedings of the Conference on Data Quality forInternational Organizations, Wiesbaden, Germany, 27-28 May.

Yamamoto, T. (2010). A Proposal for Measuring Interactivity that Brings Learning Effectiveness. Knowledge Management & E-Learning: An International Journal (KM&EL), 2(1), 6-16.

Yang, J. T. D., & Tsai, C. Y. (2003, July). An implementation of SCORM-compliant learning content management system-content repository management system. In Advanced Learning Technologies, 2003. Proceedings. The 3rd IEEE International Conference on (p. 453). IEEE.

Yao, L., & Zhao, P. (2009). Digital libraries in China: progress and prospects.Electronic Library, The, 27(2), 308-318.

Yen, B., & Park, J. (2006). An assessment of metadata quality: A case study of the national science digital library metadata repository. In Haidar Moukdad (ed.), CAIS/ACSI 2006 Information Science Revisited: Approaches to Innovation from http://www.cais-acsi.ca/proceedings/2006/bui_2006.pdf.

Yen, N. Y., Shih, T. K., Chao, L. R., & Jin, Q. (2010). Ranking metrics and search guidance for learning object repository. Learning Technologies, IEEE Transactions on, 3(3), 250-264.

Yoon, J. P., & Kim, S. (2000). Schema extraction for multimedia xml document retrieval. InProcs. of the 1st International Conference on Web Information Systems Engineering (WISE'00)(2000). IEEE Press.

Yordanova, K. (2007). Meta-data application in development, exchange and delivery of digital reusable learning content. Interdisciplinary Journal of E-Learning and Learning Objects, 3(1), 229-237.

Young, A., & Norgard, C. (2006). Assessing the quality of online courses from the students' perspective. The Internet and Higher Education, 9(2), 107-115.

Yu, S. C. (2006). The study of integrated frameworks for library and digital archives. Electronic Library, The, 24(5), 608-618.

Yu, S. C. (2008). Design of a model of publication sharing and harvesting. Electronic Library, The, 26(4), 582-593.

Yu, S. C., Lu, K. Y., & Chen, R. S. (2003). Metadata management system: design and implementation. Electronic Library, The, 21(2), 154-164.

Zeng, M. L. (1999). Metadata elements for object description and representation: a case report from a digitized historical fashion collection project. JASIS, 50(13), 1193-1208.

Zhang, J., & Dimitroff, A. (2005). The impact of metadata implementation on webpage visibility in search engine results (Part II). Information processing & management, 41(3), 691-715.

Zhang, J., & Jastram, I. (2006). A study of metadata element co-occurrence.Online Information Review, 30(4), 428-453.

Zhang, J., & Jastram, I. (2006). A study of the metadata creation behavior of different user groups on the Internet. Information processing & management,42(4), 1099-1122.

Zhang, L., Ye, P., & Liu, Q. (2011). A survey of the use of electronic resources at seven universities in Wuhan, China. Program: electronic library and information systems, 45(1), 67-77.

Zhang, Y., & Li, Y. (2008). A user-centered functional metadata evaluation of moving image collections. Journal of the American Society for Information Science and Technology, 59(8), 1331-1346.

Zhang, Y., & Li, Y. (2008). A user-centered functional metadata evaluation of moving image collections. Journal of the American Society for Information Science and Technology, 59(8), 1331-1346.

Zhou, Q. (2005). The development of digital libraries in China and the shaping of digital librarians. Electronic Library, The, 23(4), 433-441.

Zschocke, T., & Beniest, J. (2011). Adapting a quality assurance framework for creating educational metadata in an agricultural learning repository. The Electronic Library, 29(2), 181-199.

Zuccala, A., Thelwall, M., Oppenheim, C. & Dhiensa, R. (2006). Information Systems Committee Project Report. Digital Repository Management Practices, User Needs and Potential Users: An Integrated Analysis, available at: http://cybermetrics.wlv.ac.uk/DigitalRepositories/ FinalReport.pdf (accessed June 12, 2007).

Appendix A: Literature Review Dimensions

RIGOR	
Research Type	
Descriptive	Describes a phenomenon in its appearance without any use of theory.
Philosophical	Reflects upon a phenomenon without data or reference to any theory.
Theoretical	Reflects upon a phenomenon based on some theory but without empirical data or with only anecdotal and particular such.
Theory Generation	Attempts to analyze / interpret quantitative or qualitative data in a systematic manner for the purpose of model building.
Theory Testing	Attempts to test a theory, using quantitative or qualitative data in a systematic manner, i.e. not just strict theory testing.
Applied (Implementation)	Implements a theory or model in a specific case or product, without attempting to test it in any systematic manner.
Research Method	
Argument	Logical argument but not based on any particular theory or relating explicitly or by clear implication to any theory.
Case Story	Tells about a case but as opposed to a case study there is no strict data collection method. Usually own experiences or anecdotal evidence.
Survey	Covers qualitative overviews of several documents or cases.
Literature study	Only documents used, be they scientific, policy documents or other. Not necessarily strict method or even explicitly labeled as literature study.
Ethnography	Any attempt to understand actions by systematic observation and interpretation.
Grounded Theory	Theoretical examination of a phenomenon based on mathematical propositions and proofs.
Interpretative	Any kind of more strictly performed data collection than 'case story' but not necessarily strictly explained or spelled-out method for interpretation. A case study belongs here, but also more limited studies where qualitative or quantitative data is analyzed.
Experiment	Field experiment included.
Experiment (simulation)	Simulated experiment reproducing / imitating field conditions.
Product description	IT product, method, or similar, described by its developer.
Unclear	Other method, which cannot be classified in the other categories, or not clear method used.

Claim	
Ongoing research	Research which is not completed and paper which makes no claim as to the validity or scope of the findings.
Lessons	Only claims of anecdotal value, lessons learned
Descriptive	Claims validity but not generality. Authors claim to have described the situation correctly and/or credibly.
Normative	Claims generality beyond case.

Table A.1: Aspects related with the Rigor category of research characteristics

RELEVANCE	
Focus Unit	
Method	Focus is on the specific method or product proposed or used.
Individual	Focus is on an individual person or issues are addressed from the individuals' perspective.
Group	Focus is on a group of people with common characteristics.
Organization	Focus is on a specific organization or type of organizations (e.g. firms).
Sector	Focus is on a specific industry or sector of research field.
Society	The focus is general and a general perspective is adopted.
Target Audience	
Researchers	Results explicitly or implicitly primarily aim to guide other researchers.
Practitioners / Implementers	Results explicitly or implicitly primarily aim to people that can make practical use and/or implementation of a product or method.
Managers	Results explicitly or implicitly primarily aim to guide managers take decisions about the appropriate use and/or implementation of a product or method.
Unclear	The primary target audience of the results cannot be clearly identified.
Origin (Institution Type)	
Company	The affiliation of most of the authors (or of the primary author) is a company.
Research Institute	The affiliation of most of the authors (or of the primary author) is a research institute.
University	The affiliation of most of the authors (or of the primary author) is a university.
NGO	The affiliation of most of the authors (or of the primary author) is a non-governmental organization that does not belong to one of the above types.
Government	The affiliation of most of the authors (or of the primary author) is a governmental organization
Library	The affiliation of most of the authors (or the primary author) is a library

Discipline	
Computer Science	The background discipline of most of the authors (or of the primary author) is computer science, electrical / computer engineering, etc.
Information Systems	The background discipline of most of the authors (or of the primary author) is information systems etc.
Management	The background discipline of most of the authors (or of the primary author) is management, marketing etc.
Economics	The background discipline of most of the authors (or of the primary author) is economics, financial engineering etc.
Mathematics	The background discipline of most of the authors (or of the primary author) is mathematics.
Education	The background discipline of most of the authors (or of the primary author) is education
Libraries	The background discipline of most of the authors (or of the primary author) is libraries, information science, etc.
Other	The background discipline of most of the authors (or of the primary author) is some other discipline, different from the above ones.
Country	
Country	The country of most of the authors (or of the primary author).

Table A.2: Aspects related with the Relevance category of research characteristics

No	Element Name	Element Description
1	ID	A unique number assigned to each paper that was retrieved, following a specific coding scheme
2	Title	The title of the paper
3	Description	A short description of the paper or alternatively, its abstract
4	Year	The year that the paper was published
5	Volume	The volume of the journal where the paper was retrieved from
6	Issue	The issue of the journal where the paper was retrieved from
7	No of Authors	The number of authors that wrote the paper
8	No of Pages	The number of pages of the paper
9	No of References	The number of references contained within the paper
10	Source	The journal from where the paper was retrieved
11	Comments	Other comments related to the paper, as well as its actual source, if no file was downloaded (URL)
12	Research Type	The type of research that is carried out in the paper (see also Table A.1)

No	Element Name	Element Description
13	Method	The research method used within the paper (see also Table A.1)
14	Claim	The type of claim that the paper makes (see also Table A.1)
15	Focus Unit	The focus of the research carried out within the paper (see also Table A.2)
16	Target Audience	The primary target audience of the paper (see also Table A.2)
17	Institution	The type of institution where the authors of the paper come from (see also Table A.2)
18	Discipline	The type of discipline that the primary author of the paper comes from (see also Table A.2)
19	Country	The country of which the main author of the paper come from (one value – see also Table A.2)

Table A.3: Remaining categories used to classify literature

Appendix B: Literature Review Tables & Figures

No	Journal Title	1st Edition Year	Review Year Start	Review Year End	First Paper Found	Last Paper Found	# of papers identified
1	**The Electronic Library**	1983	1983	2012	1999	2011	**64**
2	**Program: electronic library and information systems**	1966	1966	2012	2002	2011	**51**
3	**Performance Measurement and Metrics**	2000	2000	2012	2001	2011	**35**
4	**Online Information Review**	1977	1977	2012	2002	2011	**34**
5	**D-Lib Magazine**	1995	1995	2012	1996	2011	**33**
6	**Interdisciplinary Journal of E-Learning and Learning Objects (IJELLO)**	2005	2005	2012	2005	2011	**33**
7	**Library Hi Tech**	1983	1983	2012	2001	2012	**28**
8	**British Journal of Educational Technology**	1970	1970	2012	2004	2011	**28**
9	**Information Processing & Management**	1994	1994	2012	1976	2012	**27**
10	**Journal of Information Science (JIS)**	1979	1979	2012	1995	2011	**27**
11	**Journal of Digital Information**	1997	1997	2012	2001	2012	**22**
12	**Computers & Education**	1976	1976	2012	1996	2012	**21**
13	**Journal of the American Society for Information Science and Technology**	1950	1950	2012	1999	2012	**21**
14	**International Journal of Technology Enhanced Learning (IJTEL)**	2008	2008	2012	2008	2011	**17**
15	**Journal of Computer Assisted Learning (JCAL)**	1994	1994	2012	1979	2012	**15**

No	Journal Title	1st Edition Year	Review Year Start	Review Year End	First Paper Found	Last Paper Found	# of papers identified
16	**IEEE Transactions on LT**	2008	2008	2012	2008	2012	**13**
17	**Journal of Library Trends**	2005	2005	2012	2007	2011	**13**
18	**Australasian Journal of Educational Technology (AJET)**	1985	1985	2012	2004	2011	**12**
19	**The Internet and Higher Education**	1998	1998	2012	2000	2012	**11**
20	**Canadian Journal of Learning & Technology**	1986	1986	2012	2002	2011	**11**
21	**Journal of Librarianship and Information Science**	1969	1969	2012	1979	2006	**10**
22	**Journal of Library Management**	1979	1979	2012	2004	2010	**9**
23	**Journal of Knowledge Management & E-Learning**	2009	2009	2012	2009	2011	**9**
24	**Code4lib Journal**	2007	2007	2012	2008	2012	**8**
25	**Advances in Library Administration and Organization**	2000	2000	2012	2001	2011	**8**
26	**Liber Quarterly**	1999	1999	2012	2003	2010	**6**
27	**Web Semantics: Science, Services and Agents on the World Wide Web**	2004	2004	2012	2004	2012	**6**
28	**International Journal of Metadata, Semantics and Ontologies (IJMSO)**	2006	2006	2012	2006	2006	**5**
29	**Interlending & Document Supply**	1973	1973	2012	1999	2010	**5**
30	**Data & Knowledge Engineering**	1990	1990	2012	2006	2010	**5**
31	**Innovations in Education & Teaching International**	1964	1964	2012	2005	2010	**4**

No	Journal Title	1st Edition Year	Review Year Start	Review Year End	First Paper Found	Last Paper Found	# of papers identified
32	**Cataloguing & Classification Quarterly**	2005	2005	2012	2010	2011	**3**
33	**Journal of Online Learning & Teaching (JOLT)**	2005	2005	2012	2004	2009	**3**
34	**Journal of Knowledge Management**	1997	1997	2012	2008	2008	**2**
35	**Journal of Library Innovation**	2010	2010	2012	2011	2011	**2**
36	**Journal of Systems and Information Technology**	1997	1997	2012	2009	2010	**2**
37	**VINE Journal**	1971	1971	2012	2004	2004	**1**
38	**Information Technology & People**	1982	1982	2012	1995	1995	**1**
39	**Journal of Library Metadata**	1997	1997	2012	-	-	**0**

Table B.1: Journals covered by this study

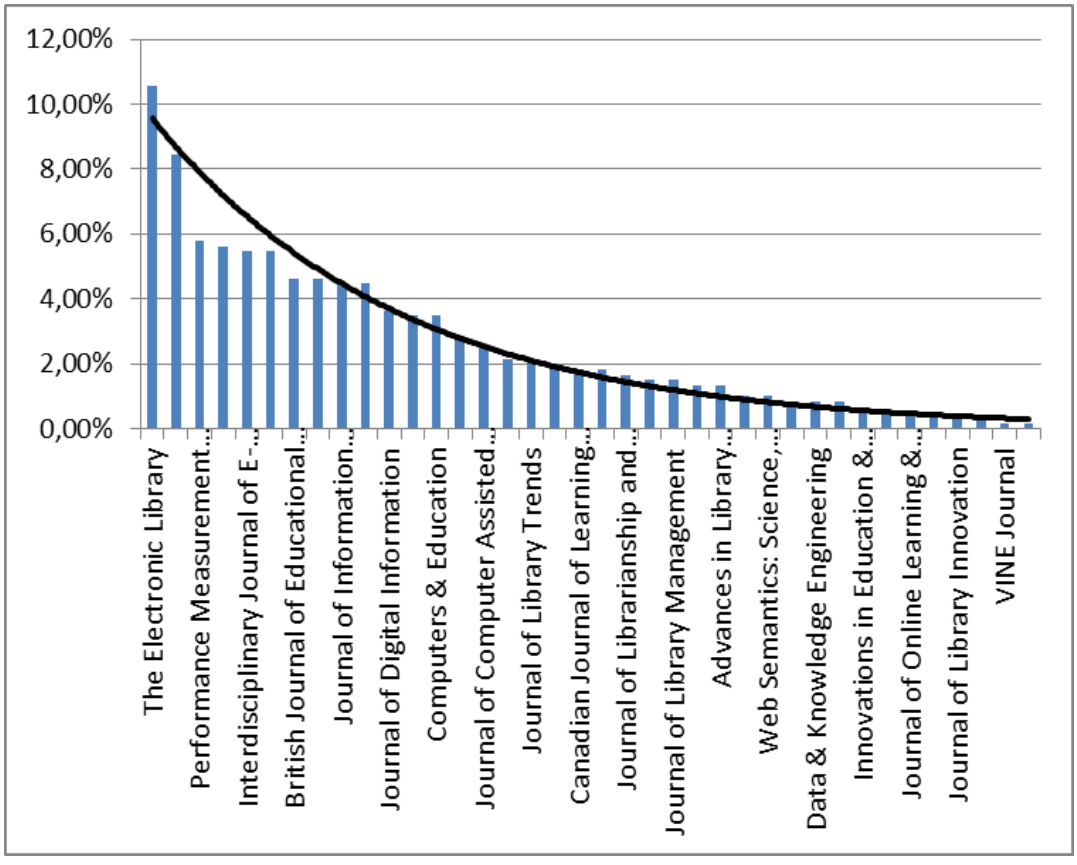


Figure B.1: Percentage of retrieved publications for the journals reviewed

No	Journal Title	Active Period	# of papers	Rate per Year	Accessible Papers
1	Program: electronic library and information systems	10	51	5.10	51
2	International Journal of Metadata, Semantics and Ontologies (IJMSO)	1	5	5.00	0
3	The Electronic Library	13	64	4.92	64
4	Interdisciplinary Journal of E-Learning and Learning Objects (IJELLO)	7	33	4.71	33
5	International Journal of Technology Enhanced Learning (IJTEL)	4	17	4.25	3
6	British Journal of Educational Technology	8	28	3.50	1
7	Online Information Review	10	34	3.40	34
8	Performance Measurement and Metrics	11	35	3.18	35
9	Journal of Knowledge Management & E-Learning	3	9	3.00	9
10	IEEE Transactions on LT	5	13	2.60	11
11	Journal of Library Trends	5	13	2.60	13
12	Library Hi Tech	12	28	2.33	28
13	D-Lib Magazine	16	33	2.06	33
14	Journal of Knowledge Management	1	2	2.00	2
15	Journal of Library Innovation	1	2	2.00	2
16	Journal of Library Trends	7	13	1.86	13
17	Journal of Digital Information	12	22	1.83	22
18	Journal of Computer Assisted Learning (JCAL)	9	15	1.67	6
19	Code4lib Journal	5	8	1.60	8
20	Journal of Information Science (JIS)	17	27	1.59	2
21	Journal of the American Society for Information Science and Technology	14	21	1.50	21
22	Australasian Journal of Educational Technology (AJET)	8	12	1.50	12
23	Cataloguing & Classification Quarterly	2	3	1.50	3
24	Journal of Library Management	7	9	1.29	9
25	Computers & Education	17	21	1.24	21

