MOI UNIVERSITY
SCHOOL OF INFORMATION SCIENCES
MASTER OF PHILOSOPHY IN INFORMATION SCIENCES

MANAGING DIGITAL ACADEMIC GREY LITERATURE USING D-SPACE AT STRATHMORE UNIVERSITY LIBRARY

A PROJECT SUBMITTED TO MOI UNIVERSITY IN PARTIAL FULFILLMENT FOR AWARD OF MASTER OF PHILOSOPHY (INFORMATION SCIENCES)

PRESENTED BY: RUTH N. GIBENDI IS/MPHIL/049/09

SUPERVISORS: PROF. JOSEPH KIPLANG'AT
DR. DAMARIS ODERO.

MARCH, 2012

ABSTRACT

The aim of this study was to implement a system for the management of digital academic grey literature using D-Space software at Strathmore University. The specific objectives were to conduct a user needs assessment survey for grey literature, develop a service model to address the identified needs based on an institutional repository framework, establish policies for the repository management and develop strategies for usage of the repository. Data was collected using two online surveys administered to producers and users of grey literature. The results showed that users were currently experiencing access problems due to controlled working hours of the special collections unit of the library, and the difficulty in tracing grey literature. The grey literature producers also wanted a structured way of storing and sharing the grey literature they produce. A grey literature management system was set up using DSpace software. Appropriate policies were formulated to run the repository and marketing strategies established to stimulate usage. The research recommends improvements in future releases of DSpace to include Web2.0 support, integrate name and subject authority control and further studies be conducted on faculty adoption of institutional repository.

DEDICATION

To my mother, Mrs. Leah Gibendi, for her unwavering support through the ages ☺.

To my late father Kadagi Elphas Gibendi, who planted in me an irresistible desire for excellence. I have a long way to go … to get anywhere close to your ideals, but, I'll keep on running towards the goal.

ACKNOWLEDGEMENTS

Many people have contributed to enable this project to be a success. I am deeply indebted to them all, though it may not be possible to mention all of them.

First, my supervisor Prof. Joseph Kiplang'at whose gentle nudging has seen this work to completion. I am immensely grateful to Dr. Damaris Odero, whom I consider to be my role model and mentor.

To Strathmore University and its fraternity, for believing in me and giving me an opportunity to pursue my potential. Further thanks for offering financial support for the completion of the Masters degree. My special thanks to Dr. Gitau Njoroge. I am truly humbled. To Bernard Shiundu, Library Head, Strathmore university. I have always counted on your support and you've never let me down. To Emily Sawe, my colleague and close friend. You helped in a big way to make DSpace a reality. Kiuna, Peninnah, Okwiri, Raymond, Adika, Richy, Chibini, Mary, Kibos, Nyamasege and Phillip – I already miss you!

To my family – I would never have made it this far without your support. Julo, you've always been my beloved and my friend. Your support through the years has seen me rise from strength to strength. I love you. Our children, Moraa, Entu, Manu and Nyancho… thus far the Lord has taken us. Special thanks to Mary and Irene for the endless hours they spend watching over the children as I did the studies. Only God can adequately repay you. To my sister-in-law Evelyne for the tremendous support you gave me while I was working in Nairobi. To my parents-in-law, Mr. and Mrs. Zebedee Toeri. Your trust in me is amazing.

But above all, To him who is able to keep us from falling, and to bring us faultless and joyful before his glorious presence – to the only God our Savior, through Jesus Christ our Lord, be glory, majesty, might, and authority, from all ages past, and now and for ever and ever. Amen. (Jude:24-25).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

DEFINITIONS

Bitstreams: In the context of this document, digital data transmitted between systems, irrespective of format.

D-Space : An open source software that is used for management of institutional repositories.

Institutional Repository: A set of services that an institution offers to its members for management of its digital academic assets.

Grey literature : That which is produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers

Metadata: Information describing an item, sometimes also called 'data about data'

URI – Uniform Resource Locator

Post-print: The final accepted and published version of an article.

Pre-print: The final edited version of an article prior to publication and peer review

Self-archiving The process whereby authors can submit the metadata and full-text item of their own publications into a database

Preservation: In this scenario, ensuring that the digital format is accessible in future by means of migration or any other action deemed necessary.

OAI-PMH : Open Archives Initiative has developed a Protocol for Metadata Harvesting

CHAPTER ONE
INTRODUCTION

1.1 INTRODUCTION

Scientific research is a quintessentially universal public record of human achievement that transcends political, sociological, cultural, and linguistic boundaries. In the scientific process, Weintraub (2000) affirms that every increment of new knowledge adds to, modifies, refines, or refutes earlier findings. Although earlier findings may be useful in order to gain a historical perspective on how the quest for a solution progressed through time, the old knowledge immediately becomes obsolete when it is replaced by the newer developments

Typically, research information in a university is first produced as grey literature, and in some cases, it remains as grey literature forever. Even when published and goes 'white', only a fraction of the entire work goes to print. Grey literature documents unequivocally document the knowledge and know-how of the organization. It represents the cutting edge of research.

Easy access to research and development information provides innovation motivation for industry and commerce leading to wealth creation and employment and simultaneously provides quality of life advantages in healthcare, environmental and cultural aspects. (Jeffery 1999). When such information exists as grey literature, it compels a case for proper management to guarantee acquisition, description, dissemination and preservation. This is more so as it becomes increasingly clear that grey literature documents the knowledge and know-how of the organization and as such is an asset that is extremely valuable.

Not all scholarly writings are published and a multitude exists as grey literature. The Fourth International Conference on Grey Literature (GL '99) in Washington, DC, in October 1999 defined grey literature as: "That which is produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers." Although this definition is fuzzy, it nonetheless offers a useful pointer of grey. In an academic institution, grey literature inexhaustively includes reports,

pre-prints, thesis and dissertations, conference proceedings, student's projects, datasets, and some forms of learning objects.

The concept of 'grey' leads to largely negative associations as evidenced in various dictionaries and thesauri definitions. Jeffery & Asserson (2005) outline some of these connotations associated with grey to include meanings of '*dull and dismal'*, '*obscured*' and '*between states'* implying uncertainty and fuzziness. To some extent, these meanings apply. Grey literature is often obscured – hidden away, poorly catalogued, not advertised and difficult to find. It is often conceived as less important than the white published literature and receives less attention in the information landscape. Yet, paradoxically, grey plays central role in development and transformation to the information age. Jeffery and Asserson (2005) are strongly convinced that if during several latest hundred years scientific information exchange was founded on printed matter, then grey literature becomes the information basis of today's knowledge society.

Grey literature indeed plays a vital role in an academic environment. The conference proceedings for instance are recognized venues for publishing new research findings way before official publication. Some grey formats (like reports) offer valuable case studies. Thesis and dissertations consist of fresh ideas and contributions to knowledge, and indeed offer a lens for viewing an institution's intellectual output over the years.

Crow (2006) naturally expects academic institutions, as producers of primary research, to take an interest in capturing and preserving the intellectual output of their faculty, students and staff. This is sadly not the case in many African libraries. Moahi (2009) presents the poignant reality in Africa where knowledge generated in universities and research centres is either disseminated in expensive international journals, (such that even the original generators of the research have no access to it), or is left to gather dust in offices and computers. After a number of years, tragically, such studies are replicated without knowledge that they had been carried out before. The far-reaching effects of neglect of access to knowledge ripples itself to university rankings. African universities are generally ranked lowest in terms of research output. The World University rankings ( ) in 2009 ranked the highest university in Africa at position 359 out of 6,000 universities. Strathmore

university is ranked 12 in the Africa and first in Kenya. Clearly, research visibility from the continent is appalling, and this is something that could be addressed by development of institutional repositories.

In the print era, librarians have had difficulty in the management of grey literature. Acquiring grey has often posed problems as these collections are not intended for commercial distribution and hence are not easily found in catalogues. Where acquisition has been successful, some libraries have faced great backlog and ended up with minimally described collections. This has often made retrieval and use for patrons very difficult. In most libraries, the grey collections are not integrated with the main collection and often are put in under closed access sometimes as 'Special Collections'. This in turn has presented challenges of access and use. No doubt therefore that grey literature has come to viewed as non-conventional, insignificant and sometimes ephemeral.

In the digital era, most academic grey is 'born digital'. Moahi (2009) concurs with this view saying that much of the knowledge produced in Africa is usually in digital form given the ubiquity of ICTs in many universities. In some instances, the print counterpart of this grey does not exist, as would be the case with some datasets and computer programs. The challenge is that this information is not captured and organized for easy access and use by others. It is only when this is done that African grown knowledge will become visible. The trend worldwide has been to establish information repositories to address visibility and accessibility.

1.2 BACKGROUND

Strathmore University is a private institution of higher learning that was chartered in June 2007. The University's charter was gazetted under the Universities Act (Cap 210B) in Kenya Gazette supplement no. 47 (legislative supplement no 27).

Strathmore University's vision is to be a centre of academic and professional excellence that provides all-round education in an atmosphere of freedom and responsibility. In its mission, the university dedicates itself to the advancement of education through teaching, scholarship

and service to society by providing an all-round education in an atmosphere of freedom and responsibility, creating a culture of continuous improvement, fostering high oral standards and developing a spirit of service and respect for others.

Strathmore University aspires to provide high quality and all-round education, which is geared towards forming students of high standards professionally, academically, morally and spiritually in an atmosphere of freedom and responsibility.

In doing this, Strathmore University will develop high standards of admission, tuition and examination administration. It shall also adopt the requirements of all relevant professional and examination bodies.

The University is ISO 9001:2000 certified. It is committed to continuously improve the effectiveness of the Quality Management System to ensure that it fulfils its purpose. The university has eight departments and a research centre, as well as an arm that provides consultancy services, serving a student population of 4,000 students. With this growth, and the commitment to improvement through quality management system, there is little doubt that research has to play a central role. The research outputs too need a sober system of management so as to minimize duplications, as well as make it visible to the entire world. Such a role would squarely be the library's.

Strathmore University Library comprises of a new library complex with a floor space of 2500 square meters. There is also a university library reading room that has a floor space of 450 square meters that has a seating capacity of 300 people.

The atmosphere in both buildings is highly conducive to private study and research. A wide range of course reference works and books of general interest available in the new two-story complex contributes to the university's excellent academic performance.

The library's vision is to provide information resources in support of the teaching, research and community endeavors of Strathmore University. In its mission statement, the library

commits itself to collect information resources and services and make them readily accessible, so as to encourage learning, research and improvement throughout life. The library strives to spread a culture of solidarity that will uphold the dignity of the human person and family values, and assist in preparing students to become competent professionals who can enrich society with their knowledge, initiative and personal responsibility.

Being at the service of the university, the library aims to support scientific research, quality teaching and community service by building up a qualitative collection of printed and non-printed materials, equipping itself with appropriate Information Communication Technologies (ICTs), and linking subscribing to library reference services

The new library, designed in accordance with the Standards and Guidelines for University Libraries issued by Kenya's Commission of Higher Education, has a seating capacity of 540 people and book stack areas for 120,000 volumes, an audio-visual section, and offices.

The library consists of three departments: Information Services, Technical Services and Information Literacy. The information services section is in charge of circulation, special collections and periodicals, audio visual and electronic services department.
Understandably, most grey literature is currently handled by the Special Collections department. The department is in charge of managing thesis, some learning materials like past examination papers, sample company reports, government publications, research publications from institutions such as IPAR, UN and the World Bank, newsletters both internal and external ones and other research publications.

The challenge of managing grey literature therefore lies in the jurisdiction of the Special Collections department. In the past, users have expressed dissatisfaction with the department's working hours. The department works from 9am to 9pm in the evening, and is closed on weekends. Evening students and those with weekend classes are hardly able to use the collections hosted in the department. This has motivated the department to think outside the box in relation to its collections. It has motivated the birth of an institutional repository that would be capable of holding some of the high-demand collection online, hence making

them accessible to students at their own flexible working hours. There has also been a desire to now capture not only the learning materials and dissertations, but also grey literature that is scattered within individuals workstations and the departmental sites to one online portal.

1.3 STATEMENT OF THE PROBLEM

Despite the fact that grey literature plays a key role in academic and research undertakings, it is currently is not represented equally well in the established information landscape.

Most libraries invest little effort in grey literature management in totality: little is done to acquire them, little done to describe them, little done to enhance access and use. The most common form of grey collected in higher learning institutions is thesis and dissertations. Even then, they are accorded basic description and access to them is often difficult and on closed access. Yet it is grey literature collections that *uniquely* define a library collection. Strathmore University library suffers no exception to these problems.

At Strathmore University, grey literature acquired is mostly in form of the thesis and past exam papers, which are then kept under closed access, yet a considerable proportion of students can only access them in the evenings when they come to college. This has seen underutilization of these resources. Lots of grey literature published by staff members lie in their individual desktops. Some grey is also found at the university web site with no clear preservation agenda. After a number of years, links to some content are dead. At the same time, there is no consistent preservation format and no meta data description to the items. Locating grey thus becomes a serendipitous task.

Poor management of grey has resulted generally to minimum visibility and hence minimum use. As a result, it is easy for students and researchers to repeat research works and to plagiarize works without being noticed. This translates to mockery to the very foundation of research.

Given the grim scenario, an institutional repository is deemed a necessity to address emerging challenges in management of grey literature at the university.

1.4 AIM

The project aims to implement a system for management of digital academic grey literature using D-Space software at Strathmore University.

1.5 OBJECTIVES

1. To conduct a user needs assessment survey for grey literature.
2. To develop a service model that addresses identified needs based on an institutional repository framework.
3. To establish policies for digital grey literature management.
4. To develop strategies to facilitate usage of the repository.

1.6 RESEARCH QUESTIONS

1. Who are the users and producers of grey literature at Strathmore University?
2. How is grey literature currently managed to be accessible to end users?
3. Can an institutional repository be used to manage grey literature collections?
4. How will issues of intellectual property, preservation, and access be handled?
5. What can the library do to ensure use of the repository by both users and producers of grey literature?