No	Journal Title	Active Period	# of papers	Rate per Year	Accessible Papers
26	Canadian Journal of Learning & Technology	10	11	1.10	11
27	Data & Knowledge Engineering	5	5	1.00	5
28	Journal of Systems and Information Technology	2	2	1.00	0
29	VINE Journal	1	1	1.00	1
30	Information Technology & People	1	1	1.00	0
31	The Internet and Higher Education	13	11	0.85	11
32	Liber Quarterly	8	6	0.75	6
33	Information Processing & Management	37	27	0.73	22
34	Advances in Library Administration and Organization	11	8	0.73	0
35	Web Semantics: Science, Services and Agents on the World Wide Web	9	6	0.67	6
36	Innovations in Education & Teaching International	6	4	0.67	0
37	Journal of Online Learning & Teaching (JOLT)	6	3	0.50	3
38	Interlending & Document Supply	12	5	0.42	5
39	Journal of Librarianship and Information Science	28	10	0.36	0

Table B.2: Statistics per journal

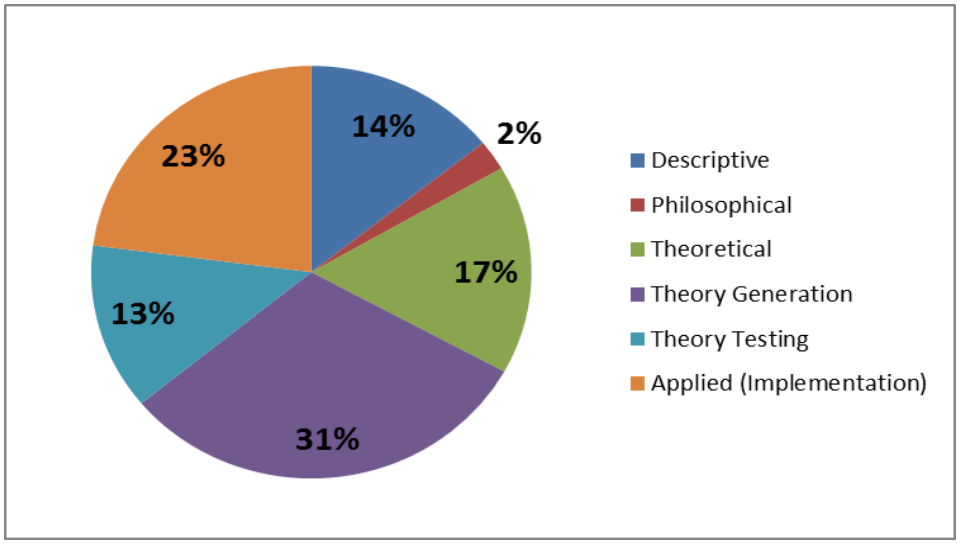


Figure B.2: Distribution of publications per Research Type

Type	# of papers	% of total papers
Descriptive	72	14,23%
Philosophical	12	2,37%
Theoretical	83	16,40%
Theory Generation	157	31,03%
Theory Testing	66	13,04%
Applied (Implementation)	116	22,92%
TOTAL	**506**	**100,0%**

Table B.3: Publications per Research Type

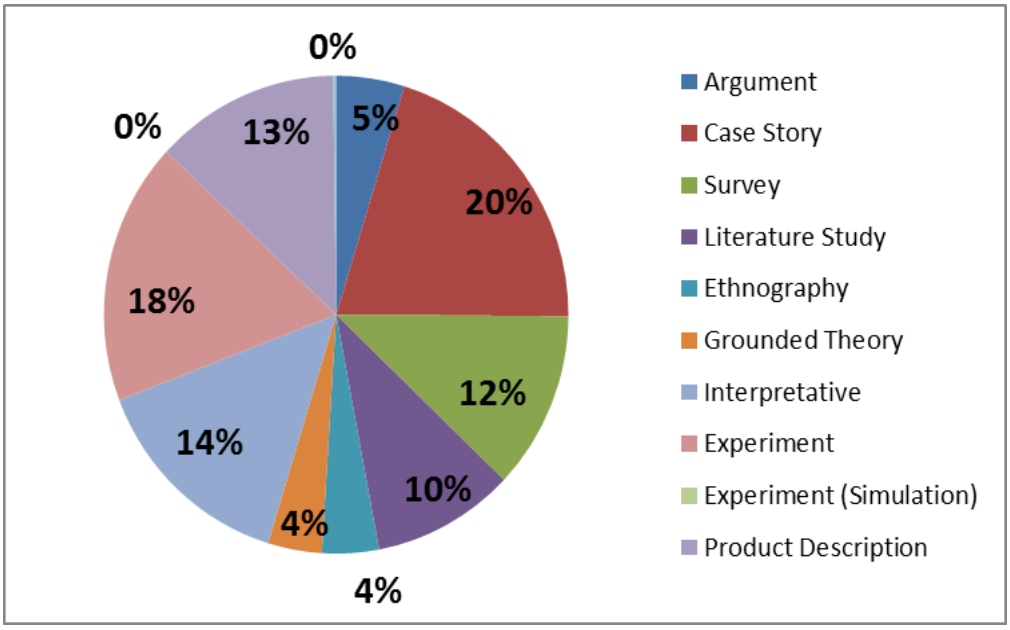


Figure B.3: Distribution of publications per Research Method

Method	# of papers	% of total papers
Case Story	103	20,36%
Experiment	90	17,79%
Interpretative	73	14,43%
Product Description	65	12,85%
Survey	61	12,06%
Literature Study	50	9,88%
Argument	24	4,74%
Ethnography	20	3,95%
Grounded Theory	19	3,75%
Unclear	1	0,20%
Experiment (Simulation)	0	0,00%
TOTAL	**506**	**100,0%**

Table B.4: Publications per Research Method

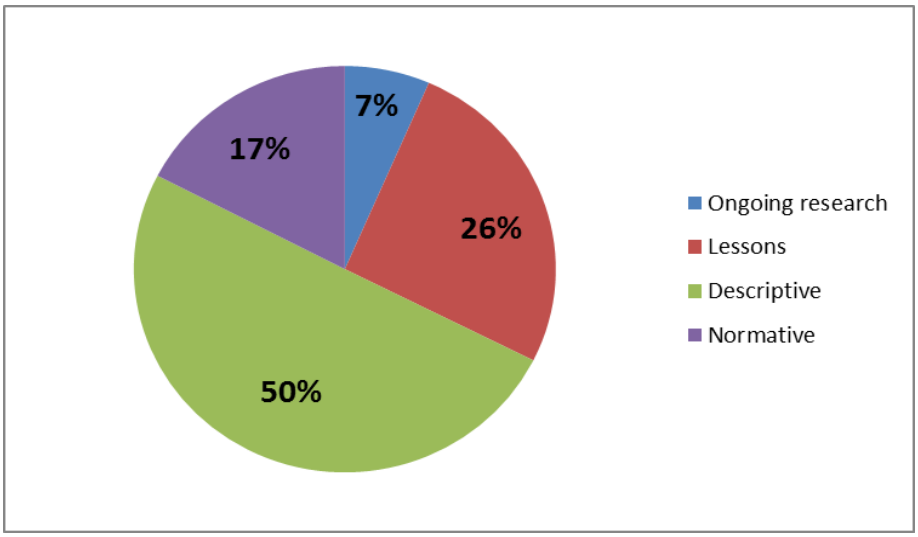


Figure B.4: Distribution of publications per Research Claim

Claim	# of papers	% of total papers
Ongoing Research	33	6,52%
Lessons	131	25,89%
Descriptive	254	50,20%
Normative	88	17,39%
TOTAL	**506**	**100,0%**

Table B.5: Publications per Claim

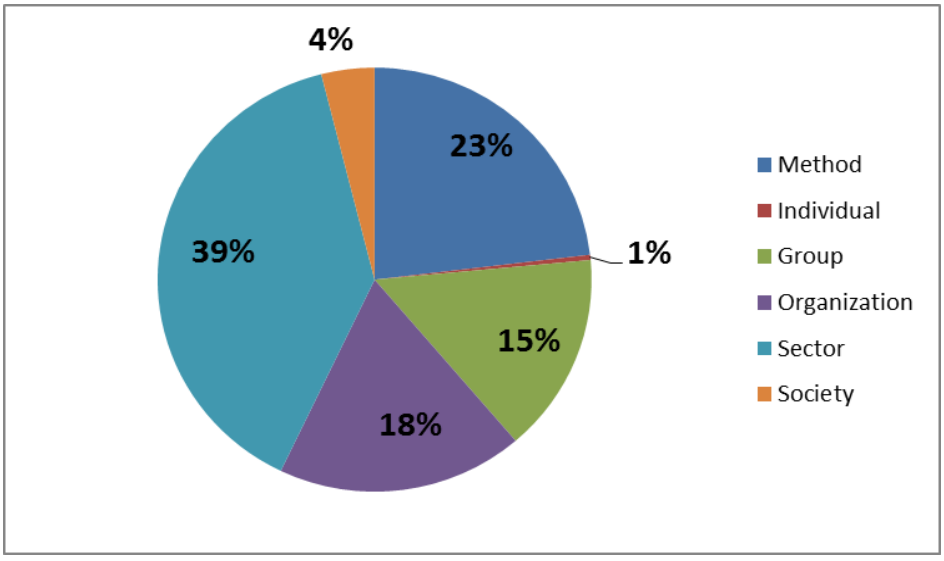


Figure B.5: Distribution of publications per Focus Unit

Focus Unit	# of papers	% of total papers
Method	117	23,12%
Individual	2	0,40%
Group	77	15,22%
Organization	93	18,38%
Sector	197	38,93%
Society	20	3,95%
TOTAL	**506**	**100,0%**

Table B.6: Publications per Focus Unit

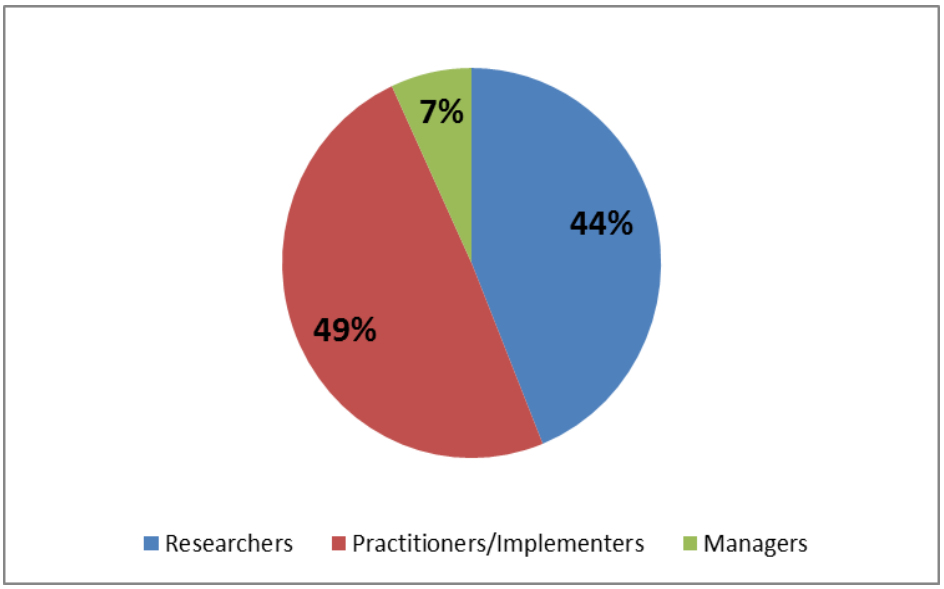


Figure B.6: Distribution of publications per Target Audience

Target Audience	# of papers	% of total papers
Researchers	222	43,87%
Practitioners/Implementers	249	49,21%
Managers	35	6,92%
Unclear	0	0,00%
TOTAL	**506**	**100,0%**

Table B.7: Publications per Target Audience

Category	No of Papers	%
Short (1-5]	10	2.0%
Medium-Short (5-10]	97	19.2%
Medium (10-15]	179	35.4%
Medium-Long (15-20]	81	16.0%
Long (20-U)	70	13.8%
Not Applicable	69	13.8%
TOTAL	**506**	**100%**

Table B.8: Publications per page length

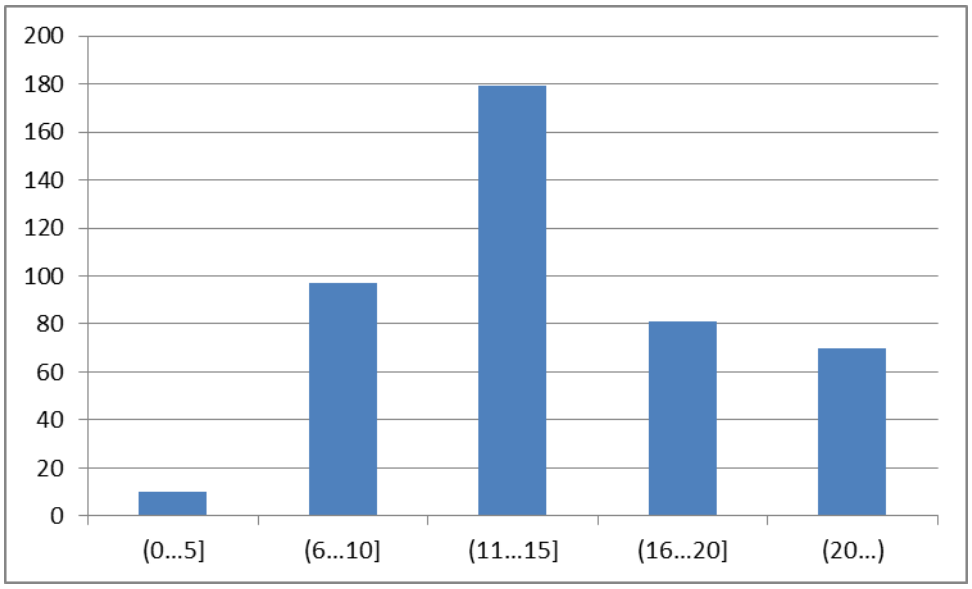


Figure B.7: Distribution of publications per page length

Category	No of References	%
0 to 10	107	21,1%
11 to 20	117	23,1%
21 to 30	129	25,5%
31 to 40	66	13,0%
41 to 50	48	9,5%
51 to 60	4	0,8%
60+	35	6,9%
TOTAL	**506**	**100%**

Table B.9: Publication distribution per references

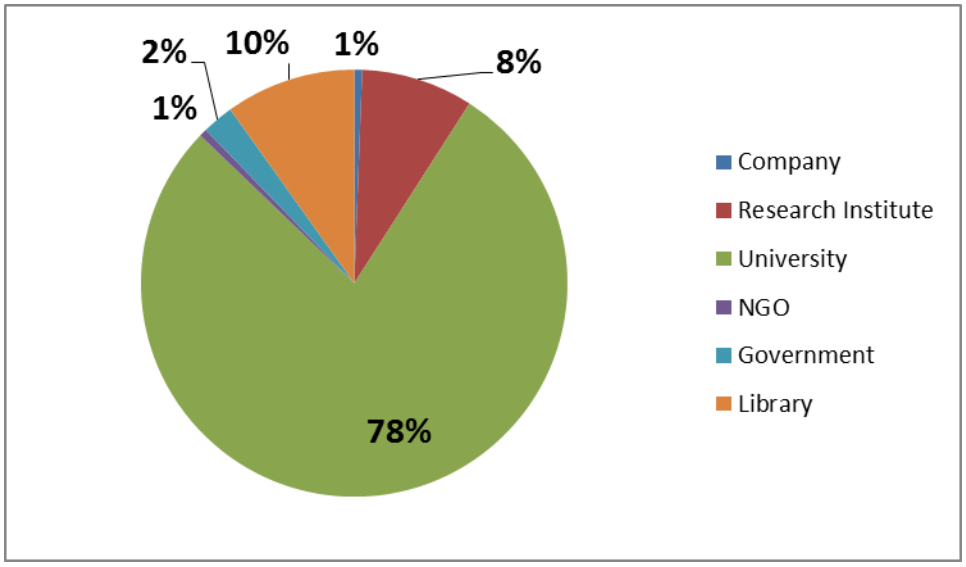


Figure B.8: Distribution of publications per Research Origin

Institution	# of papers	% of total papers
University	395	78,06%
Library	50	9,88%
Research Institute	43	8,50%
Government	12	2,37%
Company	3	0,59%
NGO	3	0,59%
TOTAL	**506**	**100,0%**

Table B.10: Publications per Research Origin

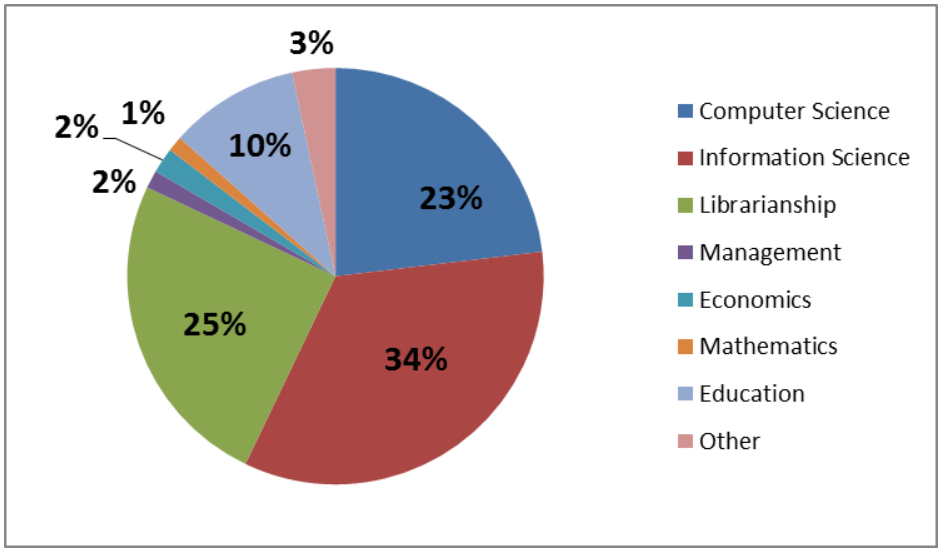


Figure B.9: Distribution of publications per Research Discipline

Disciplines	# of papers	% of total papers
Information Science	172	33,99%
Librarianship	126	24,90%
Computer Science	117	23,12%
Education	51	10,08%
Other	17	3,36%
Economics	10	1,98%
Management	7	1,38%
Mathematics	6	1,19%
TOTAL	**506**	**100%**

Table B.11: Publications per Discipline

No	Country	No of Papers	No	Country	No of Papers
1	USA	151	28	Ecuador	3
2	Great Britain	66	29	Nigeria	3
3	Canada	30	30	Singapore	3
4	Australia	21	31	Slovenia	3
5	Taiwan	18	32	South Korea	3
6	India	14	33	Bulgaria	2
7	Spain	14	34	Egypt	2
8	New Zealand	13	35	France	2
9	Greece	11	36	Ireland	2
10	China	10	37	Pakistan	2
11	Iran	10	38	Switzerland	2
12	Finland	9	39	Austria	1
13	Netherlands	9	40	Chile	1
14	Belgium	7	41	Kenya	1
15	Brazil	7	42	Korea	1
16	Japan	7	43	Mexico	1
17	Norway	7	44	Peru	1
18	Italy	6	45	Portugal	1
19	Turkey	6	46	Saudi Arabia	1
20	Malaysia	5	47	Trinidad & Tobago	1
21	Sweden	5	48	Uganda	1
22	Israel	4	49	United Arab Emirates	1
23	Poland	4	50	Uruguay	1
24	Serbia	4	51	Denmark	1
25	South Africa	4	52	Hong Kong	1
26	Bangladesh	3	53	Thailand	1
27	Croatia	3			

Table B.12: Distribution of publications per country of primary author

Continent	Number of publications	Percentage (%)
America	195	38.5%
Europe	174	34.4%
Asia	89	17.6%
Oceania	34	6.7%
Africa	14	2.8%

Table B.13: Distribution of publications per continent of primary author

ID	Journal Title	Descriptive	Philosophical	Theoretical	Theory Generation	Theory Testing	Applied
1	Journal of Knowledge Management	0	0	0	0	0	2
2	VINE Journal	0	0	0	0	0	1
3	Journal of Library Management	1	1	4	2	1	0
4	Code4lib Journal	1	0	0	1	1	5
5	Journal of Library Innovation	0	0	1	0	0	1
6	D-Lib Magazine	2	0	10	10	0	11
7	Liber Quarterly	1	1	1	2	0	1
8	Interdisciplinary Journal of E-Learning and Learning Objects (IJELLO)	0	0	9	10	7	7
9	Journal of Computer Assisted Learning (JCAL)	0	0	2	3	0	1
10	International Journal of Technology Enhanced Learning (IJTEL)	0	0	0	1	0	2
11	The Electronic Library	12	1	13	13	7	18
12	Interlending & Document Supply	2	1	1	0	0	1
13	Library Hi Tech	6	1	5	8	3	5
14	Online Information Review	4	0	1	15	8	6
15	Program: electronic library and information systems	24	3	3	9	5	7
16	Performance Measurement and Metrics	6	1	6	10	7	5
17	Information Processing & Management	1	0	0	10	4	6
18	Web Semantics: Science, Services and Agents on the World Wide Web	0	0	1	0	0	5
19	Data & Knowledge Engineering	0	0	1	1	0	3
20	Journal of Information Science (JIS)	0	0	3	7	2	5
21	IEEE Transactions on LT	1	0	2	5	0	3
22	British Journal of Educational Technology	0	0	0	1	0	0
23	The Internet and Higher Education	2	2	2	3	1	1

ID	Journal Title	Descriptive	Philosophical	Theoretical	Theory Generation	Theory Testing	Applied
24	Australasian Journal of Educational Technology (AJET)	0	0	0	4	2	5
25	Computers & Education	1	0	3	6	8	3
26	Journal of the American Society for Information Science and Technology	3	0	1	11	5	1
27	Canadian Journal of Learning & Technology	1	0	2	5	2	1
28	Journal of Digital Information	2	0	6	7	2	5
29	Cataloguing & Classification Quarterly	0	0	0	3	0	0
30	Journal of Online Learning & Teaching (JOLT)	0	0	0	3	0	0
31	Journal of Knowledge Management & E-Learning	0	0	1	3	1	4
32	Journal of Library Trends	2	1	5	4	0	1

Table B.14: Publications per Research Type per Journal

No	Journal Title	Argument	Case Story	Survey	Literature Study	Ethnography	Grounded Theory	Interpretative	Experiment	Experiment (Simulation)	Product Description
1	Journal of Knowledge Management	0	1	0	0	0	0	0	0	0	1
2	VINE Journal	0	1	0	0	0	0	0	0	0	0
3	Journal of Library Management	3	3	1	1	0	0	1	0	0	0
4	Code4lib Journal	0	1	0	0	0	1	1	0	0	5
5	Journal of Library Innovation	0	2	0	0	0	0	0	0	0	0
6	D-Lib Magazine	2	9	0	5	0	2	11	3	0	1
7	Liber Quarterly	1	2	1	1	0	0	0	1	0	0
8	Interdisciplinary	0	3	2	8	2	1	6	8	0	3

No	Journal Title	Argument	Case Story	Survey	Literature Study	Ethnography	Grounded Theory	Interpretative	Experiment	Experiment (Simulation)	Product Description
	Journal of E-Learning and Learning Objects (IJELLO)										
9	**Journal of Computer Assisted Learning (JCAL)**	0	2	0	1	1	0	1	1	0	0
10	**International Journal of Technology Enhanced Learning (IJTEL)**	0	0	0	0	0	0	1	2	0	0
11	**The Electronic Library**	4	18	12	7	0	0	3	13	0	7
12	**Interlending & Document Supply**	1	4	0	0	0	0	0	0	0	0
13	**Library Hi Tech**	1	9	4	2	2	0	1	3	0	6
14	**Online Information Review**	0	5	8	5	0	0	4	10	0	2
15	**Program: electronic library and information systems**	1	15	9	2	4	1	4	6	0	9
16	**Performance Measurement and Metrics**	3	2	3	4	2	1	6	12	0	2
17	**Information Processing & Management**	0	1	0	0	1	3	7	6	0	3

No	Journal Title	Argument	Case Story	Survey	Literature Study	Ethnography	Grounded Theory	Interpretative	Experiment	Experiment (Simulation)	Product Description
18	Web Semantics: Science, Services and Agents on the World Wide Web	0	0	0	1	0	1	0	0	0	3
19	Data & Knowledge Engineering	0	0	0	0	0	3	0	0	0	2
20	Journal of Information Science (JIS)	1	2	2	2	0	0	4	3	0	3
21	IEEE Transactions on LT	0	1	0	1	0	1	3	2	0	3
22	British Journal of Educational Technology	0	0	0	1	0	0	0	0	0	0
23	The Internet and Higher Education	4	2	0	2	0	0	1	2	0	0
24	Australasian Journal of Educational Technology (AJET)	0	4	2	0	3	0	1	1	0	0
25	Computers & Education	0	2	2	1	0	1	3	9	0	3
26	Journal of the American Society for Information Science and Technology	0	1	4	1	2	2	6	2	0	3
27	Canadian Journal of Learning & Technology	0	2	4	1	0	0	1	2	0	1

No	Journal Title	Argument	Case Story	Survey	Literature Study	Ethnography	Grounded Theory	Interpretative	Experiment	Experiment (Simulation)	Product Description
28	Journal of Digital Information	1	8	2	0	0	1	4	2	0	4
29	Cataloguing & Classification Quarterly	0	0	2	0	0	1	0	0	0	0
30	Journal of Online Learning & Teaching (JOLT)	0	0	0	0	0	0	2	1	0	0
31	Journal of Knowledge Management & E-Learning	0	0	0	1	3	0	1	1	0	3
32	Journal of Library Trends	2	3	3	3	0	0	1	0	0	1

Table B.15: Distribution of publications per Research Method and per journal

No	Journal Title	Ongoing Research	Lessons	Descriptive	Normative
1	Journal of Knowledge Management	1	0	1	0
2	VINE Journal	0	1	0	0
3	Journal of Library Management	0	1	6	2
4	Code4lib Journal	0	0	6	2
5	Journal of Library Innovation	0	0	2	0
6	D-Lib Magazine	0	4	19	10
7	Liber Quarterly	1	4	1	0
8	Interdisciplinary Journal of E-Learning and Learning Objects (IJELLO)	2	4	22	5
9	Journal of Computer Assisted Learning (JCAL)	0	3	3	0
10	International Journal of Technology Enhanced Learning (IJTEL)	0	1	1	1
11	The Electronic Library	9	22	26	7

No	Journal Title	Ongoing Research	Lessons	Descriptive	Normative
12	Interlending & Document Supply	0	3	2	0
13	Library Hi Tech	4	6	15	3
14	Online Information Review	1	5	21	7
15	Program: electronic library and information systems	6	21	21	3
16	Performance Measurement and Metrics	2	8	21	4
17	Information Processing & Management	1	4	11	5
18	Web Semantics: Science, Services and Agents on the World Wide Web	0	4	1	1
19	Data & Knowledge Engineering	2	1	1	1
20	Journal of Information Science (JIS)	1	0	10	6
21	IEEE Transactions on LT	0	3	5	3
22	British Journal of Educational Technology	0	0	0	1
23	The Internet and Higher Education	0	6	3	2
24	Australasian Journal of Educational Technology (AJET)	0	5	2	4
25	Computers & Education	1	3	11	6
26	Journal of the American Society for Information Science and Technology	0	6	9	6
27	Canadian Journal of Learning & Technology	0	1	6	4
28	Journal of Digital Information	1	12	9	0
29	Cataloguing & Classification Quarterly	0	0	3	0
30	Journal of Online Learning & Teaching (JOLT)	0	0	3	0
31	Journal of Knowledge Management & E-Learning	0	1	8	0
32	Journal of Library Trends	1	2	5	5

Table B.16: Distribution of publications per Research Claim and per journal

No	Journal Title	Method	Individual	Group	Organization	Sector	Society
1	Journal of Knowledge Management	1	0	1	0	0	0
2	VINE Journal	1	0	0	0	0	0
3	Journal of Library Management	1	1	1	4	2	0

No	Journal Title	Method	Individual	Group	Organization	Sector	Society
4	Code4lib Journal	2	0	0	4	2	0
5	Journal of Library Innovation	0	0	0	1	1	0
6	D-Lib Magazine	3	0	1	6	23	0
7	Liber Quarterly	1	0	0	2	2	1
8	Interdisciplinary Journal of E-Learning and Learning Objects (IJELLO)	8	0	8	3	14	0
9	Journal of Computer Assisted Learning (JCAL)	1	0	1	1	3	0
10	International Journal of Technology Enhanced Learning (IJTEL)	0	0	1	0	2	0
11	The Electronic Library	13	0	3	16	32	0
12	Interlending & Document Supply	0	0	0	2	3	0
13	Library Hi Tech	4	0	2	8	14	0
14	Online Information Review	11	0	4	3	15	1
15	Program: electronic library and information systems	16	0	4	18	8	5
16	Performance Measurement and Metrics	7	1	14	3	9	1
17	Information Processing & Management	10	0	6	0	5	0
18	Web Semantics: Science, Services and Agents on the World Wide Web	4	0	0	0	2	0
19	Data & Knowledge Engineering	5	0	0	0	0	0
20	Journal of Information Science (JIS)	7	0	2	0	7	1
21	IEEE Transactions on LT	4	0	1	1	4	1
22	British Journal of Educational Technology	0	0	0	0	0	1
23	The Internet and Higher Education	2	0	0	1	6	2
24	Australasian Journal of Educational Technology (AJET)	2	0	3	2	3	1
25	Computers & Education	0	0	3	3	13	2
26	Journal of the American Society for Information Science and Technology	3	0	5	4	8	1