1.7 SIGNIFICANCE OF THE PROJECT

The project will help leverage management of grey literature at Strathmore University. It will centralize, preserve and make accessible the intellectual capital of Strathmore University, first to its members and more importantly to the entire globe. The repository will therefore form part of a global system of distributed, interoperable repositories. As a result, the ranking of the university in the World University Rankings may improve.

The project will also serve as a model to other universities in Kenya. If they duplicate the model, they will immensely contribute to the visibility of scholarly works from country, as well as save the cost spend in duplicating research works.

CHAPTER TWO
LITERATURE REVIEW

2.1 INTRODUCTION

Knowledge is an indisputable resource for development of organizations. Jeffery and Asserson (2005) put forth a bold case for grey literature. They claim that if during several latest hundred years scientific information exchange was founded on printed matter then grey literature becomes the information basis of today's knowledge society. Given the case, then there is doubtless a need for management of this resource. Information, if not managed, depreciates in value, as it may not be put to use.

The role of managing grey collections squarely lies with librarians. This view is unambiguously held by a number of researchers. Gelfand (2003) concurs noting that although not all grey literature is worth collecting, much of what is worth collecting is still falling through the very large cracks of standard library acquisition practice. The role of the librarian is acknowledged as that of promoting dissemination and use of grey literature through cataloguing, searching, archiving and preservation. (Rabina 2007). Library and information professionals thus remain the vital link in the chain that makes grey literature available to researchers, students and the interested public. This views are similarly held by Mackenzie Owen (1997). If the librarians fail in this job, grey could be lost forever.

However, the very notion of grey is linked conceptually with dull and dismal. Although not peer-reviewed (as is white literature), it is not usually the unstructured, unauthorized output from a single source. Often it may well be commercial in confidence and certainly is likely to contain intellectual property of value to the organization. Jeffery and Asserson (2005)

2.2 DEFINING GREY LITERATURE

The most common definition of grey is the one given by the Fourth International Conference on Grey Literature (GL '99) in Washington, DC, in October 1999. Grey literature was de-

fined as information produced on all levels of government, academics, business and industry in electronic and print formats not controlled by commercial publishers and where publishing is not the primary activity of the producing body.

The Online Dictionary of Library and Information Science as quoted by Rabina (2007) provides a slightly different definition of the term. Grey literature is defined as "Documentary material in print and electronic formats, such as reports, preprints, internal documents (memoranda, newsletters, market surveys, etc.), theses and dissertations, conference proceedings, technical specifications and standards, trade literature, etc., not readily available through regular market channels because it was never commercially published/listed or was not widely distributed1."

Defining grey using both criteria opens a Pandora box for inclusion of all types of literature, including ephemeral ones not worthy of management. Pavlov (2004) notes that grey literature is a fuzzy set rather than a deterministic one and it is difficult if possible to define it strictly. This notion is embraced by other writers (Lambert *et al.*, 2005) who agree to the generally accepted notion of the need for narrowing the definition for particular contents.

While definitions proliferate, there is agreement on the main characteristics of grey literature: they are materials that are published not for profit and, as a result, typically not marketed or distributed by commercial publishing organizations. (Pavlov, 2004). General characteristics frequently associated with grey include difficult in identification, access and location. Grey literature often appear in limited editions, are often inaccessible in bookstores, often lack bibliographic registrations, are absent in library collections and catalogues and are rapidly distributed. (Nakohto, 2006) The most important notion therefore is to limit the definition and apply it in the context of the environment being studied. What could be ephemeral grey in one discipline could well be fundamental grey in another.

2.3 GREY LITERATURE IN THE ACADEMIC ENVIRONMENT

Siegel (2003) defines academic grey literature (scholarly grey) as that which is produced as a result study or inquiry but which is not published through the traditional channels of books or

journals. Since it receives limited distribution, it is often not given an opportunity for wider dissemination through traditional bibliographic treatment. This renders it, just like any other grey, to be invisible.

Academic grey literature is an important research output. Indeed, it forms the documented 'know-how' of an organization. (Jeffery and Asserson, 2006). In an academic institution, it forms the larger bulk of the intellectual capital of the institution.

Siegel (2003), laments of poor management practices adopted by librarians in managing grey. He says that historically, many academic libraries have shunned grey literature for a variety of reasons. Even when attention has been paid, academic concerns about collection development and grey literature tend to focus on external collections from industries or agencies and, more often than not, these collections are oriented to a particular subject or discipline. In short, academic librarians do little to harness the grey found within their confines.

This has resulted to a situation where thesis and dissertation are the most frequent grey managed by libraries. Stock (2007) attributes this presence to the fact that thesis and dissertations are a well defined and referenced document type. Rules for deposit are established at a university level and students are often 'obliged' to deposit a copy formally.

Sen (2008) presents a strong case for harnessing grey literature which is of particular interest to the academic users. This is linked to the emphasis on the review of documented past experience as a guide to what works. He argues that:

- There is little that is truly new in this world, either problems or solutions.
- A lot of time and effort is wasted on reinventing wheels.
- Too little is invested in making full use of research findings by placing them into context with other similar studies, and identifying strong messages.
- It is sensible to take advantage of past experience and knowledge, not just for lessons about 'what works' and 'what doesn't work', but 'why and in what contexts'.

Brown (1999) as cited by Ranger (2004) surveyed astronomers, chemists, mathematicians and physicists, and found out that while groups of professors in these disciplines do not often use grey literature (pre-prints and conference proceedings) for teaching, they do use it extensively in their research.

2.4 ROLE OF GREY LITERATURE

As primary sources grey literature provides un-interpreted, first hand accounts or evidence of an event or experience. These sources contain information or data and are usually written at the time of the event or research. They are usually the original source of information and allow the researcher to analyze a topic without another person's interpretation. (Bhahrati 2007)

Weintraub (2000) contends that grey literature is ubiquitous: citizens need grey literature to make informed decisions about government and the other institutions they deal with on a daily basis. For instance, if their city is planning a road project they might want to look at any related reports commissioned by the city. The literature shows that while researchers are often mentioned as frequent users of grey literature, civil servants, teachers, students and the general public also use grey literature resources, a view supported further by Ranger, 2004)

According to Jeffery and Asserson (2005), the dynamics of the landscape concern the way in which an idea, concept or knowledge is generated and transformed: from grey to grey (internal discussion within an organization with improvement), from grey to white (publication, public relations for an organization, improved evaluation scores for an organization), from grey to product or service (wealth creation or improvement in the quality of life within an organization), from white to product or service (wealth creation or improvement in the quality of life by knowledge or technology transfer).

2.5 CHALLENGES ASSOCIATED WITH MANAGING GREY LITERATURE

If grey literature is as luminescent as illustrated, then how come it still suffers from deep invisibility and use? Mostly, the answer lies in challenges associated with this fuzzy set. The most common problems associated with grey relate to acquisition, organization, training, bibliographic control and provision of access.

One characteristic of grey literature is that it is often difficult to acquire through normal bookselling channels. because it lies outside of commercial circuits of publication and dissemination. (Nakohto, 2006) For library patrons therefore, the search for and acquisition of this kind of "underground literature" can be a time-consuming, sometimes expensive and even frustrating experience. Boukacem-Zeghmouri and Schöpfel 2006 ascertain that often the acquisition of grey literature taken to the periphery, to an extend that most libraries do not even include it in their acquisition policies.

Managing grey is also often faced with another major problem - training. Grey literature management is not a subject generally dealt with in formal library training. (Gelfand, 2003). Most librarians learn it 'on-the-job'. Rabina (2007) conducted a survey on students enrolled in ALA-accredited library and information science programs to find out what they knew about grey literature and where they learnt it from. Her findings revealed that knowledge about grey literature is gained across the curriculum and that knowledge about grey literature is more intuitive and anecdotal than systematically acquired. This view is also firmly held by Gelfand (1998) who believes that the roles related to grey literature are learnt on the job. Yet, the librarian's role in the grey literature landscape is described as managing, promoting dissemination and use of grey literature through cataloguing searching, archiving and preservation (Mackenzie Owen, 1997).

Bibliographic control is also often another major difficulty encountered in managing grey. Unlike the white published counterparts, grey literature often has no cataloguing in print data available (Ranger, 2004). This often means that original cataloguing has to be performed on the items. This can be very time consuming. In a survey conducted on fifteen academic and special libraries, she found out that discrepancies also exist in grey literature bibliographic control. Some libraries catalogued each and every item in their collections, others did it at series level, and some were completely uncatalogued. The users were thus required to have had prior knowledge of the existence of the item before it could be retrieved to them. Series level cataloguing makes a lot of otherwise relevant hits to be invisible. Cataloguing is the heart of retrieval, and an uncatalogued item is as good as nonexistent.

Access to grey literature is another major challenge. Grey literature collections are often kept in special collections and hardly integrated to the large library collection. (Gelfand, 2003). Most special collections offer their items under closed access and work for restrictive hours that may not favor the end users. Yet, as Ranger (2004) observes, for the user, the only available information sources are those he can 'see'. In other words, grey literature must be shown to its potential users and it must be readily available. The general effect of this has been low usage, a view also augmented by Auger (1998).

2.6 DIGITAL GREY LITERATURE : CHALLENGES AND OPPORTUNITIES

Electronic forms of grey literature have greatly increased the number of grey literature documents. Including electronic items expands earlier definitions of grey literature. Rather than only physical objects such as paper, video, and microfiche, grey literature now includes a much larger array of documents, including websites, electronic datasets, electronic reports and articles, and software. Ranger, 2004**.** However, "Grey Literature must be distinguished and distinguishable from junk e-mail, vanity press, and other gigabytes of e-trash on the net." (Gelfand, 2000, n.p.)

There is no menace to grey literature in the Internet, on the contrary, the web has enhanced its role. Pavlov (2004) takes the view of other authors who have described the web as the 'new classical grey literature'. The Internet did change our information and communication habits, there is a bulk of grey literature on the web sites but it doesn't mean that information professionals are supposed to quit our traditional activities of collecting, archiving, and disseminating grey. The Internet is just a tool, a technical and technological instrument that solves some problems and raises others but they are not specifically grey. Organization and access to grey is one such problem, and it can not just be wished away.

Artus, as quoted by Pavlov (2004) opines that with internet, the technical facilities, in relation to grey literature has widely changed but, as of the economic structures and social functions of grey literature, only little has changed – if anything at all". He goes as far as supposing that internet technologies could blur the boundaries between white and grey literature and we would expect the white literature to perish in the long run and the grey with it. What remains is only 'literature' without any further classification as white and grey. While

Artus views could termed to be rather ambitious, what is clear is that grey literature's adaptation depends on its tactical approach to existence in digital forms. Jeffery and Asserson (2005) propose management of digital grey that is composed of a system that provides formal metadata and contextual information, a repository of research datasets, software, grey and white objects, and a workflow system that eases the threshold effort to get material to the repository. This, perhaps, are the ideals of managing digital grey.

However, several aspects of the digital problem quickly come to play. Starovoitov (2008) identifies these problems at four levels : financial, legal, technological and administrative. The financial aspect is evident – any new technology needs investments and the conversion of federal collections to electronic form needs federal funding. Legal aspects are diverse and include the necessity to develop new standards, laws, instructive documents and copyright practices. Digital technology suggests the introduction of online operating modes in network environment both for the authors and the users of documents, context search methods and computerized content analysis, computer-aided subject heading assignment, editing and proofreading.

To reach the aspirations set forth by Artus (2004) new administrative approaches will be mandatory to facilitate the complicated problem solution : and the repository framework has proved to be home to grey management in the digital era.

2.7 INSTITUTIONAL REPOSITORIES AND GREY LITERATURE

An institutional repository, as defined by Crow (2006) is a digital archive of the intellectual product created by the faculty, research staff and students of an institution and accessible to both users within and outside the intuition, with few if any barriers to access. It is a formal and managed archive of research output in the form of digital documents that is operated by a particular institution such as a university.

Crow (2006) further gives four attributes that characterize an institutional repository as being institutionally defined, containing scholarly content, the repository being cumulative and

perpetual, and open and interoperable. In contrast to discipline-specific repositories and subject-oriented or thematic digital libraries, institutional repositories

The ideals of a repository are pegged on organizational responsibility. Articulating this, Ranger (2004) says that if each country or region chooses a depository for its grey literature, publicizes the depository, the depository actively solicits documents from as many sources as possible and describes and catalogs its collection, grey literature as a whole would be in a good position to become a well-used resource. Original research and other intellectual property generated by an institutions' constituent population active would be captured. This way, an institutional repository would represent an historical and tangible bodiment of the intellectual life and output of an institution. This way, institutional repositories are perfect for capturing grey literature right from home, where it is produced.

Managing digital repositories still is the role of the library. Affirming this, Crow (2006) feels that in the long-term, organizing and maintaining digital content—as well as supporting faculty as information contributors and end users—should remain the responsibility of the library. Libraries are best-suited to provide much of the document preparation expertise (document format, control, archival standards, etc.) to help authors contribute their research to the institution's repository. Similarly, libraries can most effectively provide much of the expertise in terms of metadata tagging, authority controls, and the other content management requirements that increase access to, and the usability of, the data itself.

Depending on the goals established by each institution, an institutional repository could contain any work product generated by the institution's students, faculty, non-faculty researchers, and staff. This material might include student electronic portfolios, classroom teaching materials, the institution's annual reports, video recordings, computer programs, data sets, photographs, and art works—virtually any digital material that the institution wishes to preserve. However, given the focus on scholarly communication and on changing the structure of the scholarly publishing model, institutional repositories are defined here - whatever else they might contain - as collecting, preserving, and disseminating scholarly content. This content may include pre-prints and other works-in-progress, peer-reviewed articles, monographs, enduring teaching materials, data

sets and other ancillary research material, conference papers, electronic theses and dissertations, and other forms of grey literature.

Institutional repositories build on a growing grassroots faculty practice of posting research online, most often on personal web sites, but also on departmental sites or in disciplinary repositories. This demonstrates a desire for expanded exposure of, and access to, their work. In addition, digital publishing technologies, ever-expanding global networking, and enabling interoperability protocols and metadata standards are coalescing to provide practical technical solutions that can be implemented now. The convergence of these interrelated strands indicates that institutional repositories merit serious and immediate consideration from academic institutions and their constituent faculty, librarians, and administrators. Johnson (2002)

Jeffery and Asserson (2007) outline five attributes of an institutional repository that are fundamental to make it work.