No	Journal Title	Method	Individual	Group	Organization	Sector	Society
27	Canadian Journal of Learning & Technology	2	0	4	0	3	2
28	Journal of Digital Information	7	0	3	5	6	1
29	Cataloguing & Classification Quarterly	0	0	1	0	2	0
30	Journal of Online Learning & Teaching (JOLT)	0	0	2	1	0	0
31	Journal of Knowledge Management & E-Learning	2	0	4	2	1	0
32	Journal of Library Trends	0	0	3	3	7	0

Table B.17: Distribution of publications per Focus Unit and per journal

ID	Journal Title	Researchers	Practitioners/Implementers	Managers	Unclear
1	Journal of Knowledge Management	1	1	0	0
2	VINE Journal	0	1	0	0
3	Journal of Library Management	2	4	3	0
4	Code4lib Journal	1	7	0	0
5	Journal of Library Innovation	0	1	1	0
6	D-Lib Magazine	17	16	0	0
7	Liber Quarterly	1	3	2	0
8	Interdisciplinary Journal of E-Learning and Learning Objects (IJELLO)	25	8	0	0
9	Journal of Computer Assisted Learning (JCAL)	3	3	0	0
10	International Journal of Technology Enhanced Learning (IJTEL)	1	2	0	0
11	The Electronic Library	29	30	5	0
12	Interlending & Document Supply	2	3	0	0
13	Library Hi Tech	6	21	1	0
14	Online Information Review	7	26	1	0
15	Program: electronic library and information systems	12	36	3	0

ID	Journal Title	Researchers	Practitioners/Implementers	Managers	Unclear
16	Performance Measurement and Metrics	15	12	8	0
17	Information Processing & Management	11	9	1	0
18	Web Semantics: Science, Services and Agents on the World Wide Web	2	4	0	0
19	Data & Knowledge Engineering	2	3	0	0
20	Journal of Information Science (JIS)	6	6	5	0
21	IEEE Transactions on LT	7	4	0	0
22	British Journal of Educational Technology	1	0	0	0
23	The Internet and Higher Education	10	1	0	0
24	Australasian Journal of Educational Technology (AJET)	7	4	0	0
25	Computers & Education	15	6	0	0
26	Journal of the American Society for Information Science and Technology	10	10	1	0
27	Canadian Journal of Learning & Technology	7	4	0	0
28	Journal of Digital Information	5	13	4	0
29	Cataloguing & Classification Quarterly	2	1	0	0
30	Journal of Online Learning & Teaching (JOLT)	2	1	0	0
31	Journal of Knowledge Management & E-Learning	5	4	0	0
32	Journal of Library Trends	8	5	0	0

Table B.18: Distribution of publications per Target Audience and per journal

ID	Journal Title	Company	Research Institute	University	NGO	Government	Library
1	Journal of Knowledge Management	0	0	2	0	0	0
2	VINE Journal	0	0	1	0	0	0
3	Journal of Library Management	0	2	7	0	0	0
4	Code4lib Journal	1	0	5	0	0	2
5	Journal of Library Innovation	0	0	1	0	0	1

ID	Journal Title	Company	Research Institute	University	NGO	Government	Library
6	D-Lib Magazine	0	4	24	1	0	4
7	Liber Quarterly	0	0	4	0	0	2
8	Interdisciplinary Journal of E-Learning and Learning Objects (IJELLO)	0	0	33	0	0	0
9	Journal of Computer Assisted Learning (JCAL)	0	3	3	0	0	0
10	International Journal of Technology Enhanced Learning (IJTEL)	0	1	2	0	0	0
11	The Electronic Library	0	8	47	0	1	8
12	Interlending & Document Supply	0	1	2	0	0	2
13	Library Hi Tech	0	2	12	0	0	14
14	Online Information Review	0	2	27	1	3	1
15	Program: electronic library and information systems	0	4	35	0	5	7
16	Performance Measurement and Metrics	1	2	28	0	1	3
17	Information Processing & Management	0	2	19	0	0	0
18	Web Semantics: Science, Services and Agents on the World Wide Web	0	3	3	0	0	0
19	Data & Knowledge Engineering	0	0	5	0	0	0
20	Journal of Information Science (JIS)	1	2	14	0	0	0
21	IEEE Transactions on LT	0	0	11	0	0	0
22	British Journal of Educational Technology	0	0	1	0	0	0
23	The Internet and Higher Education	0	3	8	0	0	0
24	Australasian Journal of Educational Technology (AJET)	0	1	10	0	0	0
25	Computers & Education	0	0	21	0	0	0
26	Journal of the American Society for Information Science and Technology	1	1	19	0	0	0
27	Canadian Journal of Learning & Technology	0	0	11	0	0	0
28	Journal of Digital Information	0	0	16	1	2	3

ID	Journal Title	Company	Research Institute	University	NGO	Government	Library
29	Cataloguing & Classification Quarterly	0	0	2	0	0	1
30	Journal of Online Learning & Teaching (JOLT)	0	0	3	0	0	0
31	Journal of Knowledge Management & E-Learning	0	0	9	0	0	0
32	Journal of Library Trends	0	2	10	0	0	1

Table B.19: Distributions per Type of Institute and per journal

ID	Journal Title	Computer Science	Information Science	Management	Economics	Mathematics	Education	Other	Librarianship
1	Journal of Knowledge Management	2	0	0	0	0	0	0	0
2	VINE Journal	0	1	0	0	0	0	0	0
3	Journal of Library Management	0	8	0	1	0	0	0	0
4	Code4lib Journal	6	1	0	0	0	0	0	1
5	Journal of Library Innovation	0	1	0	0	0	0	0	1
6	D-Lib Magazine	11	16	0	1	0	0	0	5
7	Liber Quarterly	0	3	0	1	0	0	0	2
8	Interdisciplinary Journal of E-Learning and Learning Objects (IJELLO)	11	11	0	0	0	10	0	1
9	Journal of Computer Assisted Learning (JCAL)	0	1	0	0	0	5	0	0
10	International Journal of Technology Enhanced Learning (IJTEL)	2	1	0	0	0	0	0	0

ID	Journal Title	Computer Science	Information Science	Management	Economics	Mathematics	Education	Other	Librarianship
11	**The Electronic Library**	12	19	0	0	1	1	0	31
12	**Interlending & Document Supply**	0	3	0	0	0	0	0	2
13	**Library Hi Tech**	6	5	0	0	0	0	1	16
14	**Online Information Review**	11	15	1	1	1	1	0	4
15	**Program: electronic library and information systems**	13	13	0	2	1	0	3	19
16	**Performance Measurement and Metrics**	1	12	0	1	1	0	2	18
17	**Information Processing & Management**	7	8	0	1	0	0	1	4
18	**Web Semantics: Science, Services and Agents on the World Wide Web**	2	3	0	0	0	0	0	1
19	**Data & Knowledge Engineering**	2	3	0	0	0	0	0	0
20	**Journal of Information Science (JIS)**	4	10	0	0	0	0	0	3
21	**IEEE Transactions on LT**	8	1	0	0	0	1	1	0
22	**British Journal of Educational Technology**	1	0	0	0	0	0	0	0
23	**The Internet and Higher Education**	0	2	1	0	0	8	0	0

ID	Journal Title	Computer Science	Information Science	Management	Economics	Mathematics	Education	Other	Librarianship
24	Australasian Journal of Educational Technology (AJET)	1	3	0	0	0	6	1	0
25	Computers & Education	4	1	1	1	1	7	6	0
26	Journal of the American Society for Information Science and Technology	3	11	1	0	0	0	0	6
27	Canadian Journal of Learning & Technology	0	2	1	0	1	7	0	0
28	Journal of Digital Information	3	11	0	1	0	1	2	4
29	Cataloguing & Classification Quarterly	0	1	0	0	0	0	0	2
30	Journal of Online Learning & Teaching (JOLT)	0	0	1	0	0	2	0	0
31	Journal of Knowledge Management & E-Learning	5	1	1	0	0	2	0	0
32	Journal of Library Trends	2	5	0	0	0	0	0	6

Table B.20: Distribution of publications per Discipline and per journal

Appendix C: Metadata Quality Assurance Certification Process Documentation & Instruments

PHASES	STEPS	QUALITY ASSURANCE METHODS	QUALITY TOOLS / INSTRUMENTS	OUTCOMES	QUALITY ACTORS
A. **Metadata Design**	AP development: • *Definition of own requirements,* • *Selection of elements,* • *Semantics refinement,* • *Required extensions,* • *Application profile binding,* • *Evaluation of AP*	Focus group with domain experts that participate in preliminary hands-on exercise	Application Profile print out form to complete (with metadata elements)	Completed paper-based metadata records	Metadata experts & Domain experts
		Focus group with domain experts that participate in metadata understanding session	Element assessment form	Input for application profile revision	
B. **Testing**	Test implementation • Implement in test environment/tool	Metadata Quality Review of test sample of resources	Pre-check/Core Metadata Quality Criteria[1]	Good & Bad metadata practices guide	Metadata experts & Content Annotators
	Hands-on annotation • Hands-on annotation of sample of resources				
C. **Calibration**	Implement in controlled environment • Hands-on annotation of representative sample of resources	Metadata Quality Peer Review of representative sample of resources	Metadata Quality Assessment Grid[1]	Targeted comments to Content Providers	Metadata experts & Domain experts
D. **Critical Mass**	Implement in final environment/tool • Intensive annotation of critical mass of content	Analysis of Usage Data from Annotation Tools	Completeness Metrics[1]	Recommendations for Content Providers & Metadata AP	Metadata experts
		Introduce Quality Certification of each provider in metadata	Validation information and "Conforms To" pointer in metadata records	Quality Mark	Content Providers
E. **Regular Operation**	Regular annotation • Content providers provide new resources	Analysis of Usage Data from Annotation Tool	Completeness Metrics Form / Grid	Recommendations for Content Providers	Metadata experts & Content users/ consumers
		Online Peer Review of metadata	Core Criteria / Quality of metadata / "Conforms To"		

PHASES	STEPS	QUALITY ASSURANCE METHODS	QUALITY TOOLS / INSTRUMENTS	OUTCOMES	QUALITY ACTORS
	• Smaller numbers of resources being added than in "critical mass"	Quality Prizes & Awards	Rating Mechanism (*Topic relevance / Educational Usefulness / Metadata*)	Quality Mark	Content Providers
		Quality Certification in Metadata	Validation information and "Conforms To" pointer in metadata records		

Table C.1: Overview of Metadata Quality Assessment Certification Process (MQACP)

Appendix D: Data from Learning Federations Experiment

Metadata Understanding Session Data

Element	Is this element easy to understand?					Is it useful for describing Organic.Edunet content resources?					Is the selection of its possible values clear and appropriate?					Should it be mandatory / recommended / optional???		
1. General																		
1.1 Identifier	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**	**M**	**R**	**O**
1.1.1 Catalog	3	0	0	1	8	2	0	0	0	6	2	0	1	0	5	8	1	0
1.1.2 Entry	3	0	0	3	6	2	0	0	0	6	2	0	1	0	5	9	0	0
1.2 Title	3	0	0	1	11	3	0	0	0	13	3	0	1	0	11	16	0	0
1.3 Language	3	0	0	1	11	3	0	0	0	13	3	0	1	0	11	17	0	0
1.4 Description	3	0	0	0	12	3	0	0	3	10	2	1	0	0	12	12	4	0
1.5 Keyword	2	1	1	1	12	3	0	6	1	6	3	0	2	3	7	10	3	2
1.6 Coverage	2	2	2	3	6	1	1	5	4	3	3	3	3	3	1	1	4	9
1.7 Structure	6	0	5	3	3	3	2	6	1	3	6	1	5	1	3	0	5	10
1.8 Aggregation Level	6	1	4	3	3	2	4	5	3	1	3	3	6	2	1	0	7	9
2. Life Cycle																		
2.1 Version	2	0	2	3	8	0	3	5	5	2	1	1	3	4	5	2	5	9
2.2 Status	3	0	1	4	7	0	1	4	7	2	1	1	1	4	6	1	7	7
2.3 Contribute																		
2.3.1 Role	3	1	3	4	6	3	0	3	5	5	4	1	0	7	4	6	7	4
2.3.2 Entity	2	2	1	6	6	2	1	3	3	7	2	2	1	7	4	5	8	4
2.3.3 Date	3	1	2	1	10	3	1	3	3	5	4	3	0	2	6	5	7	3
3. Meta-Metadata																		
3.1 Identifier																		
3.1.1 Catalog	4	0	2	1	5	3	3	1	0	4	2	2	2	0	5	3	3	3
3.1.2 Entry	4	0	3	1	4	3	3	1	0	4	2	2	2	0	5	3	3	3

Element	Is this element easy to understand?					Is it useful for describing Organic.Edunet content resources?					Is the selection of its possible values clear and appropriate?					Should it be mandatory / recommended / optional???				
3.2 Contribute																				
3.2.1 Role	2	1	2	1	8	3	2	2	2	5	3	0	2	2	6	4	8	2		
3.2.2 Entity	2	2	1	2	7	3	3	1	2	5	3	1	2	2	5	3	9	2		
3.2.3 Date	2	1	0	1	10	3	2	1	3	5	3	0	1	2	7	4	7	2		
3.3 Metadata Schema	2	2	0	2	6	3	3	1	0	5	3	0	1	2	5	2	4	4		
3.4 Language	2	0	0	0	9	2	3	0	0	6	2	0	0	0	8	4	3	2		
4. Technical																				
4.1 Format	2	0	0	0	14	0	0	3	2	10	0	0	2	3	8	10		5	1	
4.2 Size	2	0	0	0	14	1	0	2	3	9	1	0	1	3	9	9		6	1	
4.3 Location	2	0	2	0	11	1	4	1	2	6	1	1	1	4	5	6		5	3	
4.4 Requirement																				
4.4.1 OrComposite																				
4.4.1.1 Type	2	1	5	2	7	4	3	1	4	4	4	0	3	4	3	2		8	6	
4.4.1.2 Name	2	2	3	3	7	4	3	3	2	4	3	1	4	4	3	3		6	7	
4.4.1.3 Minimum Version	4	0	2	3	8	4	3	2	5	2	3	0	5	6	1	1		4	11	
4.4.1.4 Maximum Version	4	0	2	2	8	4	4	2	4	2	3	0	5	6	1	1		4	11	
4.5 Installation Remarks	2	1	2	5	6	3	2	3	4	3	3	1	2	7	1	1		4	12	
4.6 Other Platform Requirements	2	1	3	4	6	4	2	2	3	3	2	1	4	5	1	1		4	10	
4.7 Duration	2	0	1	3	6	3	0	1	3	5	3	4	4	1	1	1		5	8	
5. Educational																				
5.1 Interactivity Type	3	3	2	2	5	2	2	3	1	6	3	2	3	3	5	1		8	6	