Ease of deposit : The first requirement – without which no grey literature repositories exist – is that it must be easy to deposit grey objects with associated metadata. This is best achieved by reducing the threshold barrier for the end-user; appropriate techniques include workflow and incremental metadata provision, automated metadata extraction and re-use and knowledge-assisted metadata input so reducing the amount of actual input required from the end-user.

Easy to retrieve **:** The major purpose of a grey literature repository is to provide grey objects upon request; the request usually being for one or more 'full text' objects (it may of course be hypermedia) of which the metadata satisfies the criteria of the search. It is clear that the relevance (accuracy, precision) of the response to the request and the recall (completeness) of the response both depend critically on the quality of the metadata.

Easy to transition : Grey objects usually record a stage or step in a research process. They are related to earlier grey material and white material. They are related to future grey and white material. The relationships or linkages are critically important to understanding the relevance of a grey object to the request (or more specifically to the intent of the request). Moreover tracking the evolution of research thought through a time-series of grey and white objects is instructive. It may also be important legally in claims of prior publication and other rights.

Easy to track provenance: Related to the above, provenance information may improve the ability of the end-user to assess the quality and relevance of the grey object of interest. Provenance includes not only predecessor objects but also the contextual information mentioned above. It is important to have correct recording of versions of a grey object for the same reasons. Finally, for provenance information to be provided it is essential that facilities for preservation and curation are provided.

*Easy to relate :*From the foregoing it is clear that relationships (hyperlinks) are of the greatest importance for effective and efficient provision of grey resources. The relationships concern not only other objects in the grey repository (which should be open access and institutional).

According to Lambert *et al.* (2007), an institutional repository, being a central point within the organization for literature and data, is a component of the integration of processes, which promises benefits both to the organization itself and to the researchers within it. Grey literature can be recorded and retrieved according to accurate and up to date personnel, project and business unit structure, leading to benefits in accuracy of reporting, quality control, etc. From the researchers' point of view, the repository assists them by storing multiple expressions or manifestations for the different parts of the process, and so brings clarity to work and version relationships. It also allows retrieval of grey literature from other projects that might otherwise be invisible to those who could benefit from it. In short, the institutional repository is integrated as a part of the overall institutional memory.

Institutional repositories, by capturing, preserving, and disseminating a university's collective intellectual capital, serve as meaningful indicators of an institution's academic quality. Crow (2006) notes that under the current system of scholarly communication, much of the intellectual output and value of an institution's intellectual property is diffused through thousands of scholarly journals. While faculty publication in these journals reflects positively on the host university, an institutional repository concentrates the intellectual product created by
a university's researchers, making it easier to demonstrate its scientific, social and financial value.

To sum it up, Digital repositories have a number of functions or foci (Swan, 2007)

- To open up and offer the outputs of the institution or community to the world
- To impact on and influence developments by maximizing the visibility of outputs and providing the greatest possible chance of enhanced impact as a result
- To showcase and sell the institution to interested constituencies – prospective staff, prospective students and other stakeholders
- To collect and curate digital outputs (or inputs, in the case of special collections)
- To manage and measure research and teaching activities
- To provide and promote a workspace for work-in-progress, and for collaborative or large-scale projects
- To facilitate and further the development and sharing of digital teaching materials and aids
- To support and sustain student endeavors, including providing access to theses and dissertations and providing a location for the development of e-portfolios

2.8 REPOSITORIES IN AFRICA

According to Anbi (2002), access to scholarly information is the biggest problem in African Universities. The dearth of Research and Development programs in African universities, access barriers to high priced scholarly literature and lack of quality research infrastructures have dampened the research spirit in African Universities. The west-centred scholarly paradigm and the high costs of scholarly literature, such as journals and books, have kept the intellectual output of Africa under constant check. A close look at the costs of journals in major scientific fields will clearly portray the fact that access to a single journal in a particular field itself is very hard to obtain, while procuring enough scholarly information in that particular field remains highly impossible.

Moahi (2009) in giving his case for institutional repositories, says that Africa actually produces knowledge. And that much of the knowledge that is produced is usually in digital form given the ubiquity of ICTs in many universities and research centers. However, the challenge is that the information and knowledge is not captured, organized for easy access

and use by others. With developments in digital scholarship, more and more scholars are creating content in digital form. However, for the most part, that content may end up in some out of reach, expensive journals, or in the author's computer or even in a subject database – really out of reach of most academics that would benefit from its use. Further, this content that is kept by authors is not guaranteed long term preservation and curation. It is partly to address this challenge that repositories were first mooted.

Anbi (2002) proposes that with the current status of Scholarly communication in general and the African Research status in particular and the advantages of Institutional Repositories the need to establish such repositories in African Universities is highly essential. The advantages and success of Institutional Repositories around the globe clearly ignites the case for initiating such repositories in African Institutions also. With access and availability of scholarly information on the downside and with the visibility of African scholarship also too low, it becomes imperative that the institutional repositories are the only possible remedies for these institutions. The main role of these repositories will be to collect all the intellectual outputs of each institution and publish them in their respective servers and preserve them for long-term usage and dissemination. There is no doubt that with Africa being a vast and diversified land, with promises of high research potential, it will certainly gain in leaps and bounds once the research outlets are opened through the initiations of these institutional repositories.

CHAPTER THREE
RESEARCH METHODOLOGY

3.1 INTRODUCTION

This chapter outlines the research methodology employed in the project. It is aligned to address the underlying research problem posed in Chapter one: the possibility of managing digital grey literature using an institutional repository model.

The chapter also justifies the choice of methodology applied to conduct the research as well as details the methods that were used in data collection and analysis.

3.2 CASE STUDY METHODOLOGY

The case study method allows the researcher to investigate a contemporary phenomenon within its real-life context when the boundaries between phenomenon and context are not clearly evident, and in which multiple sources of evidence are used (Yin, 1994). Case studies are therefore are used to provide illustrations of situations thought to be of wide interest for purposes of getting best practice or provide exercise in problem solving. (Jennings, 1997). The unit of analysis of the case was Strathmore University.

3.3 RESEARCH DESIGN

Yin (2004) elaborates that data to be used in case study research can be sought from both statistical and non-statistical sources. That is, a case can be qualitative or quantitative. For the purpose of this research, the case study was built quantitative.

A survey design was used to collect data for the research problems identified in chapter one.

Two surveys were undertaken – one targeting the end users of digital grey literature, and another targeting the producers. This was with a view of understanding their needs so as to design a system that is responsive to the needs identified.

3.4 POPULATION AND SAMPLING

The case study is confined to Strathmore University as its unit of analysis. The study population was therefore composed of students and staff of Strathmore University. At the

time of undertaking the project, the student population was 4,000 students and the staff population was 254.

The study employed a pure random sample to gather data from the population by running two surveys targeted at different populations. This was found appropriate to gather the general system requirements for the end users of the system.

3.5 DATA COLLECTION INSTRUMENTS

Questionnaires were the main instrument used to collect data. To collect data from the grey literature users, an online questionnaire was employed. Closed ended questions were preferred as this would aid in structured needs analysis. Questionnaires were administered online through the main library catalogue page such that it was easily visible to the library users. The instrument was piloted for one week, the questions were refined and the final instruments run for a month. The results of the pilot were not used in data analysis. The survey was posted for a duration of one month. The online questionnaire was found appropriate as it suffered minimal problems of non-return / non-completion. The online questionnaire also enabled real-time capture of data. The questionnaire featured mostly closed ended questions to enable ease of analysis, but also contained open ended questions to gather additional system information.

The second survey targeting post-graduate students and academic staff as producers of digital grey literature employed mailed questionnaires. The purpose was to limit respondents to only grey literature producers.

3.6 DATA PRESENTATION AND ANALYSIS

Statistical measures of central tendency were employed to analyze data collected. The data collected was collected and summarized in tables and graphs. Percentages and means were used to analyze the data collected. The detailed data analysis is presented in Chapter 4.

CHAPTER FOUR
DATA PRESENTATION ANALYSIS

4.1 INTRODUCTION

This chapter presents the results from the data collection instruments which were administered. The purpose of the survey was to generate the user needs profile that would consequently aid in the creation of a system to address the needs specified.

Two online surveys were administered. The first questionnaire was aimed at the end users of grey literature, while the second was aimed at producers of grey literature. The first questionnaire was administered to get the preferences of the end users as far as their information needs were concerned. The second questionnaire was administered to get the views of the producers of grey literature so as to create a system its subsequent management (acquisition, storage and dissemination. Below is a summary and analysis of the two surveys

4.2 GREY LITERATURE USERS SURVEY

4.2.1 DEMOGRAPHIC DATA

The online survey was administered from the catalogue home page (so that users could see it before searching the catalogue). The survey instrument was run for a period of three weeks between June and July 2011.

The survey elicited 314 responses, which was considered a representative sample of the target population of 4,200.

The bar graph on the next page visually summarizes the composition of the respondents.

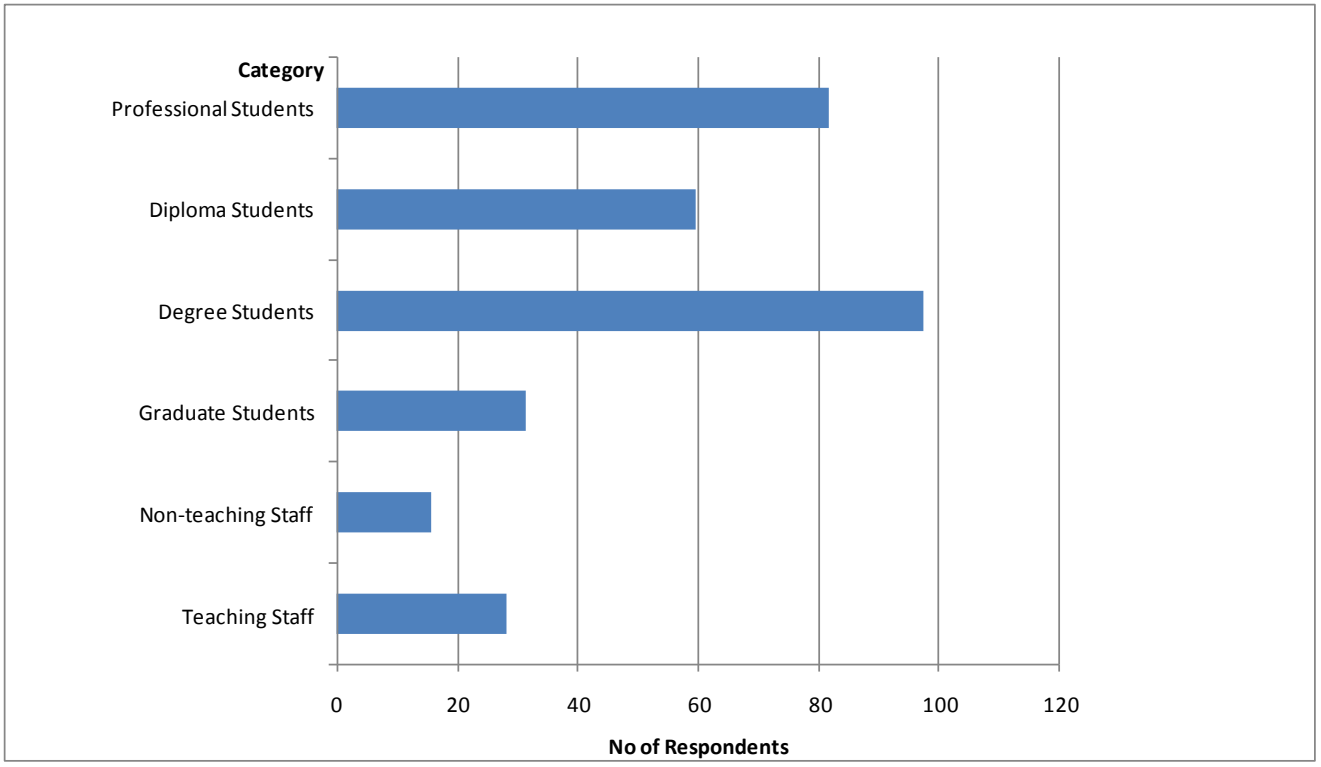


Figure 1: *Demographic profile of respondents*

The survey also sought to find out the current year of study for the students and for staff members, the duration they had been employed at the university. 19.4% of the respondents were with the university for one year or less. 39.6% of the respondents had been in the university for 2 years, 27.30% had been in the university for 3 years, 13.4% had been in the university for 4 years while only 0.3% had been in the university for more than 4 years.

4.2.2 USE OF GREY LITERATURE

The survey sought to find out exposure of the respondents to grey literature. The participants were asked to select from a list of grey literature the type that they had used. The list consisted conference papers and presentations, theses and dissertations, past examination papers, revision kits, students projects, software codes and datasets. The table below summarizes the general use of grey literature by the participants.

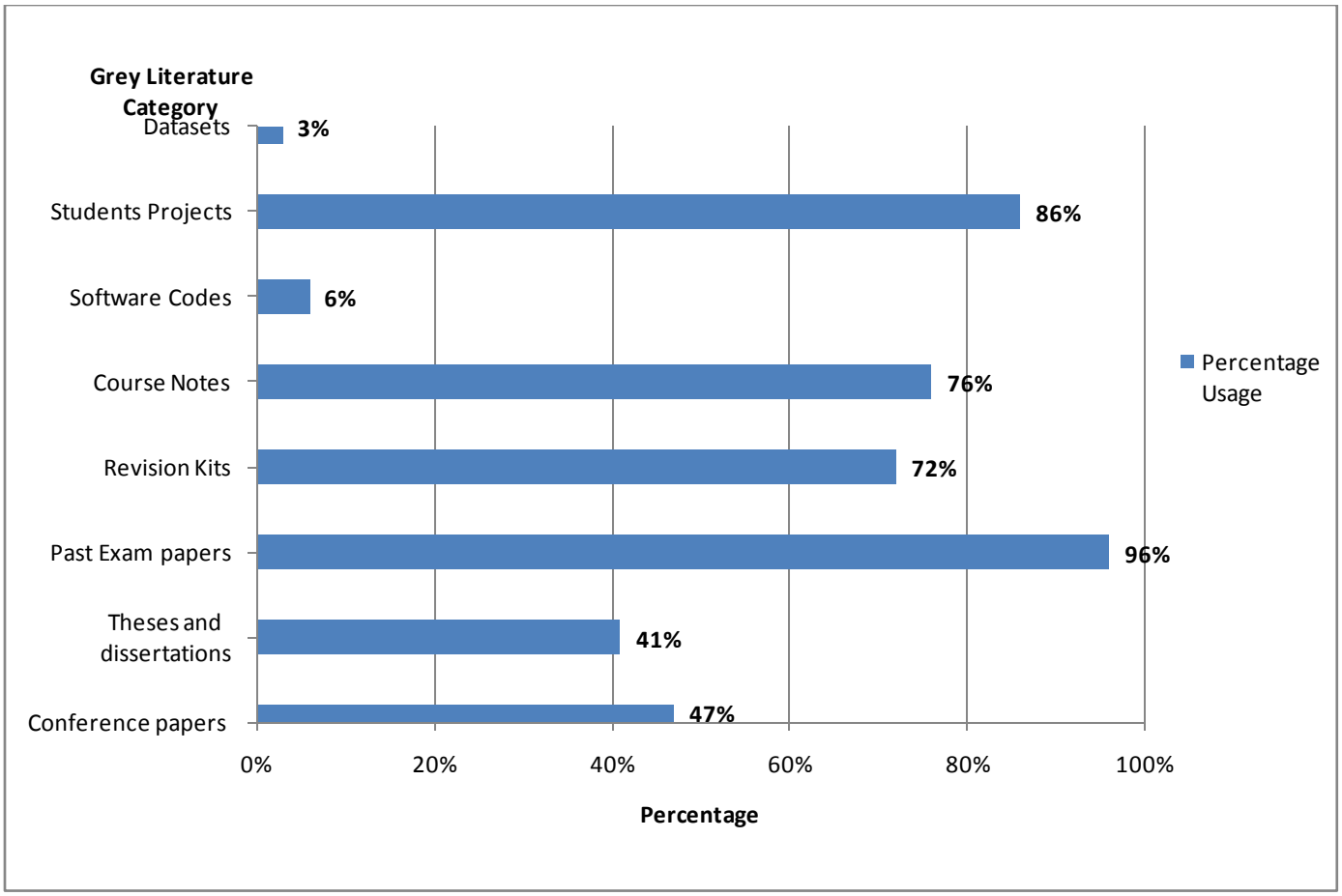


Figure 2 : *Grey Literature usage by category*

The results indicated that there was general use of all grey literature sources, with different categories using a particular kind of grey literature more than others.

Upon further classification by category, the patterns of usage tended to reflect the user category. Teaching staff tended to use conference papers and articles more, followed by theses and dissertations, course notes and past examination papers. Graduate students on the other hand placed maximum value at accessing theses and dissertations, software codes, conference papers and past examination papers respectively. Degree students placed greatest emphasis in accessing past examination papers, followed by course notes, students' projects and course notes. Professional students ranked the revision kits highest, followed by course notes and past examination papers. Diploma students' usage mirrored closely that of the

degree students. The use of datasets seemed confined to selected staff and post-graduate students who ranked it highly with the sharp contrast to the non-users.

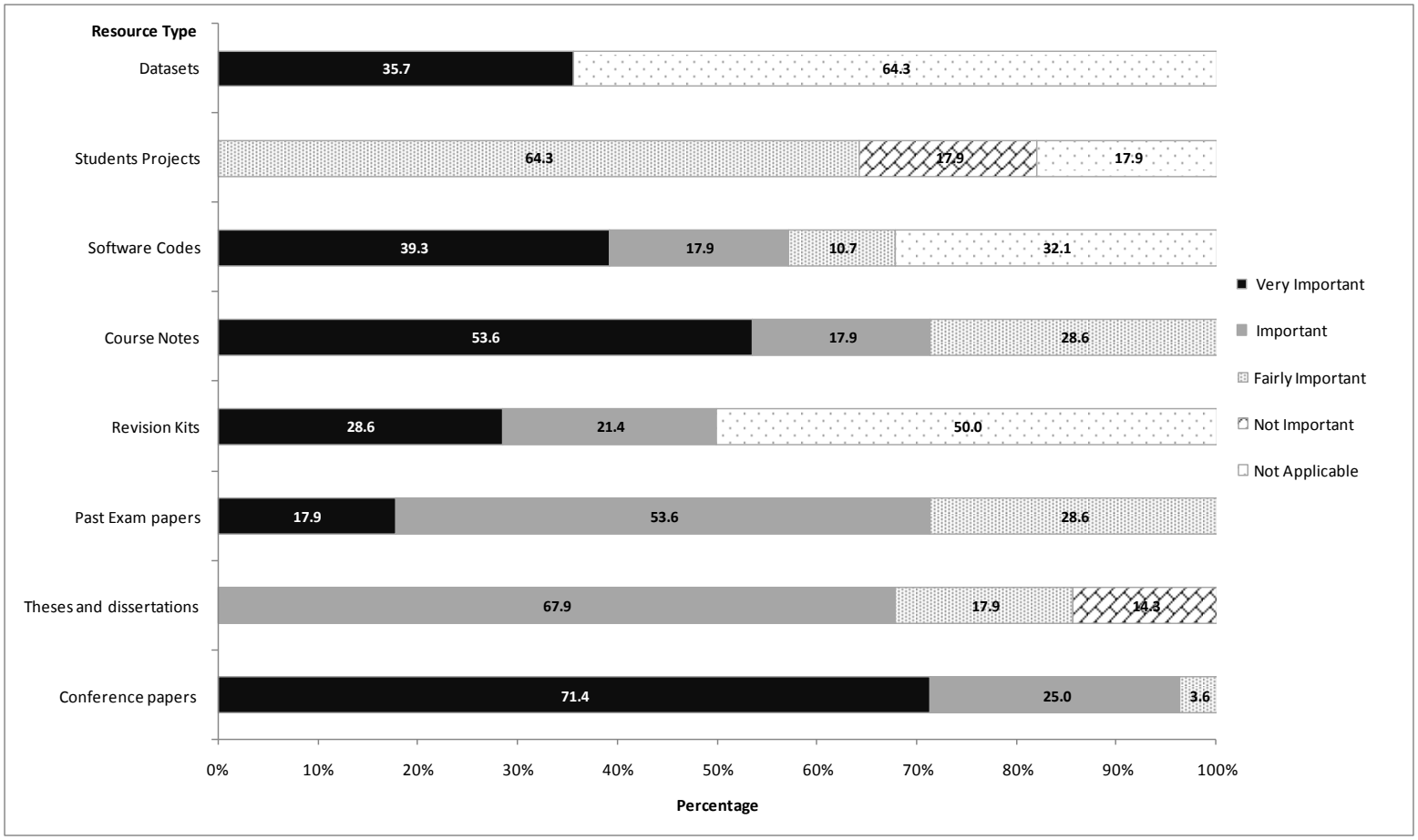


Figure 3 : *Faculty rating of Grey Literature Sources*

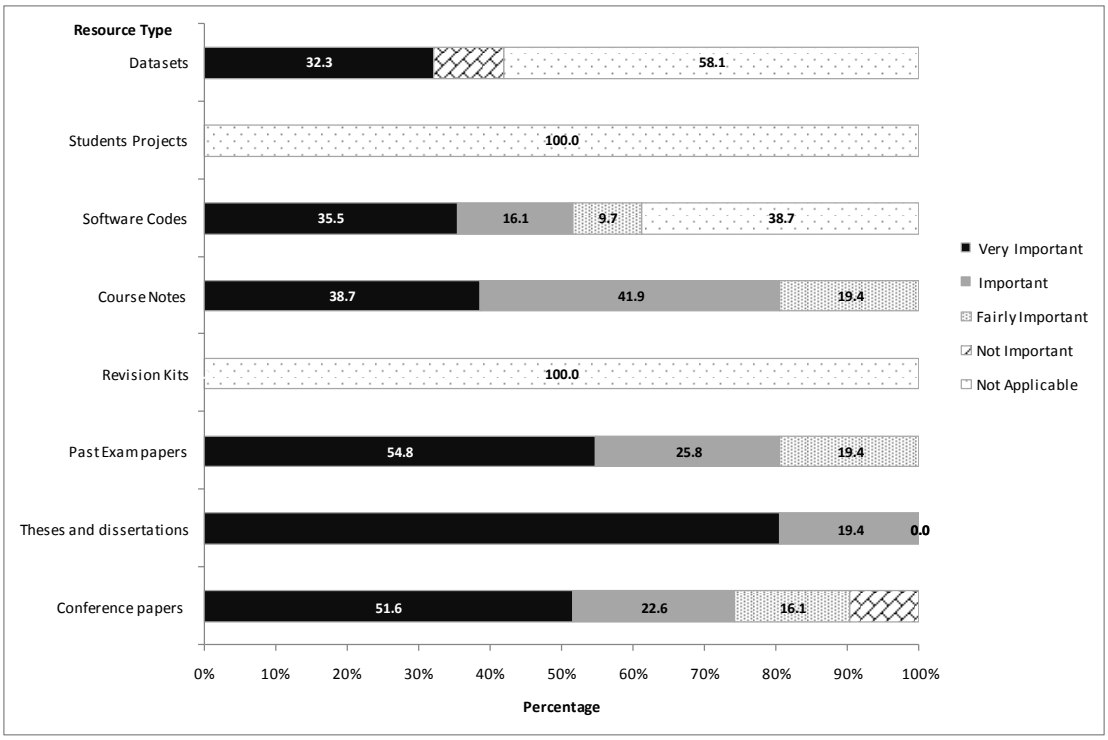


Figure 4 : *Graduate students' rating of Grey Literature Resources*

4.2.3 SOURCES OF GREY LITERATURE

Participants were asked to state where they access the digital grey literature from. Of the 183 respondents, 88% of them accessed grey literature from the Library Special Collections, 92% of them from the internet and 43% of the respondents access it from colleagues. The internet and library feature therefore as the most common sources of digital grey literature. Other sources listed by participants included sourcing from their lecturers, from government departments and from other libraries.

The internet as a core source of digital grey literature was thought to have a positive influence on the proposed system as it was already largely familiar and used by the users.

4.2.4 PROBLEMS ENCOUNTERED IN ACCESSING GREY LITERATURE

The survey sought to find out the difficulties encountered by users in accessing grey literature as well as seek suggestions to the identified difficulties. Users were asked to for a the extent to which they felt particular statements were correct in relation to use of the special collections. 276 respondents (88%) identified the special collections working hours as a

hindrance to their access of grey literature. The special collections policy which hinders users from borrowing items from the collection was found limiting by a total of 281(89%). Diploma and undergraduate students ranked the crowding for resources as a big problem at 82% and 84% respectively. The problem of difficulty in identifying what has been done was more pronounced in the faculty and post-graduate respondents with responses in favor standing at 83% and 85% respectively. Access to resources by evening students was rated a problem largely by all categories of users with by a cumulative 255 respondents (92.75)% identifying it as a major stumbling block of access to grey literature.

The study further sought to get suggestions on how the identified problems could be addressed. Table 1 below summarizes the main responses:

Suggestion	**% Respondents Suggesting**
Open special collections from 8am to 9pm	83.2
Open special collections during the weekend	93.7
Provide resources in electronic format	95.3
Make the resources more visible	68.2

Table 1: *Suggestions on improving access to Grey Literature*

4.3 GREY LITERATURE PRODUCERS SURVEY

4.3.1 DEMOGRAPHIC PROFILE

The second survey sought to establish the viewpoint of digital grey literature producers in the context of designing a system for its effective acquisition, organization and distribution. The survey targeted staff and graduate students. A total of 99 respondents participated with 32.3% comprising teaching staff, 23.2% non teaching staff and 44.5% comprised of the graduate students.

The survey sought to find out the length of time they had been at the university. 17.2% of the respondents had been at the university for less than one year, 21.8% had been at the university for duration of one to two years, 35.4% had stayed for between two to three years while 13.6% were at the university for a period exceeding three years.

4.3.2 GREY LITERATURE PRODUCED

Participants were asked to select from a list the category of *digital* grey literature that they had produced in the context of teaching and research. They were further asked to select their preference for greater dissemination for each category of material produced. The table below summarizes the cumulated percentages for the various categories, and the corresponding preference for distribution.

Grey Literature Category	**% of Respondents that have Produced Grey Literature**	**% Corresponding preference for Dissemination**
Conference papers/presentations	43%	40%
Theses and dissertation	52%	47%
Projects	68%	63%
Seminar Paper/Working Paper	13%	12%
Past Exam papers	42%	40%
Revision Kits	16%	12%
Course Notes	47%	32%
Software Code and databases	7%	3%

Table 2: *Grey literature produced and corresponding preference for dissemination.*

The bar graph below puts this information in context:

It is evident that most producers of digital grey literature outputs would prefer to have their works broadly disseminated. There is apparent reservation on dissemination of course notes, with only 76% of producers of course notes preferring broader distribution.

4.3.3 DISSEMINATION PRACTICES

The respondents were asked to select channels that they currently use in dissemination of grey literature. They selected from a list of provided dissemination channels. Email was the most frequently used manner used to disseminate the digital grey literature that they had produced with 88% of the participants having used it. It was followed by the e-learning platform at 47%, institutional and departmental web pages at 45%. Personal web pages and blogs did not seem popular with only 3% of the respondents having used them.

Participants were requested to provide additional ways that they used to disseminate their digital outputs produced. The responses featured online pages such as Slideshare, ResearchGate, the funding project pages and the library catalogue.