Element	Is this element easy to understand?					Is it useful for describing Organic.Edunet content resources?					Is the selection of its possible values clear and appropriate?					Should it be mandatory / recommended / optional???		
5.2 Learning Resource Type	3	1	0	5	7	2	0	2	4	7	1	2	2	4	8	4	11	0
5.3 Interactivity Level	4	3	3	3	2	2	4	5	3	0	3	3	5	3	2	0	6	9
5.4 Semantic Density	5	5	2	1	2	3	5	2	4	1	3	4	3	5	1	0	3	11
5.5 Intended End User Role	3	1	2	4	6	1	5	2	2	5	0	6	1	4	4	2	10	4
5.6 Context	3	0	3	2	8	1	5	0	4	5	2	4	1	2	7	2	11	3
5.7 Typical Age Range	4	0	0	4	8	2	2	5	4	2	2	3	3	2	6	0	9	7
5.8 Difficulty	4	0	3	3	5	1	1	5	4	1	2	3	4	3	4	0	5	10
5.9 Typical Learning Time	4	0	3	4	4	3	2	4	2	1	4	1	3	1	4	0	5	9
5.10 Description	4	1	2	2	3	3	0	4	1	2	4	1	3	1	4	1	2	8
5.11 Language	5	0	0	2	7	3	1	2	2	4	2	2	2	2	4	1	7	6
6. Rights																		
6.1 Cost	3	0	1	0	10	2	0	0	4	7	1	0	1	3	9	8	7	2
6.2 Copyright and Other Restrictions	3	0	2	0	9	3	0	0	2	8	2	1	2	1	8	11	4	1
6.3 Description	3	1	2	0	8	3	0	2	3	5	2	2	2	1	7	6	7	3
7. Relation																		
7.1 Kind	2	0	1	5	4		1	2	6	1	0	3	5	1	4		4	10
7.2 esource																		
7.2.1 Identifier																		
7.2.1.1 Catalog	1	0	3	1	3	0	0	1	4	2	1	0	2	2	4	2	1	9
7.2.1.2 Entry	1	1	2	2	3	1	1	2	1	3	2	1	3	1	3	3	1	9
7.2.2 Description	1	2	2	2	2	0	1	3	2	2	1	1	5	1	2	0	3	10

Element	Is this element easy to understand?					Is it useful for describing Organic.Edunet content resources?					Is the selection of its possible values clear and appropriate?					Should it be mandatory / recommended / optional???				
8. Annotation																				
8.1 Entity	1	0	2	4	4	1	1	2	2	1	1	1	2	2	5	2		1		9
8.2 Date	1	0	1	3	6	0	1	4	1	1	0	0	4	3	4	1		2		9
8.3 Description	1	0	1	3	6	0	0	5	0	2	0	0	5	2	4	0		3		9
9. Classification																				
9.1 Purpose	2	0	2	1	5	3	0	2	1	2	4	0	3	1	2	3		4		5
9.2 Taxon Path																				
9.2.1 Source																				
9.2.2 Taxon																				
9.2.2.1 Id	2	1	0	2	2	1	2	0	2	1	3	1	2	1	1	3		3		4
9.2.2.2 Entry	2	1	0	2	2	1	2	1	1	1	3	1	1	2	1	3		3		4
9.3 Description	2	2	1	2	4	2	2	1	2	3	4	1	2	1	4	4		3		5
9.4 Keyword	2	1	2	1	5	2	1	1	2	4	3	1	1	3	4	5		3		5

Metadata Record Peer Review Data

Reviewer	URL	1. In which degree is this metadata record completed?	2. Overall accuracy of the metadata provided	3. Values provided consistent to metadata standard	4. Describe the resource in an objective way?	5. Values provided, appropriate for the use in the Portal?	6. Degree of correctness of the language used	7. Overall score for the metadata of this resource	Publish ?	Comments ?
A. Steen-Holm	here	4	5	4	5	4	4	5	No	Yes
A. Steen-Holm	here	4	5	5	5	5	5	5	No	Yes
A. Steen-Holm	here	3	5	5	5	3	5	4	Yes	Yes
A. Steen-Holm	here	2	4	5	5	3	5	3	Yes	Yes
A. Steen-Holm	here	2	4	5	4	3	5	4	Yes	Yes

Reviewer	URL	1. In which degree is this metadata record completed?	2. Overall accuracy of the metadata provided	3. Values provided consistent to metadata standard	4. Describe the resource in an objective way?	5. Values provided, appropriate for the use in the Portal?	6. Degree of correctness of the language used	7. Overall score for the metadata of this resource	Publish ?	Comments ?
T. Jasinski	here	5	5	5	5	5	5	5	Yes	Yes
T. Jasinski	here	5	5	5	5	5	5	5	Yes	Yes
T. Jasinski	here	5	5	5	5	5	5	5	Yes	Yes
T. Jasinski	here	5	5	5	5	5	5	5	Yes	Yes
T. Jasinski	here	5	5	5	5	5	5	5	Yes	Yes
G. Lieblein	here	4	5	5	5	4	5	5	No	Yes
G. Lieblein	here	4	5	5	5	5	5	5	No	Yes
G. Lieblein	here	2	4	5	5	1	5	3	No	Yes
G. Lieblein	here	3	5	5	5	3	5	4	No	Yes
G. Lieblein	here	2	5	4	5	4	5	4	No	Yes
D. Rodriguez	here	5	5	5	5	5	5	5	Yes	Yes
D. Rodriguez	here	5	5	5	5	5	5	5	Yes	Yes
D. Rodriguez	here	-	-	-	-	-	-	-	No	No
D. Rodriguez	here	5	4	5	5	4	5	4	Yes	Yes
D. Rodriguez	here	5	4	5	5	4	5	5	Yes	Yes
C. Wagner-Alt	here	5	5	5	5	5	5	5	Yes	No
C. Wagner-Alt	here	5	5	5	5	5	4	5	Yes	No
C. Wagner-Alt	here	4	4	5	4	5	3	4	Yes	No
C. Wagner-Alt	here	5	5	5	5	5	5	5	Yes	No
C. Wagner-Alt	here	5	5	5	5	5	5	5	Yes	No
E. Chryssafidou	N//A	5	5	5	5	5	5	5	Yes	No
E. Chryssafidou	N//A	5	5	5	4	5	5	5	Yes	Yes
E. Chryssafidou	N//A	5	5	5	5	5	5	5	Yes	No
E. Chryssafidou	N//A	5	5	5	4	5	5	-	No	Yes

Reviewer	URL	1. In which degree is this metadata record completed?	2. Overall accuracy of the metadata provided	3. Values provided consistent to metadata standard	4. Describe the resource in an objective way?	5. Values provided, appropriate for the use in the Portal?	6. Degree of correctness of the language used	7. Overall score for the metadata of this resource	Publish ?	Comments ?
E. Chryssafidou	N//A	4	2	5	2	4	4	3	No	Yes
A. Kaasik	here	4	3	3	4	4	3	3	Yes	Yes
A. Kaasik	here	4	5	4	5	4	3	4	Yes	Yes
A. Kaasik	here	5	5	4	5	4	3	4	Yes	Yes
A. Kaasik	here	5	5	5	3	5	4	4	Yes	No
A. Kaasik	here	4	5	4	5	5	5	5	Yes	No
J. Wickham	here	5	4	-	5	5	5	5	Yes	Yes
J. Wickham	here	5	4	-	5	5	5	5	Yes	Yes
J. Wickham	here	5	4	-	3	3	4	4	Yes	Yes
J. Wickham	here	5	3	-	4	4	3	4	Yes	Yes
J. Wickham	here	3	2	-	3	2	3	3	No	Yes
A. Extremeno	here	2	3	3	4	3	4	3	No	Yes
A. Extremeno	here	1	2	3	4	2	4	3	No	Yes
A. Extremeno	here	2	3	3	4	3	4	3	No	Yes
A. Extremeno	here	2	1	3	2	1	1	1	No	Yes
A. Extremeno	here	2	3	3	4	3	4	3	No	Yes
C. Wurzer	here	5	5	5	4	4	5	5	Yes	Yes
C. Wurzer	here	5	5	5	5	5	5	5	Yes	Yes
C. Wurzer	here	5	5	5	5	5	5	5	Yes	Yes
C. Wurzer	here	5	4	5	4	5	5	4	No	Yes
C. Wurzer	here	5	-	5	-	5	5	-	No	Yes
L. Csambalik	here	5	5	4	4	5	5	5	Yes	Yes
L. Csambalik	here	5	4	4	5	4	4	4	Yes	Yes
L. Csambalik	here	4	5	4	5	4	5	4	Yes	Yes

Reviewer	URL	1. In which degree is this metadata record completed?	2. Overall accuracy of the metadata provided	3. Values provided consistent to metadata standard	4. Describe the resource in an objective way?	5. Values provided, appropriate for the use in the Portal?	6. Degree of correctness of the language used	7. Overall score for the metadata of this resource	Publish ?	Comments ?
L. Csambalik	here	4	4	4	4	5	4	4	Yes	Yes
L. Csambalik	here	4	3	4	4	3	4	4	Yes	Yes
G. Barbagiannis	here	4	-	4	-	3	5	3	No	Yes
G. Barbagiannis	here	4	5	5	5	3	5	4	Yes	Yes
G. Barbagiannis	here	4	5	3	5	3	5	4	Yes	Yes
G. Barbagiannis	here	3	5	4	5	3	5	3	No	Yes
G. Barbagiannis	here	4	5	3	5	4	5	4	Yes	Yes
V. Protonotarios	here	4	4	4	5	3	5	4	No	Yes
V. Protonotarios	here	4	4	4	5	3	5	3	No	Yes
V. Protonotarios	here	4	5	5	5	4	5	4	Yes	Yes
V. Protonotarios	here	4	5	5	5	4	5	4	Yes	Yes
V. Protonotarios	here	3	5	5	5	4	5	3	No	Yes
D. Mikohazi	here	5	4	5	5	5	4	5	Yes	Yes
D. Mikohazi	here	5	4	4	5	5	5	5	Yes	Yes
D. Mikohazi	here	4	4	5	5	4	4	4	Yes	Yes
D. Mikohazi	here	4	5	5	5	4	5	5	Yes	Yes
D. Mikohazi	here	5	4	5	5	5	5	5	Yes	Yes
Maria Toader	here	4	5	5	5	4	5	5	Yes	No
Maria Toader	here	4	5	5	5	5	5	5	Yes	No
Maria Toader	N/A	4	5	4	5	5	5	5	Yes	No
Maria Toader	here	4	4	4	4	4	5	4	Yes	No
Maria Toader	here	4	4	5	5	4	5	5	Yes	No
A. Ionescu	here	4	5	5	4	4	5	4	Yes	No
A. Ionescu	here	4	5	4	5	4	5	5	Yes	No

Reviewer	URL	1. In which degree is this metadata record completed?	2. Overall accuracy of the metadata provided	3. Values provided consistent to metadata standard	4. Describe the resource in an objective way?	5. Values provided, appropriate for the use in the Portal?	6. Degree of correctness of the language used	7. Overall score for the metadata of this resource	Publish ?	Comments ?
A. Ionescu	here	2	5	2	5	4	5	4	Yes	No
A. Ionescu	here	5	5	5	5	4	5	5	Yes	No
A. Ionescu	here	5	5	5	5	4	5	5	Yes	No
U. Moor	here	4	5	4	4	3	4	4	Yes	No
U. Moor	here	5	5	4	4	4	5	4	Yes	No
U. Moor	here	4	3	3	3	2	5	3	No	Yes
U. Moor	here	4	3	3	3	2	5	3	No	Yes
U. Moor	here	4	3	3	3	2	5	3	No	Yes
A. Katrakilis	N/A	4	4	5	5	5	4	4	Yes	Yes
A. Katrakilis	N/A	5	5	4	4	5	3	3	No	Yes
A. Katrakilis	N/A	5	5	5	5	5	4	5	Yes	Yes
A. Katrakilis	N/A	5	5	5	5	5	5	5	Yes	Yes
A. Katrakilis	N/A	5	5	5	5	5	5	5	Yes	No
A. Katrakilis	N/A	5	4	5	5	5	4	4	Yes	Yes
A. Katrakilis	N/A	4	4	3	5	4	5	3	Yes	Yes
A. Katrakilis	N/A	4	4	3	5	3	5	4	Yes	Yes
A. Katrakilis	N/A	5	5	5	5	5	5	5	Yes	No
A. Katrakilis	N/A	4	4	4	5	4	4	4	Yes	Yes
A. Katrakilis	N/A	4	4	4	5	4	4	4	Yes	Yes
A. Katrakilis	N/A	4	3	3	4	3	4	3	No	Yes
A. Katrakilis	N/A	4	4	4	5	5	5	4	Yes	No
A. Katrakilis	N/A	4	4	4	5	4	5	4	Yes	No
A. Katrakilis	N/A	4	4	4	5	5	5	5	Yes	Yes
A. Katrakilis	N/A	4	4	4	5	4	5	4	Yes	Yes

Reviewer	URL	1. In which degree is this metadata record completed?	2. Overall accuracy of the metadata provided	3. Values provided consistent to metadata standard	4. Describe the resource in an objective way?	5. Values provided, appropriate for the use in the Portal?	6. Degree of correctness of the language used	7. Overall score for the metadata of this resource	Publish?	Comments?
A. Katrakilis	N/A	4	4	3	5	3	3	3	No	Yes
A. Katrakilis	N/A	4	4	3	5	4	3	4	Yes	Yes
A. Katrakilis	N/A	4	4	4	5	4	4	4	Yes	Yes
A. Katrakilis	N/A	4	4	4	4	5	5	4	Yes	Yes

Appendix E: Data from Cultural Federations Experiment

Metadata Understanding Session Data

	Easyness							Usefulness							Obligation			Obligation		
Dublin Core	***1***	***2***	***3***	***4***	***5***	***AVG***	***Blanks***	***1***	***2***	***3***	***4***	***5***	***AVG***	***Blanks***	*Mand*	*Rec*	*Opt*	*Mand*	*Rec*	*Opt*
dc.title			1		10	4,82	0	1				10	4,64	0	9	1	1	81,82%	9,09%	9,09%
dc.type		1	2	1	7	4,27	0			2		9	4,64	0	9	0	1	90,00%	0,00%	10,00%
dc.identifier	1	3	1	1	5	3,55	0		1	2		7	4,3	1	7	0	2	77,78%	0,00%	22,22%
dc.description					11	5	0				1	10	4,91	0	8	1	1	80,00%	10,00%	10,00%
dcterms.created	1	2		1	7	4	0	1		2	1	6	4,1	1	4	3	1	50,00%	37,50%	12,50%
dcterms.issued	3	1	1	1	5	3,36	0	2		4		4	3,4	1	2	1	5	25,00%	12,50%	62,50%
dc.language					11	5	0				1	10	4,91	0	10	0	0	100,00%	0,00%	0,00%
dc.source			2	2	7	4,45	0			2	1	8	4,55	0	9	1	0	90,00%	10,00%	0,00%
dc.rights			1	1	9	4,73	0			1	1	9	4,73	0	10	0	0	100,00%	0,00%	0,00%
dc.subject	1	1	4	1	4	3,55	0			1	1	8	4,7	1	5	2	1	62,50%	25,00%	12,50%
dc.relation	3	2	2	1	3	2,91	0	1	1	4		3	3,33	2	2	0	5	28,57%	0,00%	71,43%
dc.contributor	3	1	4		2	2,7	1	2	1	3	1	1	2,75	3	0	4	2	0,00%	66,67%	33,33%
dc.format			2	2	7	4,45	0		3	1	1	5	3,8	1	4	1	4	44,44%	11,11%	44,44%
dc.creator	1		1	1	8	4,36	0		1		2	8	4,55	0	7	3	0	70,00%	30,00%	0,00%
dc.publisher	1	1	2	2	5	3,82	0		1	1	3	5	4,2	1	5	2	2	55,56%	22,22%	22,22%
dc.provider	2	1	1	2	5	3,64	0	1	1	3	2	4	3,64	0	4	4	1	44,44%	44,44%	11,11%
dc.coverage	5	2	1		2	2,2	1	1	2	2		3	3,25	3	1	1	5	14,29%	14,29%	71,43%
dcterms.provenance	4	1	4		2	2,55	0	1	2	3	2	2	3,2	1	1	3	4	12,50%	37,50%	50,00%
Europeana																				
object			4		6	4,2	1					9	5	2	7	1	0	87,50%	12,50%	0,00%
URI	1		1		9	4,45	0			1		9	4,8	1	9	0	0	100,00%	0,00%	0,00%

country					11	5	0				1	10	4,91	0	9	1	0	90,00%	10,00%	0,00%
language	1				9	4,6	1		1		1	9	4,64	0	8	1	1	80,00%	10,00%	10,00%
isShownBy	3	2	1		4	3	1	1	1	1	1	4	3,75	3	4	1	2	57,14%	14,29%	28,57%
isShownAt	3	2			5	3,2	1	1	1	1	2	3	3,63	3	3	3	1	42,86%	42,86%	14,29%
provider	2	1	2	1	5	3,55	0	1	1	2	1	4	3,67	2	4	2	2	50,00%	25,00%	25,00%
data.provider	2		1	2	5	3,8	1	1	1	1	2	3	3,63	3	4	2	1	57,14%	28,57%	14,29%
CollectionName				2	9	4,82	0		1		1	7	4,56	2	6	3	1	60,00%	30,00%	10,00%
Rights	1				10	4,64	0		1			8	4,67	2	9	0	1	90,00%	0,00%	10,00%
IEEE LOM																				
meta-metadata	1	2	3		3	3,22	2	1	2	2		3	3,25	3	1	2	3	16,67%	33,33%	50,00%