4.3.4 CHALLENGES IN MANAGING AND DISSEMINATING GREY LITERATURE

Respondents were asked to enlist problems that they encountered using the tools identified above. The results were categorized and are summarized below in Table 3:

Challenge	**% experienced**
Broken Links	53%
Minimum visibility to research works despite availing it	25%
Access is limited to intended audience e.g. elearning	58%
I don't know the procedures to get my work visible on the website	74%
I do not have a personal website/blog	86%
Most of my work is in my personal computer	92%

Table 3: *Challenges associated with managing and disseminating digital grey literature*

4.3.5 SYSTEM REQUIREMENTS

Respondents were asked to select what they would consider important if a system for dissemination of grey literature was to be provided. 83% of respondents rated long term

preservation of their works as very important and important. A cumulative 72% of the respondents rated long term preservation of their works as very important/important. 78% of the respondents rated ability to accommodate a variety of file formats as very important/important. 87% value ease of use of the system by rating it as very important/important. Ability to post content from anywhere is not greatly valued with a cumulative 58% rating it as very important/important. Global visibility was rated as important/very important by 71% of the respondents. Other desirable features enlisted by participants included ease of retrieval of information and ability to search through the collections, and provision of full text documents.

4.4 CONCLUSION

This chapter has presented and analyzed data that was collected with the view of creating an effective system for the management of digital grey literature.

From the data collected and analyzed in this chapter, the researcher was able to create a system specification and design of an appropriate system. The next chapter details the system implementation process for the institution repository software selected to manage the digital grey literature produced.

CHAPTER FIVE: IMPLEMENTATION

5.1 INTRODUCTION

This chapter details the implementation of the system based on the problem identified in chapter one and analyzed in chapter four. It expands on the systems analysis and design methodology highlighted in chapter three from system selection, implementation and training.

The project's aim was to implement a system. Subsequently, system design methodologies were employed to manage the implementation of project from the initial design idea to the actual physical design. The project utilized the Structured Systems Analysis and Design Method, (SSADM). SSADM was found appropriate as methodology is appropriate because it results in improved quality of systems developed, improved project management, planning and control, more effective use of inexperienced staff. ( )

Using SSADM, the project was divided to six modules that were further broken down into a hierarchy of stages, steps and tasks.

i. Feasibility studies
ii. Requirements analysis
iii. Requirement specification
iv. Logical system specification
v. Implementation
vi. Training and Support

i. Feasibility Study

The role of the study in the planning stage is to ensure that the proposed project is viable economically, technically and operationally. A project is said to be feasible if the overall benefits of the project outweigh the costs.

A survey was conducted to establish the current needs of the users in grey literature and the current needs of grey literature producers. The details of the results of the feasibility study are presented in Chapter four.

a) Time Feasibility

The project has to completed within the required period of time. The duration of the project is one semester – approximately four months. Given this constraint, the scope of the project will be scaled down from a complete institutional repository and concentrate on only grey literature. Populating the repository may also not be completed within this span of time, and the researcher will have to be content with less items in the repository.

b) Skill Feasibility

This was specifically concerned with the ability of the librarian to develop the system. The researcher arrived to the conclusion that the project enjoyed skill feasibility. The researcher is well versed with library management and management of grey literature, as well as use of automated library system. Besides, additional help was got from the Systems Librarian especially for the initial installation and configuration of the system. Additional skills were also elicited from the Library Committee members.

c) Economic Feasibility

This was concerned with the costs that would be incurred in carrying out the project. It also sought to find out whether it would be economically possible to maintain the system.

The hardware and software necessary to carry out the project are already available at the university, and the existing hardware was used. The software is open source and minimal costs will be incurred in downloading. Installation, customization and configuration of the software to the fit the needs of the institution cost considerable staff time.

d) Technological feasibility

This was concerned with whether or not the system to be built could be realized within the available technology. The developed system should be completed successfully, with all its functions and processes working as they should.

The project was technologically feasible as the software to carry out the task already exists. Three common software used in managing repositories include D-Space, Fedora and E-Prints. An analysis was been done and a decision reached to adopt D-Space based on the existing users globally, vibrancy of the user community, skill availability and most importantly ability to effectively manage digital grey literature.

e) Social Feasibility

The system adopted has to be one that will be acceptable by users in the replacement of the manual system. To ensure social feasibility, the new system would have to offer users convenience to use resources from their workstations, saving them much effort, yet offering them with the familiarity of current system that they are used to.

The system would also have to adopt a simple interface that is familiar to users. The codes used to represent communities were customized to reflect the physical departments already familiar to the users.

5.2 USER REQUIREMENTS AND SYSTEM SPECIFICATIONS

5.2.1 USER REQUIREMENTS

From the analysis of data in chapter four the following user specifications were derived for the grey literature producers and users:

The user requirements and specifications for the grey literature producers were identified as:

a) Ease of submitting the material
b) Ease of use
c) Dependability(reliable, long term preservation)
d) Ability to accommodate a wide range of resources and formats
e) Ability to be able to post the content from anywhere

f) Global visibility

5.2.2 SOFTWARE SELECTION CRITERIA

From the user requirements and specifications, an institutional repository system was considered appropriate. The market place offers very few commercial products for repository implementers (such as DigiTool and Digital Commons). However, these proprietary or hosted systems were considered too costly an option for Strathmore University. The university has a policy on use of Free and Open Source Software (FOSS) wherever possible in comparison to proprietary software. A search for viable FOSS repository management software narrowed to three options: Greenstone, Fedora and DSpace as possible platforms for the repository service.

A software comparison was then done and DSpace was selected. The selection was based on software functionality vis-à-vis identified user requirements. Available competency and the fact that DSpace has been widely tried and tested made it also a natural choice.

DSpace is a repository management software that was developed by Massachusetts Institute of Technology (MIT) to manage the digital output of its faculty. As faculty and other researchers develop research materials and scholarly publications in increasingly complex digital formats, there is a need to collect, preserve, index and distribute them: a time-consuming and expensive chore for individual faculty and their departments, labs, and centers to manage themselves. The DSpace system provides a way to manage these research materials and publications in a professionally maintained repository to give them greater visibility and accessibility over time.

5.3 MANAGING DIGITAL GREY LITERATURE USING D-SPACE

Managing digital grey literature from an information perspective involves three core tasks: Acquisition, Organization and Dissemination. This works in the context of provision of related administrative tasks. What follows is an overview of conducting these tasks using D-Space, starting with core concepts and relationships.

DSpace is designed to operate as a centralized, institutional service. Different communities within the institution such as labs, centers, schools or departments can have their own separate areas within the system. Members of these communities deposit content directly via a Web user interface designed to make this depositing as simple as possible. Alternatively the system features a batch item importer for the bulk loading of content. Each community may also appoint people as 'gatekeepers', who may review and edit submissions before their inclusion in the main repository. The DSpace system then indexes the metadata submitted with the digital item and makes it available according to the access privileges determined by the community. In order to provide a workable service in the available time, DSpace was developed 'breadth-first'. In other words, each of the basic requirements of an institutional digital repository system was addressed in a relatively simple manner, so that functionality can evolve with the service already in production.

5.3.1 ACQUIRING DIGITAL GREY LITERATURE USING D-SPACE

The challenges of acquiring grey literature were highlighted in the literature review. One core attribute of grey literature is its temporary nature and limited distribution . The problem was further highlighted by the user requirements desire for a system that provides an easy process for submission of their digital grey literature outputs.

For DSpace to be of value in the management of digital grey literature there is need for the system to provide of mechanism of capturing the grey literature at the point of production by the authors themselves. There is also need for provision for a submission to be made on behalf of another .

Acquiring grey literature to the repository signifies a commitment by the organization to preserve and make it accessible. As such there are legal implications and the creator must grant the repository non-exclusive rights to redistribute their content this way without compromising them of their exclusive bundle of rights.

Acquiring grey literature to the repository is technically simple involving a process of user registration, submission of basic metadata, uploading the accompanying files and signing an end user license. These processes will be explained in details here:

a) Registration of New User

The first process of submission involves authenticating who can submit content. Only users who are registered with DSpace can submit items, administer items, collections or communities or view items that are not globally accessible. The system will ask the user to authenticate themselves.

Figure 5: *DSpace authentication screenshot*

Users that have been 'bulk identified' by an administrator may register with SU-Portal by providing a valid email address to uniquely identify them and a corresponding password. The user registration usually gather the user's email address, verifies that individuals are registering using an email address that they can access. It is not possible to register DSpace using someone else's email address. The system stores the password securely within it as well as provides a mechanism for securely resetting forgotten passwords.

Once a user logs in to their profile, they acquire rights to submit to a collection. DSpace counters the problem of grey literature acquisition by providing an easy way for materials to enter the system through a distributed system accessible via a web interface. This subsequently means that a user can submit an item from any global location.

b) Submitting Items

Submitting an item to the repository is a simple process of identifying a collection to submit to and clicking the submit button.

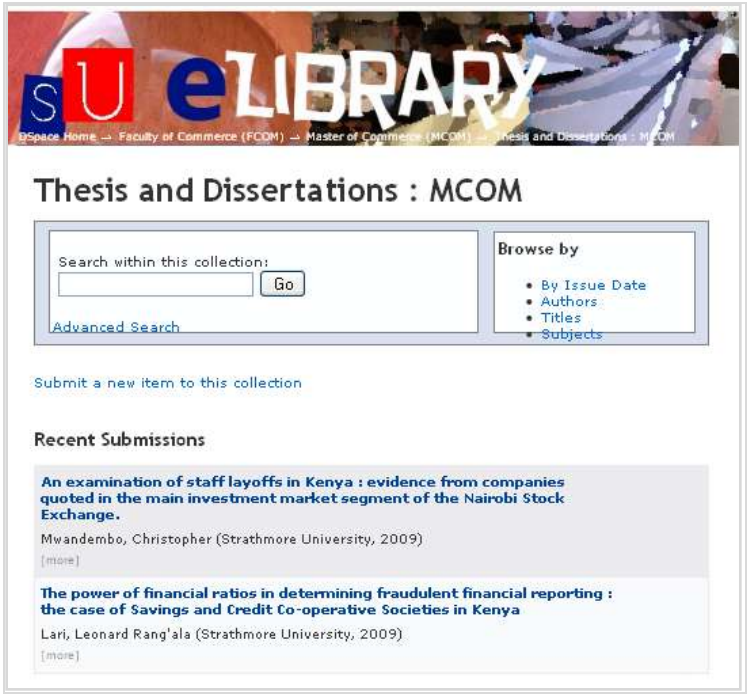


Figure 6: *DSpace Submission Screenshot*

The user is then required to provide *metadata* about the item they are submitting. Usually, the mandatory fields are author, title and date. The user then attaches the associated bitstreams (the file) and signs an agreement with the repository granting it non-exclusive distribution rights. Depending on the workflow submission set, the item could then become part of DSpace and accessible, or could move forward through the workflow for review.

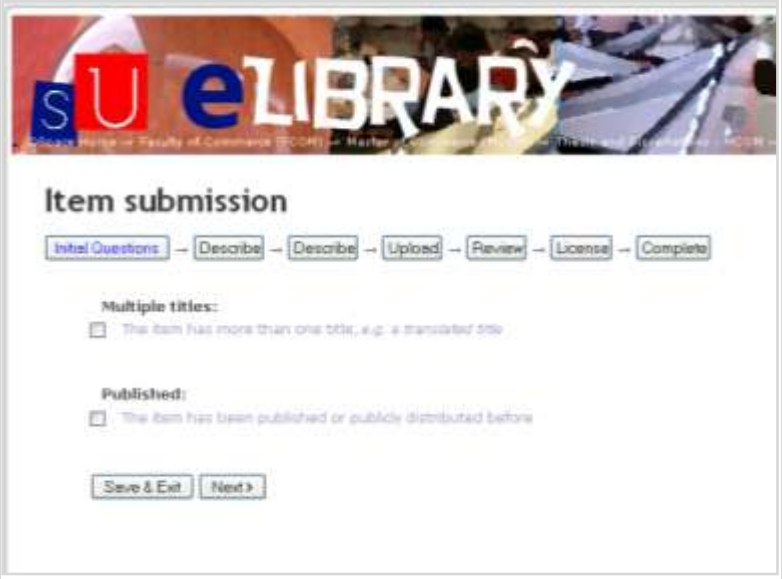


Figure 7: *DSpace Submission Process Screenshot*

c) Metadata

During submission, users are presented with a fill-in form that allows them to fill in the metadata. The same form is filled for all submitted items. The metadata is fundamental as it enables discovery of resources upon searching the repository. Good metadata allows easy discovery, synonymous to traditional cataloguing in the library systems.

The Dublin Core Metadata Element Set Schema is used to guide on metadata. The elements captured from the end-user interface are:

i. **Author(s) :** DSpace supports single and multiple authors. Currently no authority control for authors exist. (i.e. DSpace does not currently know that "Wahome Mutahi" and "Whispers" are the same author, nor does it distinguish well between two authors that share the same name).

ii. **Title(s):** A title field is provided. An extra alternative title field is also provided for use if applicable.

iii. **Date of Issue:** Date when the material was made publicly available. Date of publication.

iv. **Series Name and Report Number:** The field is applicable for working papers and reports that may have local series names and numbers. This is particularly suitable for grey literature.

v. **Identifiers:** Including ISBN, ISSN, ISMN, URI, DOI and others. The field is not mandatory.

vi. **Language :** The language in which the material submitted is written in.

vii. **Subject Keywords.** Multiple keywords are supported. Currently no thesauri or authority control for subject keywords exist.

viii. **Abstract :** A summary of the resource being submitted.

ix. **Sponsors / Funding Codes :** If the research received funding and sponsorship, the fields can be used to provide the name and/or codes.

x. **Other Description :** Other descriptions that may be appended to the resource.

Additional elements are modeled and stored internally. Some of these are generated automatically by the system. Others may be managed by DSpace administrators using the administrator user interface.

d) Upload File(s) to Item

Submitted DSpace Items are destined to become "archival atoms", that is some amount or boundary of material that make sense together. As such, each item can include multiple pieces of content. Each piece of content might comprise several files. Users must submit one or more files to be included with their item. For example, a user might submit a conference paper, along with presentation materials actually used at the conference. A researcher might submit a pre-print of an article, along with dataset(s) that would enable other researchers to independently reproduce the research results. DSpace calculates and retains a checksum of each file uploaded with the item that can be used by DSpace administrators and by users to verify the integrity of the content and metadata within the system. This checksum can be obtained through the end-user interface.

DSpace attempts to recognize the bitstream format of each uploaded file. If it cannot match the bitstream format to a known format from the system's bitstream format registry, it asks the user to describe the format so that library administrators can track and support important emerging formats over time.

e) Grant Distribution License

To enable the host institution to administer, preserve, and distribute the submitted material, DSpace asks the user to grant to the institution a non-exclusive license to distribute the material, and to translate it for the purposes of preservation. License agreements can vary by Collection, and are specified by the collection administrator(s) for the Collection. Because license terms are likely to change over time based on the needs of submitters and the host institution, DSpace stores a copy of the license that was granted at the time of submittal as a bitstream within the item, so that the specific terms agreed upon are always available.

The distribution license is a legal document that is always stored with a copy of the item submitted. This way, specific terms agreed upon are always available. A default license can also be used in absence of a collection specific license. Appendix 4 provides the Default Distribution License for the SU-Portal instance.

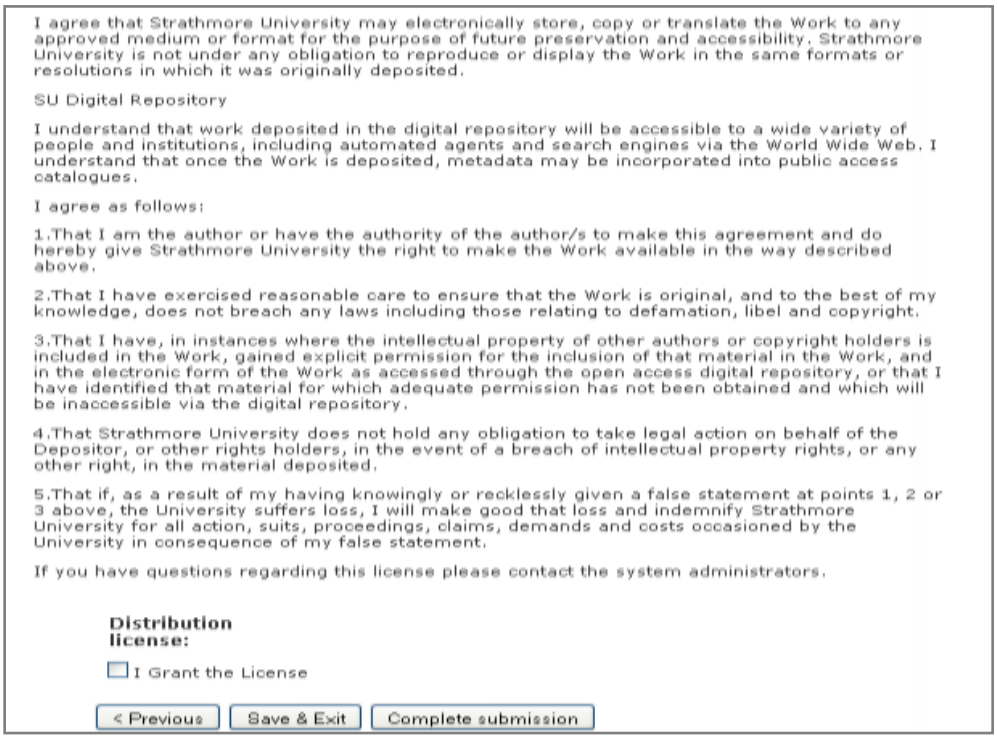


Figure 8: *DSpace Grant License Screenshot*

f) Submission Workflow

When the user submits an item to a DSpace collection, the system routes the submission through the approval process previously configured for that collection. The approval process can vary by collection, and can include any combination of the approval roles defined below. Further, groups of users can be used to associate any number of users with each of the roles that are configured for the process.

The system routes the submission to individuals (if any) who have been chosen to assume each of the following roles for the target collection. If none have been selected, the item becomes part of DSpace after submission.

a) Reviewers: review the content of the submission for appropriateness to the collection. Reviewers act as the gatekeepers for the collection. Reviewers are empowered to return a submission back to the submitter because it is deemed inappropriate for the collection. Reviewers do not edit submission metadata. The y simply accept or reject a submission.

b) ***Approvers:*** check the submission for completeness and/or obvious errors (e.g. wrong file uploaded). Approvers can edit the submission's metadata to fix obvious errors, and are empowered to return a submission back to the submitter because it is incomplete or in error with appropriate comments.

c). ***Metadata Editors****:* check and/or augment the submission's metadata. For example, a metadata editor may be assigned to add the appropriate series name and number to each submission in a collection. Metadata Editors can only edit the submission's metadata.

These roles are summarized in the table below:

Reviewer	**Approver**	**Metadata Editor**
Can accept submission for inclusion, or reject submission. .	Can edit metadata provided by the user with the submission, but cannot change the submitted files. Can accept submission for inclusion, or reject submission	Can edit metadata provided by the user with the submission, but cannot change the submitted files. Must then commit to archive; may not reject submission.

Table 4: *Submission Workflow Roles and privileges*

Each collection may specify zero or more individuals to assume each of these roles. If none is selected, the system skips the roles and the item becomes available once submitted. The choice for the process is determined by the nature of the items being submitted, the person submitting and organizational policies. For instance, in submitting past examination papers, there was no need for reviewers and approvers of the item. However, when a student submits a thesis, there is need for reviewers (who could be faculty administrators to just check for appropriateness) and approvers (the student supervisors) to check for submission completeness and obvious errors) and metadata editors (librarians, to complete the metadata for the item) before the item becomes globally accessible. The intent of the various checks and controls is to allow communities flexibility in meeting their collection management needs while avoiding the institutional paralysis that often results from too much flexibility or too many choices.

DSpace users in each role are notified via email when they have pending submissions that require their attention. These tasks are also available to them from the "My DSpace" section of the DSpace system. As each task is completed, the submission is routed to the individual(s) responsible for the next stage of review, if any. The system sends email(s) to each of these Users, with links to their pending task in the DSpace system. When the submission either enters the DSpace archive or is returned to the submitter, the system notifies the submitter by email about the status of their submission.

The workflow steps are designed such that if a collection has no e-person group associated with the group, the particular task is skipped and the workflow continues. In a case where there are no reviewers, approvers and metadata editors, then the collection is installed directly to the main archive.

g) My DSpace

DSpace offers each user personalized access to information within the system through their My DSpace page. As appropriate given their role(s) in the system, users can view their:

- Items being assembled : incomplete submissions or submissions that were interrupted while partially complete.

- submissions pending archive :
- archived Items that they submitted
- review tasks for Items pending archive: users who are taking part in the approval process for one or more collections can view, select and perform tasks that require their attention.

DSpace filters these sections so that only the sections relevant to each user are presented. When a task is taken, the task is automatically removed from the pool. This avoids a situation where more than one person are dealing with a task.

5.3.2 ORGANIZING DIGITAL GREY LITERATURE USING D-SPACE

a) DSpace Organizational Model

Each DSpace site is divided into *communities*; these typically correspond to a laboratory, research center or department. Communities contain *collections*, which is a grouping of related content. Each collection is composed of *items*, which are the basic archival elements of the archive. Items are further subdivided into *bundles* of *bitstreams*. Bitstreams are, as the name suggests, streams of bits, usually ordinary computer files. Bitstreams that are somehow closely related are organized into bundles, for example HTML files and images that compose a single HTML document.

The way data is organized in DSpace is intended to reflect the structure of the organization using the DSpace system. Typically, a DSpace site is divided to communities that reflect the functional categorization in the organizational chart.

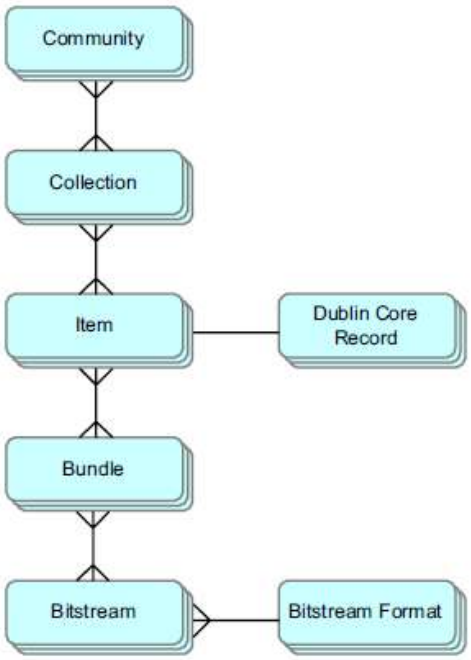


Figure 9: *DSpace Data Model Diagram*

In the case of the SU-Portal, communities were set up to reflect the existing Faculties, Schools and Institutes. Sub-communities reflected the various courses offered by the given Schools and collections echoed the nature of the materials for depositing. This is illustrated below, and a select structure of the communities and collections is reflected in the Appendix 3.

A data item is composed of the metadata and the accompanying full-text. This allows for discovery of the resource. An example from the SU-Portal is depicted below:

Object Type	Example in Su-Portal
Community	Faculty of Information Technology (FIT)
Sub-Community (more sub-communities may be created as desired)	Master of Science in Information Technology (MSIT)
Collection	Theses and Dissertations MSIT: 2008
Item	A thesis from this community. Mobile phone voting using blind signature protocol Olembo, Maureen (Metadata and accompanying files)
Bitstream	PDF Full text file of the thesis
Bitstream Format	Portable Document Format (PDF)

Table 5: *Example Data Model from Su-Portal*

b) Administering Collections

Administering collections in DSpace involves a number of tasks as discussed below:

i) Adding or Removing authorized Submitters

A collection administrator usually has the powers to select epeople who may submit items to a given collection. In our instance, for the collection *Master of Commerce Thesis* the authorized submitters were only current students who are undertaking the unit. The authorized submitters may be added as a group. Removal of epeople equally follows the simple process of clicking their email address and selecting remove.

ii) Customizing Collection Home Page

Using basic HTML the collection home page can be customized. News item can be added and additional items on the community can be offered. Additional links to relevant pages can also be provided from the community home page. This task can be assigned to the community administrator.

c) Administering Items

Once an item has been submitted (as described in the previous section on item submission) it is still possible to perform additional activities to the items. This includes editing the item metadata, withdrawing the item or mapping the item to other collections.

Editing an item metadata may be desirable so as to add additional fields or correct spelling errors. The edit metadata fields however do not offer a graphical user interface as was provided in the submit procedure. No validation is also done, so usually one has to be careful in performing these tasks. Additional bitstreams can also be added. This feature was particularly useful in administering past examination papers since the metadata remained the same, and only the bitstreams change to reflect the immediate past examination paper. This made the update process fairly simple involving an item search and an edit to its bitstreams. Withdrawing an item removes the item temporarily from all public view (searching and browsing). However it does not remove the item from the database. Reasons for withdrawal could include legal breaches or institutional policy procedures. This allows that the item could be reinstated in future.

Items can appear in more than one collection. In our DSpace instance, the School of Graduate Studies wanted all thesis to appear in their collection page. But the individual faculty also wanted the item to appear in their collections. The same was witnessed in shared courses that are offered by School of Humanities and Social Sciences. The items had to appear both at this school and at the various departments offering the unit. (e.g. Faculty of Information Technology).

The Item Mapping tool which allows the administrator to search and map collections to other collections was used. The item mapped from a collection usually inherits its parent rights. E.g. if it was limited to a particular group of viewers, mapping it to a new collection does not alter or modify those rights.

Preservation

Each bitstream is associated with one Bitstream Format. Because preservation services are an important aspect of the DSpace service, it is important to capture the specific formats of files that users submit. In DSpace, a bitstream format is a unique and consistent way to refer to a particular file format. An integral part of a bitstream format is an either implicit or explicit notion of how material in that format can be interpreted. For example, the interpretation for bitstreams encoded in the JPEG standard for still image compression is defined explicitly in using standards. The interpretation of bitstreams in Microsoft Word 2000 format is defined implicitly, through reference to the Microsoft Word 2000 application. Bitstream formats can be more specific than MIME types or file suffixes. For example, application/ms-word and .doc span multiple versions of the Microsoft Word application, each of which produces bitstreams with presumably different semantics. Supported The format is recognized, and the hosting institution is confident it can make bitstreams of this format useable in the future, using whatever combination of techniques (such as migration, emulation, etc.) is appropriate given the context of need.

.

Metadata

Broadly speaking, DSpace holds three sorts of metadata about archived content: Descriptive, administrative and structural metadata.

Capturing metadata in DSpace is straightforward through forms provided that provide basic Dublin Core metadata structural elements. The screenshot on the next page illustrates.

The following fifteen Dublin Core metadata elements are captured:

i. **Title** : The name given to the resource by the creator or publisher.

ii. **Creator:** The person(s) or organization(s) primarily responsible for the intellectual content of the resource; the author.

iii. **Subject:** The topic of the resource; also keywords, phrases or classification descriptors that describe the subject or content of the resource.

iv. **Description:** A textual description of the content of the resource, including abstracts in the case of document-like objects; also may be a content description in the case of visual resources.

v. **Publisher:** The entity responsible for making the resource available in its present form, such as a publisher, university department or corporate entity.

vi. **Contributors:** Person(s) or organization(s) in addition to those specified in the CREATOR element, who have made significant intellectual contributions to the resource but on a secondary basis.

vii. **Date:** The date the resource was made available in its present form.

viii. **Type:** The resource type, such as home page, novel, poem, working paper, technical report, essay or dictionary. It is expected that TYPE will be chosen from an enumerated list of types.

ix. **Format:** The data representation of the resource, such as text/html, ASCII, Postscript file, executable application or JPG image. FORMAT will be assigned from enumerated lists such as registered Internet Media Types (MIME types). MIME types are defined according to the RFC2046 standard.

x. **Identifier:** A string or number used to uniquely identify the resource. Examples from networked resources include URLs and URNs (when implemented).

xi. **Source:** The work, either print or electronic, from which the resource is delivered (if applicable).

xii. **Language:** The language(s) of the intellectual content of the resource.

xiii. **Relation : T**he relationship to other resources. Formal specification of Relation is currently under development.

xiv. **Coverage:** The spatial locations and temporal duration characteristics of the resource. Formal specification of COVERAGE is also now being developed.

xv. **Rights Management:** A link (URL or other suitable URI as appropriate) to a copyright notice, a rights-management statement or perhaps a server that would provide such information in a dynamic way.