Metadata Record Peer Review Data (March 2012)

Reviewer	Resource	No	1. In which degree is this metadata record completed?	2. Overall accuracy of the metadata provided	3. Values provided consistent to metadata standard	4. Describe the resource in an objective way?	5. Values provided, appropriate for the use in the Portal?	6. Degree of correctness of the language used	7. Overall score for the metadata of this resource	AVERAGE	Publish?
MNHNL	TNHM	1	3	1	5	2	1	5	3	2,86	No
MNHNL	TNHM	2	4	1	5	2	1	5	3	3,00	No
MNHNL	TNHM	3	4	1	5	2	1	5	3	3,00	No
MNHNL	JME	4	4	5	5	5	5	5	5	4,86	Yes
MNHNL	AC	5	4	1	5	2	1	5	3	3,00	No
MNHNL	AC	6	4	1	5	2	1	5	3	3,00	No
MNHNL	AC	7	4	1	5	2	1	5	3	3,00	No
MNHNL	JME	8	4	5	5	5	5	5	5	4,86	Yes
MNHNL	JME	9	4	5	5	5	4	5	4	4,57	Yes

Reviewer	Resource	No	1. In which degree is this metadata record completed?	2. Overall accuracy of the metadata provided	3. Values provided consistent to metadata standard	4. Describe the resource in an objective way?	5. Values provided, appropriate for the use in the Portal?	6. Degree of correctness of the language used	7. Overall score for the metadata of this resource	AVERAGE	Publish?
MNHNL	JME	10	4	5	5	5	4	5	4	4,57	Yes
MNHNL	NHMC	11	1	1	2	1	1	3	1	1,43	No
MNHNL	NHMC	12	1	4	3	4	4	3	3	3,14	No
MNHNL	NHMC	13	1	1	1	1	1	3	1	1,29	No
MNHNL	NHMC	14	1	4	3	5	5	5	3	3,71	No
MNHNL	NHMC	15	1	4	3	5	4	4	3	3,43	No
MNHNL	NHMC	16	1	3	3	3	3	3	3	2,71	No
MNHNL	NHMC	17	1	4	4	5	3	4	3	3,43	No
MNHNL	NHMC	18	1	3	3	3	2	3	2	2,43	No
MNHNL	TNHM	19	1	4	4	5	4	4	4	3,71	No
MNHNL	TNHM	20	5	5	5	5	4	5	4	4,71	Yes
NHMC	MNHNL	1	3	4	4	5	4	5	4	4,14	Yes
NHMC	MNHNL	2	3	4	4	5	4	5	4	4,14	Yes
NHMC	MNHNL	3	3	4	4	5	4	5	4	4,14	Yes
NHMC	MNHNL	4	3	4	4	5	4	5	4	4,14	Yes
NHMC	MNHNL	5	3	4	4	5	4	5	4	4,14	Yes
NHMC	HNHM	6	5	4	4	5	5	3	4	4,29	Yes
NHMC	HNHM	7	4	4	4	5	5	4	4	4,29	Yes
NHMC	HNHM	8	5	5	4	5	5	4	5	4,71	Yes
NHMC	TNHM	9	4	4	4	5	4	5	4	4,29	Yes
NHMC	TNHM	10	4	4	4	5	4	5	4	4,29	Yes
NHMC	TNHM	11	4	3	4	5	4	4	4	4,00	Yes
NHMC	JME	12	3	3	3	4	3	3	3	3,14	No
NHMC	JME	13	3	4	5	5	4	5	4	4,29	Yes

Reviewer	Resource	No	1. In which degree is this metadata record completed?	2. Overall accuracy of the metadata provided	3. Values provided consistent to metadata standard	4. Describe the resource in an objective way?	5. Values provided, appropriate for the use in the Portal?	6. Degree of correctness of the language used	7. Overall score for the metadata of this resource	AVERAGE	Publish?
NHMC	JME	14	4	4	5	5	4	3	4	4,14	Yes
NHMC	JME	15	4	4	5	5	3	4	4	4,14	Yes
NHMC	AC	16	4	4	4	5	2	4	3	3,71	No
NHMC	HNHM	17	5	4	4	5	5	5	5	4,71	Yes
NHMC	HNHM	18	5	4	5	5	5	5	5	4,86	Yes
NHMC	TNHM	19	4	4	5	5	5	4	5	4,57	Yes
NHMC	TNHM	20	4	5	5	5	5	4	5	4,71	Yes
NHMC	AC	21	4	4	5	5	4	5	5	4,57	Yes
NHMC	AC	22	4	4	5	5	4	5	5	4,57	Yes
NHMC	AC	23	4	4	5	5	4	5	5	4,57	Yes
NHMC	AC	24	4	4	5	5	4	5	5	4,57	Yes
TNHM	MNHNL	1	4	1	3	1	1	4	3	2,43	No
TNHM	MNHNL	2	4	5	4	5	5	5	5	4,71	Yes
TNHM	MNHNL	3	4	5	4	5	5	5	5	4,71	Yes
TNHM	MNHNL	4	4	4	5	4	5	5	4	4,43	No
TNHM	HNHM	5	4	4	4	4	4	5	4	4,14	No
TNHM	HNHM	6	4	4	5	5	5	5	5	4,71	Yes
TNHM	NHMC	7	5	5	5	5	5	5	5	5,00	Yes
TNHM	NHMC	8	4	4	4	5	5	5	4	4,43	Yes
TNHM	NHMC	9	4	4	4	5	5	5	4	4,43	Yes
GRNET	HNHM	1	2	4	3	4	3	4	3	3,29	No
GRNET	HNHM	2	2	3	3	4	3	3	3	3,00	No
GRNET	HNHM	3	2	5	4	5	2	5	4	3,86	No
GRNET	MNHNL	4	2	4	4	5	2	5	4	3,71	No

Reviewer	Resource	No	1. In which degree is this metadata record completed?	2. Overall accuracy of the metadata provided	3. Values provided consistent to metadata standard	4. Describe the resource in an objective way?	5. Values provided, appropriate for the use in the Portal?	6. Degree of correctness of the language used	7. Overall score for the metadata of this resource	AVERAGE	Publish?
GRNET	MNHNL	5	2	4	4	5	2	5	4	3,71	No
GRNET	MNHNL	6	3	3	2	5	1	4	3	3,00	No
GRNET	MNHNL	7	3	4	4	5	1	3	3	3,29	No
GRNET	MNHNL	8	3	3	4	5	4	3	3	3,57	No
JME	MNHNL	1	3	-	3	5	2	5	3	3,50	Yes
JME	MNHNL	2	3	-	3	5	3	-	3	3,40	Yes
JME	HNHM	3	-	-	5	5	5	5	5	5,00	Yes
JME	HNHM	4	3	-	4	5	5	5	5	4,50	Yes
JME	HNHM	5	3	-	4	5	5	4	4	4,17	Yes
JME	HNHM	6	-	-	3	4	5	5	3	4,00	No
JME	HNHM	7	4	-	4	5	5	5	5	4,67	Yes
JME	NHMC	8	3	-	5	5	5	5	5	4,67	Yes
JME	NHMC	9	5	-	5	5	5	5	5	5,00	Yes
JME	AC	10	3	-	4	5	5	5	3	4,17	No
HNHM	TNHM	1	5	5	5	5	5	4	5	4,86	Yes
HNHM	TNHM	2	5	5	5	5	4	5	5	4,86	Yes
HNHM	TNHM	3	1	5	5	5	5	5	4	4,29	No
HNHM	NHMC	4	1	5	5	5	2	5	4	3,86	No
HNHM	NHMC	5	5	4	5	5	5	5	5	4,86	Yes
HNHM	NHMC	6	4	5	5	5	5	5	5	4,86	Yes
HNHM	NHMC	7	4	5	5	5	5	5	5	4,86	Yes
HNHM	NHMC	8	1	3	5	5	2	5	3	3,43	No
HNHM	NHMC	9	5	4	5	5	5	5	5	4,86	Yes
AC	NHMC	1	4	4	4	4	4	4	4	4,00	Yes

Reviewer	Resource	No	1. In which degree is this metadata record completed?	2. Overall accuracy of the metadata provided	3. Values provided consistent to metadata standard	4. Describe the resource in an objective way?	5. Values provided, appropriate for the use in the Portal?	6. Degree of correctness of the language used	7. Overall score for the metadata of this resource	AVERAGE	Publish?
AC	MNHNL	2	2	3	3	4	4	2	3	3,00	Yes
AC	MNHNL	3	3	4	3	4	4	4	4	3,71	Yes
AC	HNHM	4	4	3	4	1	2	4	3	3,00	No
AC	HNHM	5	3	2	3	3	3	3	3	2,86	No
AC	NHMC	6	4	4	4	4	4	4	4	4,00	Yes
AC	NHMC	7	4	4	4	4	4	4	4	4,00	Yes
AC	NHMC	8	4	4	4	4	4	4	4	4,00	Yes
AC	NHMC	9	-	-	2	3	2	1	2	2,00	No
AC	MNHNL	10	4	4	4	5	5	5	4	4,43	Yes
HNHM	AC	10	5	5	4	5	3	5	4	4,43	No
HNHM	AC	11	5	5	4	5	3	5	4	4,43	No
HNHM	AC	12	5	5	4	5	3	5	4	4,43	No
HNHM	MNHNL	13	4	3	3	5	2	5	3	3,57	No
HNHM	MNHNL	14	4	3	3	5	2	5	3	3,57	No
HNHM	MNHNL	15	4	3	3	5	2	5	3	3,57	No
HNHM	MNHNL	16	5	5	4	5	3	4	4	4,29	Yes
HNHM	MNHNL	17	5	5	4	5	3	3	4	4,14	Yes
HNHM	NHMC	18	1	4	3	5	2	5	2	3,14	No

Metadata Record Peer Review Data (May 2013)

Reviewer	Reviewed resources of Partner	1. In which degree is this metadata record completed?	2. Overall accuracy of the metadata provided	3. Values provided consistent to metadata standard	4. Describe the resource in an objective way?	5. Values provided, appropriate for the use in the Portal?	6. Degree of correctness of the language used	7. Overall score for the metadata of this resource	Publish?
AC	MNHNL	5	4	5	5	5	5	5	Yes
AC	MNHNL	5	4	5	5	5	5	5	Yes
TNHM	MNHNL	5	1	5	5	4	5	5	No
TNHM	MNHNL	5	5	5	5	5	5	5	Yes
JME	MNHNL	5	5	5	5	5	5	5	Yes
JME	MNHNL	5	5	5	5	5	5	5	Yes
NHMC	MNHNL	5	5	5	5	5	5	5	Yes
NHMC	MNHNL	5	5	5	5	5	5	5	Yes
NHMC	MNHNL	5	5	5	5	5	5	5	Yes
NHMC	MNHNL	5	5	4	5	5	5	5	Yes
NHMC	MNHNL	5	4	5	5	5	5	5	Yes
NHMC	MNHNL	5	5	5	5	5	5	5	Yes
NHMC	MNHNL	5	5	5	5	5	5	5	Yes
NHMC	MNHNL	5	5	5	5	5	5	5	Yes
HNHM	MNHNL	3	3	3	5	3	4	3	No
HNHM	MNHNL	3	5	3	5	4	4	4	Yes
HNHM	MNHNL	3	5	4	5	4	4	4	Yes
HNHM	MNHNL	5	4	5	5	5	5	5	Yes
HNHM	MNHNL	5	4	5	5	4	4	3	Yes
HNHM	MNHNL	5	4	5	5	5	5	5	Yes
AC	HNHM	5	3	4	5	5	4	4	Yes
AC	HNHM	4	4	3	5	4	3	3,5	Yes
TNHM	HNHM	3	5	5	5	4	5	4	Yes
TNHM	HNHM	3	5	5	5	4	5	4	Yes

Reviewer	Reviewed resources of Partner	1. In which degree is this metadata record completed?	2. Overall accuracy of the metadata provided	3. Values provided consistent to metadata standard	4. Describe the resource in an objective way?	5. Values provided, appropriate for the use in the Portal?	6. Degree of correctness of the language used	7. Overall score for the metadata of this resource	Publish?
JME	HNHM	5	5	5	5	5	5	5	Yes
JME	HNHM	4	5	5	5	5	5	5	Yes
MNHNL	HNHM	5	5	5	5	5	5	5	Yes
MNHNL	HNHM	5	5	5	5	5	5	5	Yes
MNHNL	HNHM	5	5	5	5	5	5	5	Yes
MNHNL	HNHM	2	4	4	4	4	3	3	Yes
MNHNL	HNHM	2	2	2	2	2	3	2	No
NHMC	HNHM	4	5	5	5	5	5	5	Yes
NHMC	HNHM	4	5	5	5	5	5	5	Yes
NHMC	HNHM	4	5	5	5	5	5	5	Yes
NHMC	HNHM	5	5	5	5	5	5	5	Yes
NHMC	HNHM	5	5	5	5	5	5	5	Yes
NHMC	HNHM	5	5	5	5	5	5	5	Yes
NHMC	HNHM	4	5	5	5	5	5	5	Yes
NHMC	HNHM	4	5	5	5	5	5	5	Yes
NHMC	HNHM	4	5	5	5	5	5	5	Yes
AC	NHMC	5	5	5	5	5	3	5	Yes
AC	NHMC	5	5	5	5	5	5	5	Yes
AC	NHMC	5	5	5	5	5	5	5	Yes
TNHM	NHMC	3	5	5	5	4	5	4	Yes
TNHM	NHMC	3	5	5	5	4	5	4	Yes
TNHM	NHMC	3	5	5	5	4	5	4	Yes
JME	NHMC	4	5	5	5	5	5	5	Yes
JME	NHMC	5	5	5	5	5	5	5	Yes
JME	NHMC	5	5	5	5	5	4	5	Yes
HNHM	NHMC	4	4	5	5	5	5	4	Yes