5.3.3 DISSEMINATING DIGITAL GREY LITERATURE USING D-SPACE

a) User Interface

DSpace's current user interface is web-based. There are several interfaces: one for submitters and others involved in the submission process, one for end-users looking for information, and one for system administrators.

The end-user or public interface supports search and retrieval of items by browsing or searching the metadata. Once an item is located in the system, retrieval is accomplished by clicking a link that causes the archived material to be downloaded to the user's web browser. "Web-native" formats (those which will display directly in a web browser or with a plug-in) can be viewed immediately; others must be saved to the user's local computer and viewed with a separate program that can interpret the file (e.g. a Microsoft Excel spreadsheet).

b) Downloading Item Bitstreams

As mentioned earlier, an item in DSpace is composed of its description (metadata) and accompanying bitstreams. The item display after searching or browsing will display first the items metadata and on the bottom, associated bitstreams. The user can then download or open it in a browser if its corresponding mime-type is a digital native. (e.g. PDF, JPG)

c) Search and Browse

DSpace allows end-users to discover content in a number of ways, including:

- **A.** Searching for one or more keywords
- **B.** Browsing though title, date and author indices
- **C.** Via external reference, such as a **Handle**

A. Searching

Search is an essential component of discovery in DSpace. Users' expectations from Web search engines are quite high, so a goal for DSpace is to supply as many search features as possible. DSpace employs full text indexing of archived items so in searching, one is searching on the entire corpus.

It is also possible to limit the search to a particular community or collection by navigating to that collection and then clicking on the search button. The screenshot below shows the DSpace search feature from the home page of the repository.

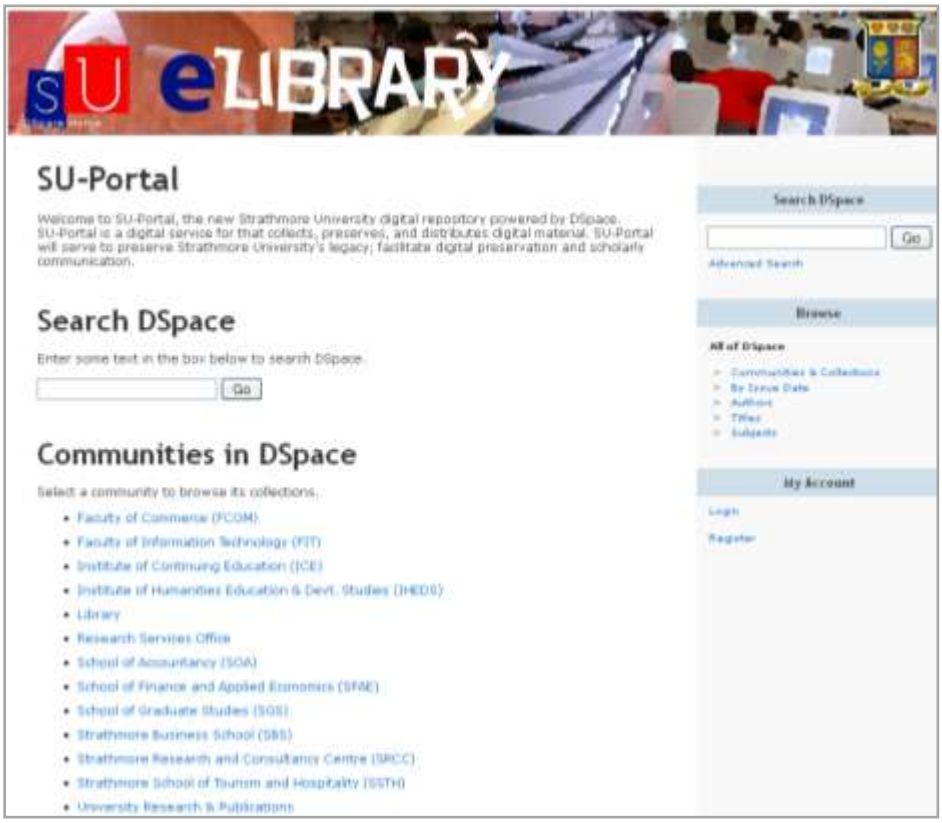


Figure 10: *DSpace Search Interface Screenshot*

DSpace employs the *Jakarta Lucene* search engine. The following search features are therefore provided:

i) Indexing: DSpace provides full-text indexing feature. Where this is selected, all fields and the full text attachments are searched. As such the title, author, subject, abstract, series, sponsor and identifier fields will be searched in addition to the items full text attachment.

ii) Stop Words : The search engine ignores certain words that occur frequently in English, but do not add value to the search. These are: "a", "and" , "are" , "as" , "at" , "be" , "but" , "by" , "for" , "if" , "in" , "into", "is" ,"it" ,"no" , "not" , "of" , "on" , "or" , "such", "the" , "to" , "was". The stop word list is not editable to the administrator.

***iii) Truncation*:** Use of an asterisk (*) after a word stem introduces this wild card operator that searches word stems. e.g. financ* will retrieve finance, financers, financing, and financial.

***iv) Stemming* :** The Lucene search engine automatically expands words with common endings to include plurals, past tenses etc.

***v) Phrase Searching*:** Phrase searching is supported and is introduced by putting quotation marks around the phrase e.g. "information technology"

***vi) Exact word match*:** A + sign before a word signals that the word must appear in the search results. e.g. Information +literacy will retrieve information as optional but literacy must appear in the search results.

***vii) Boolean searching* :** The Boolean operators (AND, OR, NOT) are supported and can be used to combine terms. They must be capitalized. The AND operator finds items containing all words or phrases. The OR operator retrieves items that contain either search term while the NOT operator is used to eliminate a search term. A minus sign (-) can be used in place of NOT.

***viii) Advanced Searches*:** The advanced search page allows a user to conduct field searches as well combine Boolean operators. A user can restrict their searches to a particular field by clicking on the drop down menu and selecting the field as well choosing their Boolean operator. However, the input boxes must be used in the order provided. Leaving a blank middle box will lead to invalid results. This simple interface is depicted below:

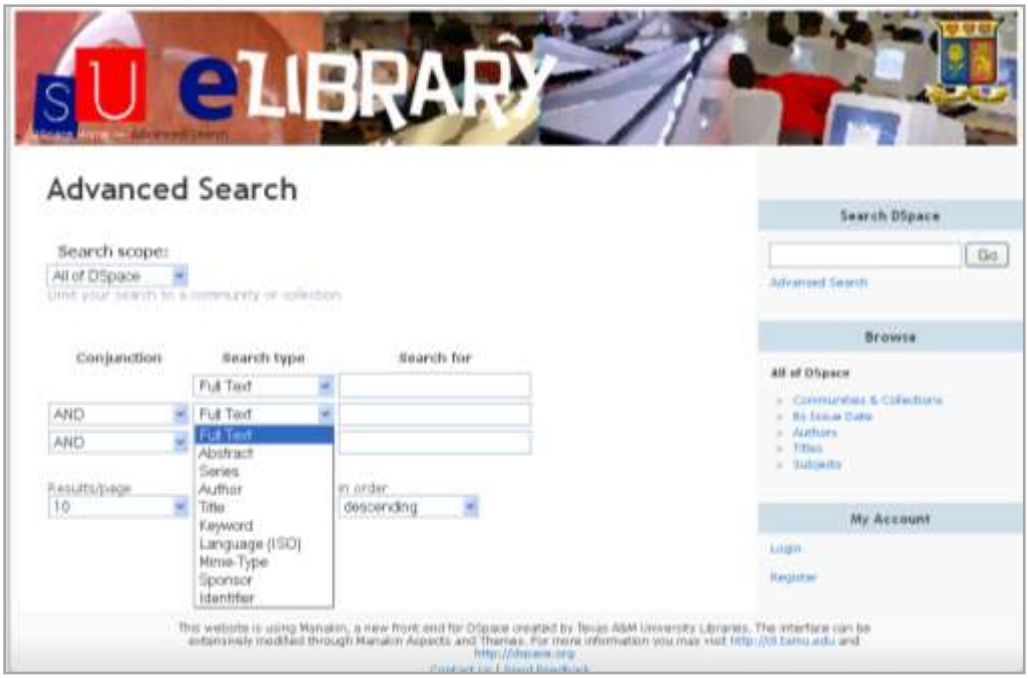


Figure 11: *DSpace Advanced Search Interface Screenshot*

B) Browse

Another important mechanism for discovery in DSpace is the browse. This is the process whereby the user views a particular index, such as the title index, and navigates around it in search of interesting items.

DSpace allows users to browse by Issue Date, Author, Title and Subject. DSpace users may browse the contents of DSpace in a number of ways. They could browse by communities and collections through the structure in an outline view. The outline entry includes text describing the contents and purpose of a given community or collection.

Users can brose a collection page by titles, subjects, author or issue date. This supports results that paginated with links to the previous and next pages. Browse by author produces a list that allows the user to navigate by the alphabet letters. In browsing by author, all items by a single author are collapsed and displayed in a single location. Where such a list is long, it is possible to further sort the list by title or date.

C. Via external reference, such as a Handle

The third way of discovering items in DSpace is through external references such as handles. The rationale for the handle system is that researchers require a stable point of reference for their works. Broken links are a common problem in grey literature management as highlighted in the literature review. To address this, DSpace employs a feature that allows creation of persistent a identifier for each item, collection or community. DSpace employs the CNRI ( ) handle system for creating those identifiers. Handles are primarily a way of assigning globally unique identifies to objects.

Each site running DSpace needs to identify a unique handle prefix from CNRI. The handle system features a global resolution infrastructure. An end user can therefore search for the handle in any service and they will be directed to the object.

Handles can be written in two forms:

hdl: 2262/56169

http://hdl.handle.net/2262/56169

Writing it the second way allows it to be resolved by any web browser. The SUPortal has not yet registered for use of the CNRI handles and uses local identifiers instead.

Below is a capture of an item in DSpace, showing the item handle.

d) OAI Support

The Open Archives Initiative has developed a Protocol for Metadata Harvesting (OAI-PMH). This allows sites to programmatically retrieve or 'harvest' the metadata from several sources, and offer services using that metadata, such as indexing or linking services. Such a service allows users to access information from a large number of sites that are collated in a central catalog. DSpace exposes the Dublin Core metadata for items that are publicly accessible.

To take advantage of this, SU-Portal was registered in OpenDOAR (Directory of Open Access Repositories) and ROAR (Registry of Open Access Repositories). DSpace's OAI-PMH service exposes the metadata in the repository and allows users to search the repository from its interface, hence grating further global visibility. Below is the screen capture of OpenDOAR and ROAR depicting the SU-Portal and visibility of its digital outputs.

Figure 12: *The Open Directory of Open Access Repositories listing Strathmore's Repository.*

e) Selective Dissemination of Grey Literature: Subscription and RSS Feeds

As noted above, end-users (e-people) may 'subscribe' to collections through the Web user interface in order to be alerted when new items are added to those collections. Each day, end-users who are subscribed to one or more collections will receive an e-mail giving brief details of all new items that appeared in any of those collections the previous day. If no new items appeared in any of the subscribed collections, no e-mail is sent. Users can unsubscribe themselves at any time.

DSpace also comes with inbuilt support for Really Simple Syndication (RSS). This allows users to subscribe to feeds of their interest and they are automatically notified based on the way they have selected to receive the feeds whenever a new item is posted to a given collection.

5.3.4 ADMINISTERING THE REPOSITORY

a) Administering E-people

Many of DSpace's features such as document discovery and retrieval can be used anonymously, but users must be authenticated to perform functions such as submission, email notification ('subscriptions') or administration. Users are also grouped for easier administration. DSpace calls users *e-people*, since some users may be machines rather than actual people.

For each e-person, DSpace holds the email address, first and last names, authentication information, and a list of collections subscribed to. DSpace also stores rights associated with the e-person for each item category in the repository. This could be Read (searching and viewing metadata), Write (modifying metadata associated with an object, but not ability to delete), Add (making additional items available) Remove (deleting or excluding an item from the repository) and Workflow (how the user participates in the repository. e.g. reviewer, metadata editor).

b) Institutional Repository Policy

DSpace is a software, a tool that allows collecting, describing, and making accessible digital resources. There was need for a policy to be put in place to define how exactly the tool was to be used, by whom, what would be acceptable for submission and for how long the materials would be preserved.

Policy decisions are complex and the Library Executive Committee was tasked with the policy drafting task. The policies are outlined below:

i) Acquisition Policy: What would the repository accept? By whom
ii) Submission Policy: How do we go about recruiting content to the repository?
iii) Metadata Policy: Which metadata is to be included with items.
iv) Preservation and Withdrawal Policy: For how long should a document be retained? Under what circumstances can an item be withdrawn from the collection?

c) Training, Marketing and Outreach

Marketing of the repository was undertaken to both the producers of digital grey literature and to the end-users. The bottom-up and top-bottom approaches were both for maximum outreach.

For the content creators, the repository was marketed to the top level management and live sessions were held to them. This was largely successful and influenced the passing of the organization-wide open access policy. Follow up sessions were held in large academic staff meetings where demonstrations were conducted on usage and submission procedures.
One-on-one sessions were further conducted to faculty on demand and informal request sessions.

For the end-users, training was conducted as part of the mandatory library training sessions. Users were shown how to access the institutional repository both while in campus and while out of campus. To further discourage the use of print resources, some category of the print grey literature were completely withdrawn (e.g. the past examination). This compelled the users to use and get used to the digital versions as well as drive traffic to the institutional repository. Print and electronic guides were further designed and distributed for training sessions.