Reviewer	Reviewed resources of Partner	1. In which degree is this metadata record completed?	2. Overall accuracy of the metadata provided	3. Values provided consistent to metadata standard	4. Describe the resource in an objective way?	5. Values provided, appropriate for the use in the Portal?	6. Degree of correctness of the language used	7. Overall score for the metadata of this resource	Publish?
HNHM	NHMC	4	4	5	5	5	5	4	Yes
HNHM	NHMC	4	4	4	5	5	5	4	Yes
HNHM	NHMC	5	5	5	5	5	5	5	Yes
HNHM	NHMC	5	5	5	5	5	5	5	Yes
MNHNL	NHMC	5	5	5	5	5	5	5	Yes
MNHNL	NHMC	5	5	5	5	5	5	5	Yes
MNHNL	NHMC	5	5	5	5	5	5	5	Yes
MNHNL	NHMC	4	4	4	4	4	4	4	Yes
MNHNL	NHMC	4	4	4	4	4	4	4	Yes
MNHNL	NHMC	4	4	4	4	4	4	4	Yes
AC	TNHM	2	3	3	5	3	3	3	Yes
JME	TNHM	3	3	5	5	4	5	4	No
JME	TNHM	3	3	5	5	4	5	4	No
NHMC	TNHM	3	4	5	5	5	5	4	No
NHMC	TNHM	2	5	5	5	2	5	3	No
NHMC	TNHM	4	4	5	5	5	5	5	No
HNHM	TNHM	2	4	2	5	4	5	3	No
HNHM	TNHM	4	5	5	5	5	5	2	Yes
MNHNL	TNHM	5	4	5	5	5	5	4	Yes
MNHNL	TNHM	4	4	3	4	4	4	4	Yes
TNHM	AC	2	5	3	5	2	5	3	Yes
TNHM	AC	1	1	1	1	1	1	1	No
JME	AC	2	3	5	5	4	5	4	No
NHMC	AC	4	5	5	5	5	5	5	Yes
NHMC	AC	4	5	5	5	3	5	4	Yes
NHMC	AC	4	4	5	5	5	5	5	No

Reviewer	Reviewed resources of Partner	1. In which degree is this metadata record completed?	2. Overall accuracy of the metadata provided	3. Values provided consistent to metadata standard	4. Describe the resource in an objective way?	5. Values provided, appropriate for the use in the Portal?	6. Degree of correctness of the language used	7. Overall score for the metadata of this resource	Publish?
HNHM	AC	2	5	5	5	5	5	3	No
HNHM	AC	1	1	1	5	1	1	1	No
MNHNL	AC	2	1	1	1	1	1	1	No
MNHNL	AC	1	2	1	1	1	3	1	No
AC	JME	3	4	4	5	4	4	4	Yes
AC	JME	2	2	2	5	2	3	2	No
TNHM	JME	3	5	5	5	2	5	3,5	Yes
NHMC	JME	4	5	5	5	5	5	5	No
NHMC	JME	3	5	5	5	4	5	4	Yes
NHMC	JME	3	4	5	5	5	5	4	No
HNHM	JME	2	5	5	5	5	4	4	Yes
HNHM	JME	4	5	5	5	5	5	5	Yes
MNHNL	JME	3	5	5	5	4	5	4	Yes
MNHNL	JME	2	3	2	3	3	3	3	Yes

Metadata Record Peer Review Data (May 2013)

Reviewer	From	Resource	1. In which degree is this metadata record completed?	2. Overall accuracy of the metadata provided	3. Values provided consistent to metadata standard	4. Describe the resource in an objective way?	5. Values provided, appropriate for the use in the Portal?	6. Degree of correctness of the language used	7. Overall score for the metadata of this resource	Publish?
NHMC	HNHM	here	5	5	5	5	5	5	5	Yes
NHMC	JME	here	5	5	5	5	5	5	5	Yes
NHMC	MNHNL	here	5	5	4	5	5	5	5	Yes
NHMC	MNHNL	here	5	5	5	5	5	5	5	Yes

Reviewer	From	Resource	1. In which degree is this metadata record completed?	2. Overall accuracy of the metadata provided	3. Values provided consistent to metadata standard	4. Describe the resource in an objective way?	5. Values provided, appropriate for the use in the Portal?	6. Degree of correctness of the language used	7. Overall score for the metadata of this resource	Publish?
NHMC	TNHM	here	5	5	5	5	5	5	5	Yes
NHMC	TNHM	here	5	5	5	5	5	5	5	Yes
HNHM	MNHNL	here	4	5	5	4	4	5	4	Yes
HNHM	NHMC	here	4	5	4	5	4	5	5	Yes
HNHM	NHMC	here	5	5	5	5	4	5	5	Yes
HNHM	AC	here	2	3	2	5	1	5	2	Yes
HNHM	TNHM	here	5	5	4	5	4	5	4	Yes
HNHM	TNHM	here	4	4	4	5	3	5	4	Yes
MNHNL	NHMC	N/A	3	2	2	2	2	3	2	No
MNHNL	HNHM	N/A	5	3	5	3	4	4	4	Yes
MNHNL	TNHM	N/A	3	2	2	2	2	3	2	No
MNHNL	JME	N/A	3	2	3	2	3	3	3	No
MNHNL	NHMC	N/A	4	4	3	4	3	4	4	Yes
JME	AC	here	4	4	5	5	5	5	5	Yes
JME	HNHM	here	4	4	5	5	5	5	5	Yes
JME	MNHNL	here	4	4	5	5	5	5	5	Yes
JME	MNHNL	here	4	4	5	5	5	5	5	Yes
JME	NHMC	here	5	5	5	5	5	5	5	Yes
JME	NHMC	here	5	4	5	5	5	5	5	Yes
AC	HNHM	N/A	2	2	3	5	3	4	3	Yes
AC	JME	N/A	3	3	3	3	3	3	3	Yes
AC	MNHNL	N/A	4	5	5	5	5	5	5	Yes
AC	MNHNL	N/A	4	5	5	5	5	5	5	Yes
AC	NHMC	N/A	5	5	5	5	5	5	5	Yes
AC	NHMC	N/A	5	5	5	5	5	5	5	Yes
TNHM	AC	N/A	3	5	5	5	4	5	4	Yes

Reviewer	From	Resource	1. In which degree is this metadata record completed?	2. Overall accuracy of the metadata provided	3. Values provided consistent to metadata standard	4. Describe the resource in an objective way?	5. Values provided, appropriate for the use in the Portal?	6. Degree of correctness of the language used	7. Overall score for the metadata of this resource	Publish?
TNHM	HNHM	N/A	3	5	5	5	4	5	4	Yes
TNHM	JME	N/A	3	5	5	5	4	5	4	Yes
TNHM	MNHNL	N/A	3	5	5	5	4	5	4	Yes
TNHM	NHMC	N/A	3	5	5	5	4	5	4	Yes

Appendix F: Data from Research Federations Experiment

Metadata Understanding Session Data (June 2011)

Dublin Core	Easyness							Usefulness							Obligation			Obligation		
	1	*2*	*3*	*4*	*5*	*AVG*	*Blanks*	*1*	*2*	*3*	*4*	*5*	*AVG*	*Blanks*	*Mand*	*Rec*	*Opt*	*Mand*	*Rec*	*Opt*
dcterms.title		1			14	**4,8**	**1**					15	**5**	**1**	16			100,00%	0,00%	0,00%
dcterms.alternative		1	2	4	8	**4,27**	**1**	1		3	5	6	**4**	**1**	3	3	10	18,75%	18,75%	62,50%
dcterms.creator	1	1	4		9	**4**	**1**			1		11	**4,83**	**4**	11	1		91,67%	8,33%	0,00%
ags.creatorPersonal			4	2	8	**4,29**	**2**	1	1	4		8	**3,93**	**2**	6	3	5	42,86%	21,43%	35,71%
ags.creatorCorporate			1	5	8	**4,5**	**2**	2		3	1	8	**3,93**	**2**	5	4	5	35,71%	28,57%	35,71%
dcterms.contributor	1	3		2	8	**3,93**	**2**	1		4	3	6	**3,93**	**2**	5	4	7	31,25%	25,00%	43,75%
dcterms.publisher				1	13	**4,93**	**2**	1				14	**4,73**	**1**	10	3	1	71,43%	21,43%	7,14%
dcterms.date	1		2		12	**4,47**	**1**				1	14	**4,93**	**1**	14	1		93,33%	6,67%	0,00%
dcterms.identifier		1	3	4	7	**4,13**	**1**			1	1	13	**4,8**	**1**	7	8		46,67%	53,33%	0,00%
dcterms.language				1	14	**4,93**	**1**					15	**5**	**1**	11	4	1	68,75%	25,00%	6,25%
dcterms.format		1	2	3	9	**4,33**	**1**			1	4	9	**4,57**	**2**	7	5	1	53,85%	38,46%	7,69%
dcterms.source	2	2	5	1	5	**3,33**	**1**		2	5	1	5	**3,69**	**3**	7	5	2	50,00%	35,71%	14,29%
dcterms.type				7	8	**4,53**	**1**			1	1	13	**4,8**	**1**	7	7		50,00%	50,00%	0,00%
meta-metadata.catalog	2	4	6	3		**2,67**	**1**	1	1	3	6	3	**3,64**	**2**	5	3	6	35,71%	21,43%	42,86%
meta-metadata.entry	2	5	3	4		**2,64**	**2**		1	3	6	4	**3,93**	**2**	4	3	6	30,77%	23,08%	46,15%
mm.contribute.role		4	5	2	1	**3**	**4**	1	2	3	4	3	**3,46**	**3**	4	3	6	30,77%	23,08%	46,15%
mm.contribute.entity	2	2	6	2	2	**3**	**2**		1	3	4	4	**3,92**	**4**	4	3	5	33,33%	25,00%	41,67%
mm.contribute.date	2	3	4	2	3	**3,07**	**2**		1	4	4	3	**3,75**	**4**	3	4	5	25,00%	33,33%	41,67%
mm.metadata schema	1	4	4	2	3	**3,14**	**2**		2	3	3	6	**3,93**	**2**	5	3	5	38,46%	23,08%	38,46%
mm.language		4	1	2	7	**3,86**	**2**		1	2	2	7	**4,25**	**4**	4	2	5	36,36%	18,18%	45,45%
dcterms.rights		3	3	2	7	**3,87**	**1**			2	4	6	**4,33**	**4**	2	5	2	22,22%	55,56%	22,22%

Dublin Core	Easyness							Usefulness							Obligation			Obligation		
	1	*2*	*3*	*4*	*5*	*AVG*	*Blanks*	*1*	*2*	*3*	*4*	*5*	*AVG*	*Blanks*	*Mand*	*Rec*	*Opt*	*Mand*	*Rec*	*Opt*
dcterms.accessrights			2	1	12	**4,67**	**1**			1	2	12	**4,73**	**1**	6	4	2	50,00%	33,33%	16,67%
dcterms.license		1	1	2	11	**4,53**	**1**		1	1	5	7	**4,29**	**2**	4	7	2	30,77%	53,85%	15,38%
ags.rightsStatement		2	2	3	8	**4,13**	**1**		1	2	4	5	**4,08**	**4**	3	3	6	25,00%	25,00%	50,00%
ags.termsOfUse	1	1	3	1	9	**4,07**	**1**		1	3	4	4	**3,92**	**4**	3	2	6	27,27%	18,18%	54,55%
dcterms.relation		5	3	1	7	**3,63**	**0**		2	3	3	6	**3,93**	**2**	3	3	8	21,43%	21,43%	57,14%
dcterms.conformsTo		2	2	6	6	**4**	**0**	1	2	4	3	3	**3,38**	**3**	1	4	7	8,33%	33,33%	58,33%
dcterms.references		2	2	1	9	**4,21**	**2**		1	2	5	5	**4,08**	**3**	2	4	7	15,38%	30,77%	53,85%
dcterms.isReferencedBy		2	1	1	10	**4,36**	**2**		2	6	1	5	**3,64**	**2**	3	1	9	23,08%	7,69%	69,23%
ags.isTranslationOf		1	2	3	8	**4,29**	**2**			9	4	1	**3,43**	**2**	1	2	10	7,69%	15,38%	76,92%
ags.hasTranslation		1	2	4	8	**4,27**	**1**			8	4	2	**3,57**	**2**	1	2	10	7,69%	15,38%	76,92%
dcterms.subject		1	1	1	13	**4,63**	**0**					14	**5**	**2**	7	6	2	46,67%	40,00%	13,33%
dcterms.description	1		2	1	11	**4,4**	**1**				5	9	**4,64**	**2**	4	6	4	28,57%	42,86%	28,57%
dcterms.abstract				2	13	**4,87**	**1**		1	1	2	9	**4,46**	**3**	4	7	3	28,57%	50,00%	21,43%
dcterms.blbiographicCitation		1		1	12	**4,71**	**2**	2		1	1	11	**4,27**	**1**	8	4	2	57,14%	28,57%	14,29%

Metadata Understanding Session Data (May 2012)

Dublin Core	Easyness							Usefulness							Obligation			Obligation		
	1	*2*	*3*	*4*	*5*	*AVG*	*Blanks*	*1*	*2*	*3*	*4*	*5*	*AVG*	*Blanks*	*Mand*	*Rec*	*Opt*	*Mand*	*Rec*	*Opt*
1.1 Title				1	12	**4,92**	**1**			1		12	**4,85**	**1**	14			100,00%	0,00%	0,00%
1.2 Alternative Title	1	4	3	1	3	**3,08**	**2**	2		3	4	3	**3,5**	**2**		2	12	0,00%	14,29%	85,71%
2.1 Creator			1	3	8	**4,58**	**2**				2	10	**4,83**	**2**	7	6		53,85%	46,15%	0,00%
2.2 Contributor	2	2	3	1	4	**3,25**	**2**			3	5	4	**4,08**	**2**		5	8	0,00%	38,46%	61,54%
2.3 Publisher			2	4	6	**4,33**	**2**				4	8	**4,67**	**2**	3	7	3	23,08%	53,85%	23,08%

Dublin Core	Easyness							Usefulness							Obligation			Obligation		
	1	*2*	*3*	*4*	*5*	*AVG*	*Blanks*	*1*	*2*	*3*	*4*	*5*	*AVG*	*Blanks*	*Mand*	*Rec*	*Opt*	*Mand*	*Rec*	*Opt*
3.1 Date		2	3	2	6	**3,92**	**1**	1			2	10	**4,54**	**1**	10	3		76,92%	23,08%	0,00%
3.2 Language	1			2	10	**4,54**	**1**	1			3	9	**4,46**	**1**	9	3		75,00%	25,00%	0,00%
3.3 Identifier	2	1	3	4	3	**3,38**	**1**		1		6	6	**4,31**	**1**	4	9		30,77%	69,23%	0,00%
3.4 Format		2	2	4	4	**3,83**	**2**		2	3	4	4	**3,77**	**1**	2	8	3	15,38%	61,54%	23,08%
4.1 Is Shown By	3		4	3	2	**3,08**	**2**		1	4	5	2	**3,67**	**2**	2	6	5	15,38%	46,15%	38,46%
4.2 Is Shown At	3		3	3	3	**3,25**	**2**			4	5	3	**3,92**	**2**	2	6	5	15,38%	46,15%	38,46%
5.1 Subject			1	7	5	**4,31**	**1**	1		1	1	10	**4,46**	**1**	5	6	2	38,46%	46,15%	15,38%
6.1 Description	2	1	2	3	5	**3,62**	**1**	2	1	3	1	6	**3,62**	**1**	3	2	8	23,08%	15,38%	61,54%
6.2 Abstract			2	2	9	**4,54**	**1**			1	3	9	**4,62**	**1**	5	7	1	38,46%	53,85%	7,69%
6.3 Bibliographic Citation		1	3	4	5	**4**	**1**		1	1	3	8	**4,38**	**1**	4	4	5	30,77%	30,77%	38,46%
6.4 Type	1	1	3	3	5	**3,77**	**1**			2	4	7	**4,38**	**1**	6	6	1	46,15%	46,15%	7,69%
7.1 Rights	1	2	3	6	1	**3,31**	**1**	1	1	3	3	5	**3,77**	**1**	3	4	6	23,08%	30,77%	46,15%
7.2 Access Rights	1	1	3	6	2	**3,54**	**1**	1		2	3	7	**4,15**	**1**	5	3	5	38,46%	23,08%	38,46%
7.3 License	1	3	3	2	3	**3,25**	**2**	1	1	2	3	5	**3,83**	**2**	2	4	6	16,67%	33,33%	50,00%
8.1 Review Status	1		2	3	7	**4,15**	**1**			4	5	3	**3,92**	**2**	2	5	6	15,38%	38,46%	46,15%
8.2 Publication Status	1		1	3	8	**4,31**	**1**			4	6	2	**3,83**	**2**	2	6	5	15,38%	46,15%	38,46%
9.1 Relation	3	4	5	1		**2,31**	**1**	2	2	4	3	1	**2,92**	**2**		4	8	0,00%	33,33%	66,67%
9.2 Conforms To	2	5	5		1	**2,46**	**1**	3	2	5	2		**2,5**	**2**		4	8	0,00%	33,33%	66,67%
9.3 References	1	2	5	3	2	**3,23**	**1**		3	1	5	3	**3,67**	**2**	1	4	7	8,33%	33,33%	58,33%
9.4 Is Referenced By		1	3	5	2	**3,73**	**3**		2	4	3	2	**3,45**	**3**	1	3	8	8,33%	25,00%	66,67%
9.5 Has Part	2	5	3	2		**2,42**	**2**	3	1	4	2	1	**2,73**	**3**		3	9	0,00%	25,00%	75,00%
9.6 Is Part Of	1	3	5	2	1	**2,92**	**2**	3		4	2	1	**2,8**	**4**		4	8	0,00%	33,33%	66,67%
9.7 Has Version	2	3	4	2	1	**2,75**	**2**	2	1	4	3	1	**3**	**3**		4	8	0,00%	33,33%	66,67%
9.8 Is Version Of	2	2	5	2	1	**2,83**	**2**	2	1	4	3	1	**3**	**3**		4	8	0,00%	33,33%	66,67%

Dublin Core	Easyness							Usefulness							Obligation			Obligation		
	1	*2*	*3*	*4*	*5*	*AVG*	*Blanks*	*1*	*2*	*3*	*4*	*5*	*AVG*	*Blanks*	*Mand*	*Rec*	*Opt*	*Mand*	*Rec*	*Opt*
9.9 Has Translation		3	4	4	1	**3,25**	**2**	2		3	5	1	**3,27**	**3**	1	2	9	8,33%	16,67%	75,00%
9.10 Is Translation Of		3	4	4	1	**3,25**	**2**	2		4	4	1	**3,18**	**3**	1	2	9	8,33%	16,67%	75,00%
9.11 Has Meta-metadata	3	3	3	1	2	**2,67**	**2**	1	1	5	2	2	**3,27**	**3**	1	3	8	8,33%	25,00%	66,67%
10.1 Object of Interest		1	4	4	3	**3,75**	**2**	1		3	5	2	**3,64**	**3**	2	1	9	16,67%	8,33%	75,00%
10.2 Variable		6	5	1		**2,58**	**2**	1	2	4	2	2	**3,18**	**3**	1	2	9	8,33%	16,67%	75,00%
10.3 Method	1	1	5	3	2	**3,33**	**2**	1	1	3	3	3	**3,55**	**3**	1	2	9	8,33%	16,67%	75,00%
10.4 Protocol	1	3	4	2	2	**3,08**	**2**	1	1	3	3	3	**3,55**	**3**	1	2	9	8,33%	16,67%	75,00%
10.5 Instrument	1	3	3	4	1	**3,08**	**2**	1	1	3	4	2	**3,45**	**3**		2	9	0,00%	18,18%	81,82%
10.6 Technique	1	2	4	3	2	**3,25**	**2**	1	1	4	4	1	**3,27**	**3**		2	9	0,00%	18,18%	81,82%
1. Identifier	3	2		2	5	**3,33**	**2**	2	1	1		7	**3,82**	**3**	6	5	1	50,00%	41,67%	8,33%
2. Type	1	2	1	2	5	**3,73**	**3**	1	2	1	1	6	**3,82**	**3**	2	6	3	18,18%	54,55%	27,27%
3. Language	2	1		3	6	**3,83**	**2**	1	2		3	5	**3,82**	**3**	4	5	2	36,36%	45,45%	18,18%
4. Date		1	1	3	7	**4,33**	**2**		2	1	1	7	**4,18**	**3**	5	4	1	50,00%	40,00%	10,00%
5. Contributor	2	1	3	2	4	**3,42**	**2**	2	1	2	1	5	**3,55**	**3**	3	5	3	27,27%	45,45%	27,27%
foaf: Agent/Name	1	1	3	1	6	**3,83**	**2**			3	4	4	**4,09**	**3**	4	2	5	36,36%	18,18%	45,45%
foaf: Organization/Name	1		4		7	**4**	**2**		1	2	3	4	**4**	**4**	3	3	5	27,27%	27,27%	45,45%
foaf: Person/First Name	1	1	3		7	**3,92**	**2**		1	3	2	5	**4**	**3**	3	4	5	25,00%	33,33%	41,67%
foaf: Person/Last Name	1	2	2		7	**3,83**	**2**		1	2	3	5	**4,09**	**3**	4	3	5	33,33%	25,00%	41,67%
foaf: Person/Personal Mailbox	1	1	3		7	**3,92**	**2**	1	1	3	3	3	**3,55**	**3**	1	2	8	9,09%	18,18%	72,73%

Metadata Record Peer Review Data (August 2011)

Reviewer	Resource	1. In which degree is this metadata record completed?	2. Overall accuracy of the metadata provided	3. Values provided consistent to metadata standard	4. Describe the resource in an objective way?	5. Values provided, appropriate for the use in the Portal?	6. Degree of correctness of the language used	7. Overall score for the metadata of this resource	Publish?
CULS	INRA	2	4	-	3	3	5	4	Yes
CULS	UHasselt	2	4	-	4	5	5	4	Yes
CULS	SLU	2	4	-	4	5	5	4	Yes
CULS	ICROFS	2	4	-	-	4	5	4	No
CULS	CINECA	2	5	-	5	5	5	5	Yes
CULS	FAO	2	4	-	4	5	5	4	Yes
CULS	iFremer	2	5	-	5	5	5	5	Yes
INRA	UHasselt	1	4	5	5	2	5	4	Yes
INRA	SLU	1	1	5	5	1	5	3	Yes
INRA	ICROFS	1	5	5	5	1	3	4	Yes
INRA	CINECA	1	1	5	5	1	5	2	No
INRA	iFremer	-	-	-	-	-	-	-	-
INRA	ICROFS	1	5	5	4	2	5	4	Yes
INRA	FAO	1	1	5	5	1	5	3	No
INRA	Cemadoc	2	5	5	5	2	5	4	Yes
AUA	CINECA	1	1	1	1	1	1	1	No
AUA	FAO	3	4	4	4	4	2	3	Yes
AUA	ICROFS	4	4	5	4	4	4	4	Yes
AUA	INRA	4	5	4	5	4	4	4	Yes
AUA	SLU	4	5	4	4	4	4	4	Yes
AUA	UHasselt	4	5	4	5	4	5	5	Yes
AUA	WUR	4	5	4	4	4	5	4	Yes
CINECA	INRA	1	4	4	5	3	5	3	No
CINECA	UHasselt	1	4	4	5	3	5	3	No

Reviewer	Resource	1. In which degree is this metadata record completed?	2. Overall accuracy of the metadata provided	3. Values provided consistent to metadata standard	4. Describe the resource in an objective way?	5. Values provided, appropriate for the use in the Portal?	6. Degree of correctness of the language used	7. Overall score for the metadata of this resource	Publish?
CINECA	SLU	1	3	4	3	1	4	2	Yes
CINECA	ICROFS	1	4	4	2	1	3	2	No
CINECA	FAO	-	-	-	-	-	-	-	-
CINECA	TrAgLor	1	4	3	2	1	3	2	No
CINECA	EKT	1	4	4	3	2	3	3	Yes
SLU	WUR	2	4	4	5	2	5	2	No
SLU	EKT	2	4	4	5	2	5	2	No
SLU	FAO	1	-	5	5	1	5	1	No
SLU	INRA	1	4	4	5	2	5	2	No
SLU	Cemadoc	2	4	4	5	2	5	2	No
SLU	UHasselt	1	5	4	5	2	5	2	No
SLU	ICROFS	1	-	4	5	2	5	2	No
SLU	CINECA	1	4	4	4	2	5	1	No
UHasselt	INRA	5	5	3	5	3	4	5	Yes
UHasselt	SLU	5	5	3	5	3	4	5	Yes
UHasselt	ICROFS	2	-	-	-	-	4	2	No
UHasselt	CINECA	5	4	3	5	3	4	4	Yes
UHasselt	FAO	3	-	-	-	-	4	3	No
UHasselt	iFremer	4	4	4	4	3	5	4	Yes
UHasselt	WUR	2	3	-	-	-	4	3	-
ACTA	INRA	5	5	5	5	5	5	5	Yes
ACTA	UHasselt	3	2	5	5	2	-	3	Yes
ACTA	SLU	5	5	5	5	5	5	5	Yes
ACTA	ICROFS	5	5	5	5	5	5	5	Yes
ACTA	CINECA	1	3	5	1	1	4	2	No
ACTA	FAO	4	4	5	5	2	5	4	Yes

Reviewer	Resource	1. In which degree is this metadata record completed?	2. Overall accuracy of the metadata provided	3. Values provided consistent to metadata standard	4. Describe the resource in an objective way?	5. Values provided, appropriate for the use in the Portal?	6. Degree of correctness of the language used	7. Overall score for the metadata of this resource	Publish?
ACTA	iFremer	3	4	4	4	4	5	4	Yes
ARI	INRA	5	5	5	5	5	5	5	Yes
ARI	UHasselt	5	5	5	5	4	5	5	Yes
ARI	SLU	5	5	4	5	4	5	5	Yes
ARI	ICROFS	5	4	3	4	2	5	4	Yes
ARI	CINECA	5	5	4	5	4	4	5	Yes
ARI	FAO	5	5	3	5	3	5	4	Yes
ARI	Cemadoc	5	3	3	4	3	4	4	Yes
ICROFS	INRA	1	4	4	5	1	4	1	No
ICROFS	UHasselt	1	4	4	5	1	5	1	No
ICROFS	SLU	1	4	4	5	1	5	1	No
ICROFS	CINECA	1	3	3	5	1	5	1	No
ICROFS	FAO	1	4	4	5	1	5	1	No
ICROFS	Cemadoc	1	4	4	5	1	5	1	No
ICROFS	WUR	1	3	4	5	1	5	1	No

Metadata Record Peer Review Data (December 2011)

Reviewer	Resource	1. In which degree is this metadata record completed?	2. Overall accuracy of the metadata provided	3. Values provided consistent to metadata standard	4. Describe the resource in an objective way?	5. Values provided, appropriate for the use in the Portal?	6. Degree of correctness of the language used	7. Overall score for the metadata of this resource	Publish?
ACTA	ARI	4	5	4	5	4	5	4	Yes
ACTA	CULS	3	5	3	4	3	5	4	Yes
ACTA	ICROFS	4	5	5	4	5	5	5	No

Reviewer	Resource	1. In which degree is this metadata record completed?	2. Overall accuracy of the metadata provided	3. Values provided consistent to metadata standard	4. Describe the resource in an objective way?	5. Values provided, appropriate for the use in the Portal?	6. Degree of correctness of the language used	7. Overall score for the metadata of this resource	Publish?
ACTA	INRA	4	5	4	5	5	5	5	Yes
ACTA	SLU	4	5	4	5	5	4	5	Yes
ACTA	UHasselt	4	5	4	5	5	4	5	Yes
ARI	ACTA	1	4	4	4	2	5	3	Yes
ARI	AUA	5	5	5	5	5	5	5	Yes
ARI	CINECA	4	4	5	4	4	5	4	Yes
ARI	ICROFS	5	5	5	5	4	5	5	Yes
ARI	INRA	5	5	5	5	4	5	5	Yes
ARI	SLU	5	5	5	5	4	5	5	Yes
ARI	UHasselt	5	5	5	5	4	5	5	Yes
AUA	ACTA	2	5	3	5	2	4	3	Yes
AUA	ARI	3	5	4	5	4	5	4	Yes
AUA	CINECA	2	5	4	5	4	4	4	Yes
AUA	ICROFS	3	5	4	5	4	5	4	Yes
AUA	INRA	3	5	4	5	4	5	4	Yes
AUA	SLU	3	5	3	5	4	5	4	Yes
AUA	UHasselt	3	5	4	5	4	5	4	Yes
CINECA	ARI	4	5	4	5	5	5	5	Yes
CINECA	AUA	4	5	3	4	4	5	5	Yes
CINECA	CULS	3	4	3	5	4	5	4	Yes
CINECA	ICROFS	4	5	4	5	5	5	5	Yes
CINECA	INRA	4	5	3	5	5	5	4	Yes
CINECA	SLU	4	5	3	5	4	5	4	Yes
CINECA	UHasselt	3	5	3	5	5	5	4	Yes
CULS	ACTA	2	4	3	2	2	1	2	Yes

Reviewer	Resource	1. In which degree is this metadata record completed?	2. Overall accuracy of the metadata provided	3. Values provided consistent to metadata standard	4. Describe the resource in an objective way?	5. Values provided, appropriate for the use in the Portal?	6. Degree of correctness of the language used	7. Overall score for the metadata of this resource	Publish?
CULS	CINECA	2	5	4	5	5	5	5	Yes
CULS	ICROFS	2	5	4	5	5	5	5	Yes
CULS	UHasselt	2	4	4	5	4	5	4	Yes
ICROFS	ACTA	1	1	2	3	1	5	2	No
ICROFS	ARI	3	4	5	5	5	5	4	Yes
ICROFS	AUA	2	3	4	5	4	5	3	Yes
ICROFS	CULS	3	2	3	4	2	5	3	No
ICROFS	INRA	3	3	2	5	5	-	3	Yes
ICROFS	SLU	3	5	5	5	5	5	5	Yes
ICROFS	UHasselt	2	2	4	4	4	4	3	Yes
INRA	ACTA	2	4	5	4	1	5	3	No
INRA	ARI	2	4	4	5	5	5	4	Yes
INRA	AUA	2	5	5	5	5	5	4	Yes
INRA	CINECA	1	4	5	5	4	5	3	No
INRA	CULS	1	4	4	5	4	5	4	No
INRA	ICROFS	2	5	5	5	5	4	4	Yes
INRA	SLU	2	5	5	5	5	5	4	Yes
INRA	UHasselt	2	4	5	5	4	5	3	No
SLU	ACTA	2	3	2	2	3	5	3	No
SLU	UHasselt	2	4	3	5	2	5	3	No
SLU	INRA	2	4	4	5	4	-	4	Yes
SLU	AUA	4	3	3	3	3	5	3	No
SLU	SLU	4	4	4	5	4	5	4	Yes
SLU	ICROFS	2	5	5	5	2	5	4	Yes
SLU	CINECA	4	3	3	4	3	5	4	No

Reviewer	Resource	1. In which degree is this metadata record completed?	2. Overall accuracy of the metadata provided	3. Values provided consistent to metadata standard	4. Describe the resource in an objective way?	5. Values provided, appropriate for the use in the Portal?	6. Degree of correctness of the language used	7. Overall score for the metadata of this resource	Publish?
SLU	CULS	2	4	4	5	2	5	3	No
UHasselt	ACTA	3	3	3	4	3	4	3	Yes
UHasselt	ARI	4	-	4	4	4	4	3	Yes
UHasselt	AUA	3	-	4	4	4	4	3	Yes
UHasselt	CINECA	3	4	4	4	4	4	3	Yes
UHasselt	CULS	3	4	4	4	4	4	4	Yes
UHasselt	ICROFS	4	4	4	4	4	4	4	Yes
UHasselt	INRA	4	4	4	4	4	4	4	Yes