CHAPTER SIX
CONCLUSIONS AND RECOMMENDATIONS

5.1 CONCLUSIONS

D-Space software has been successfully deployed in the management of digital grey literature at Strathmore University. The versatile nature of the software has allowed great customizations to be undertaken.

A student user survey conducted in July 2011 as part of the larger annual library survey revealed a user satisfaction rate of 88% with the institutional repository. Further qualitative data indicated that the accessibility and access of resources out of campus of much value particularly to the evening students.

The growth of the institutional repository is perpetual and no doubt content recruitment will be on-going. The more the repository is populated with relevant content, the higher the value will be placed to the repository.

The success of the Su-Portal, the Strathmore University DSpace instance has built in librarians a strong interest in digital library discussions. It has also opened up exciting new roles for librarians and given them confidence in their ability to contribute to the global rally for open access to knowledge and research. Digital grey literature has contributed to greater user satisfaction and opened up special collections to groups that were hitherto disadvantaged by the library's opening hours and loaning policies.

5.2 RECOMMENDATIONS

Based on the SU-Portal's experience in managing digital grey literature using a repository software, the following recommendations would be made:

a) Need for Authority Control

There is no name authority control. It is therefore difficult to differentiate between authors with the same name. e.g. George M. Njenga and George Njen'ga would currently be treated as different authors.

b) Subject controlled vocabulary

The current implementation of DSpace does not have support for controlled vocabulary. It is therefore difficult to discover relationships between objects and hence looking for objects with the same subjects terms remains a difficult task.

c) Web 2.0 tools

The broader web and its users are increasingly demanding for a participatory platform. Users may want to comment, 'like', 'tweet', add to favorites and mash-up the contents found in repositories in the spirit of open access. The only attempt D-Space has made to this end is the inclusion of the RSS feeds. For the repository to really encourage collaboration and feedback for authors, it is necessary that additional Web 2.0 tools be incorporated to future D-Space releases.

d) Personalization of DSpace

While DSpace aims to showcase the works of a creator, there is need for the creator of this works to be seen as a person, and not just the face of the university. These may motivate interaction and participation even in content creation. Individualized statistics for an item should also be provided by DSpace, so that it is not just statistics of an item, but of a person and their collection. This may need DSpace to rethink its current hierarchical format created by communities and collections and perhaps prefer a less decentralized flexible presentation.

e) Faculty DSpace Adoption

There is need for further study to be conducted on factors motivating use/disuse of the current implementation of the Su-Portal instance. A majority of faculty members still have their publications in their possession, and some are unwilling to publish their works with the repository despite numerous efforts in publicize.

REFERENCES

Anbu, J.P. (2002). Institutional Repositories: Time for African universities to consolidate the digital divide. Available online . Accessed 15/01/2010

Artus, Helmut M. Old WWWine in New Bottles? Developments in electronic information and communication: structural change and functional inertia. GL5 Conference Proceedings. *5th International Conference On Grey Literature*. Amsterdam 4-5 Dec. 2003/ GreyNet. – Amsterdam: TextRelease, Jan. 2004.

Boukacem-Zeghmouri, C. & Schöpfel, J. (2005). Access and document supply: a comparative study of grey literature. GL 7 Conference. *7th International Conference on Grey Literature*. Nancy. France. 5-6th Dec.

Farace, D.J. *et al.* (2005). Access to Grey Content: An Analysis of Grey Literature based on Citation and Survey Data, A Follow-up Study. . GL 7th Conference. *7th International Conference on Grey Literature*. Nancy. France. 5-6th Dec.

Jeffery, K.G. and Asserson, A. (2005) Grey in the R&D Process. . GL 7 Conference. *7th International Conference on Grey Literature*. Nancy. France. 5-6th Dec.

Jennings, D. (1997) 'Researching and writing strategic management cases: a systems view". *Journal of Management Decision.* 35(2). 100-105

Jeffery, K.G. and Asserson, A. (2007) Greyscape. GL9, *Ninth International Conference on Grey Literature*. Antwerp, Belgium. 10-11 Dec 2007

Johnson, R.K. (2002). Institutional repositories: partnering with faculty to enhance scholarly communication. *DLib Magazine*. 8:11

Julia Gelfand, J. (2003). The Ubiquity of Grey Literature in a Connected Content Context. Paper Presented at the *International Conference on Grey Literature 5*. Amsterdam, Netherlands

Lambert, S., Matthews, B and Jones, C. (2005). Grey literature, institutional repositories and the organizational context. GL7 Conference. *7th International Conference on Grey Literature.* Nancy. France. 5-6th Dec.

Mackenzie Owen, J. (1997) The expanding horizon of Grey Literature. In Proceedings GL3 : *Third International Conference on Grey Literature*, Luxembourg (Luxembourg).

Moahi, K.H. (2009). Institutional repositories : towards harnessing knowledge for African development. *1st International Conference on African Digital Libraries and Archives* (ICADLA-I). Addis Ababa, Ethiopia. 1-3 July.

Nahotko, M. (2007). Some Types of Grey Literature: A Polish Context. *GL 7 Conference. 7th International Conference on Grey Literature*. Nancy. France. 5-6th Dec.

Pavlov, L.P. (2004). The Problems of Grey in the Context of Postmodern Science and Culture. GL6 Conference. *6 International Conference on Grey Literature*. New York, United States, 6-7 December 2004

Rabina, D. L. (2007). A Scholarly communication perspective of Grey literature in library and information science education. GL9 Conference. *9th International Conference on Grey Literature* Antwerp, Belgium, 10-11 December 2007

Ranger, S.L. (2004). Grey Literature in Special Libraries: Access and Use GL6 Conference. *6th International Conference on Grey Literature. New York, United States, 6-7 December 2004*

Sen, B. (2008). Grey literature for development: Some case studies. - In: *GL9 Conference Proceedings*. TextRelease, 4 p.,

Siegel, G. (2003). Capturing Academic Grey Literature – Starting at Home. *GL5 Conference Proceedings* Amsterdam, Netherlands, 4-5 December 2003

Starovoitov, A. *et .al.* (2007) Digital Documents in Grey Literature: New Challenges. GL9 Conference. *9 International. Conference on Grey Literature* Antwerp, Belgium, 10-11 December 2007

Stock, C. (2007). Open access to full text and ETDs in Europe**:** improving accessibility through the choice of language? GL9 Conference. *9 International Conf. on Grey Literature.* Antwerp, Belgium, 10-11 December 2007

Sulouff, P., Bell, S., Briden, J., Frontz, S., Marshall, A. (2005). Learning about grey literature by interviewing subject librarians: A study at the University of Rochester. http://www.ala.org/ala/acrl/acrlpubs/crlnews/backissues2005/julyaugust05/learngreylit.cfm

Weintraub, Irwin. The role of grey literature in the sciences. *ACCESS*. 10. Not paginated. Fall 2000b. http://academic.brooklyn.cuny.edu/library/access/greyliter.htm

APPENDICES

APPENDIX 1: SURVEY QUESTIONNAIRE FOR GREY LITERATURE USERS

User Survey

The library is currently in the process of redesigning its Special Collections to help improve access and use to its collections. Kindly fill in this short survey. Your correct and honest answer will help the library design a a useful system. Thank you for taking part in this survey.
* Required

What is your affiliation with the University?

- ○ Teaching Staff
- ○ Non-teaching Staff
- ○ Graduate Student
- ○ Degree Student
- ○ Diploma Student
- ○ Professional Student (CPA, ACCA, CISA, CISM)

What is your current year of study? (Students) or How long have you been employed by the University? (Staff)

- ○ 1(or less)
- ○ 2 Years
- ○ 3 Years
- ○ 4 Years
- ○ Greater than 4 years

Which of the following information resources have you ever used in your research and/or study work? (tick all that apply)

- ☐ Conference papers and/or presentations
- ☐ Theses and dissertations
- ☐ Past Exam papers
- ☐ Revision Kits
- ☐ Course Notes
- ☐ Software Codes
- ☐ Students Projects
- ☐ Datasets

On a scale of 1 to 5 (5 very important, 1 Not important) how would you rate the importance of the resources listed above for your study and/or research? *

	5 Very Important	4 Important	3 Quite important	2 Not important	1 Not applicable
Conference papers and/or presentations	○	○	○	○	○
Theses and Dissertations	○	○	○	○	○
Past Exam papers	○	○	○	○	○

	5 Very Important	4 Important	3 Quite important	2 Not important	1 Not applicable
Software codes	○	○	○	○	○
Students Projects	○	○	○	○	○
Data sets	○	○	○	○	○

When using the resources in question **, where did you primarily get them from? (tick all that apply) *

- ☐ Library (special collections)
- ☐ Internet
- ☐ Colleagues
- ☐ Other:

Rate the following statements in relation to retrieval of the sources above *

	Strongly Agree	Agree	Disagree	Strongly Disagree	Not applicable
It is difficult to trace	○	○	○	○	○
Special collections opening hours	○	○	○	○	○
I prefer an electronic version of the work	○	○	○	○	○
Evening Students	○	○	○	○	○
Number of copies	○	○	○	○	○
crowding	○	○	○	○	○
tracing colleagues works	○	○	○	○	○
prefer to built on what others have done	○	○	○	○	○
a central place for placing these resources	○	○	○	○	○

Submit

APPENDIX 2 : GREY LITERATURE PRODUCERS SURVEY

Survey for Grey Literature Producers

The library is currently in the process of redesigning its Special Collections to help improve access and use to its collections. Kindly fill in this short survey. Your correct and honest answer will help the library design a a useful system. Thank you for taking part in this survey

What is your affiliation with the university?

- ○ Academic Staff
- ○ Administrative Staff
- ○ Post-graduate Student

In the course of your teaching and learning, which of the following have you produced in digital/electronic form? Tick all that apply.

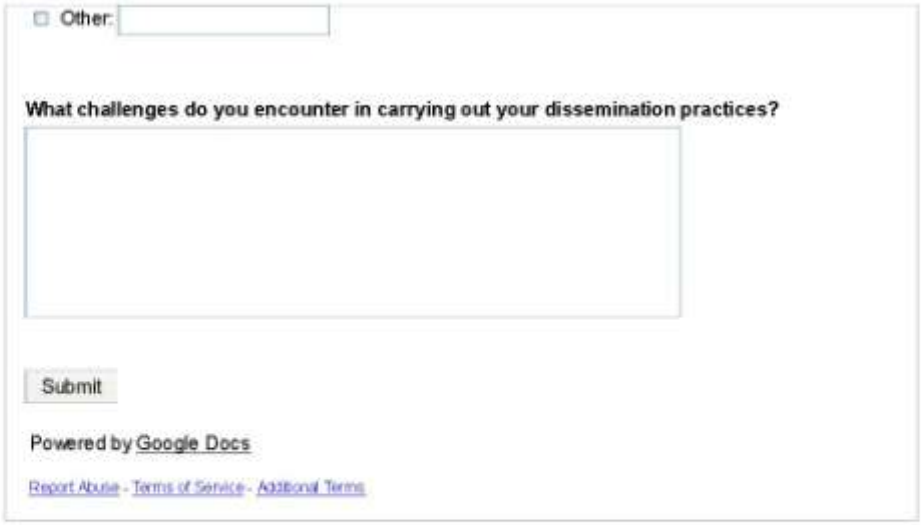
- ☐ Conference papers and/or presentation
- ☐ Theses and dissertation
- ☐ Projects
- ☐ Seminar Paper/Working Paper
- ☐ Past Exam papers
- ☐ Revision Kits
- ☐ Course Notes
- ☐ Software Code and databases

Of the choices provided abovet that you have produced, what would have been your dissemination preference?

	I prefer it to be disseminated widely	I do not prefer it to be disseminated widely
Conference papers and/or Presentations	○	○
Thesis and/or dissertation	○	○
Project	○	○
Seminar paper/working paper	○	○
Past examination paper	○	○
Revision Kits	○	○
Course notes	○	○
Software code and databases	○	○

How do you currently disseminate the outputs selected above (choose all that apply)

- ☐ Strathmore Website / departmental web site
- ☐ e-Learning platform
- ☐ Personal web page / blog
- ☐ Email
- ☐ Internet sites (Slideshare, Research gate)


Other:
What challenges do you encounter in carrying out your dissemination practices?
Submit
Powered by Google Docs
Report Abuse - Terms of Service - Additional Terms

APPENDIX 3: LIST OF SELECTED COMMUNITIES AND SUB-COMMUNITIES IN SU-PORTAL

- Faculty of Commerce (FCOM)
    - Bachelor of Commerce (BCOM)
        - Course Materials : BCOM
        - Projects and Thesis : BCOM
    - Bachelor of Science in Leadership and Management (BLM)
        - Course Materials : BLM
        - Projects and Thesis : BLM
    - Diploma in Leadership and Management (DLM)
        - Course Materials : DLM
        - Projects : DLM
    - Faculty Publications (FCOM)
        - Conference Proceedings : FCOM
        - Open Lectures and Other Presentations : FCOM
        - Press Clippings : FCOM
        - Published Articles : FCOM
        - Staff Thesis and Dissertations : FCOM
    - Master of Commerce (MCOM)
        - Thesis and Dissertations : MCOM
- Faculty of Information Technology (FIT)
    - Annual ICT Conferences : FIT
        - ICT Conference : 2007
        - ICT Conference : 2008
        - ICT Conference : 2009
    - Bachelor of Science in Business Information Technology (BBIT)
        - Course Materials : BBIT
        - Projects and Thesis : BBIT
    - Faculty Publications (FIT)
        - Conference Proceedings : FIT
        - Open Lectures and other Presentations : FIT
        - Press Clippings : FIT
        - Published Articles : FIT
        - Staff Thesis and Dissertations : FIT
    - Master of Science in Computer Based Information Systems (MSCIS)
        - Thesis and Dissertations : MSCIS

APPENDIX 4: REPOSITORY AGREEMENT (DISTRIBUTION LICENSE)

Strathmore University IR Deposit Agreement

By agreeing with and accepting this license, I (the author(s), copyright owner or nominated agent) agree to the conditions, as stated below, for deposit of the item (referred to as .the Work.) in the digital repository maintained by Strathmore University, or any other repository authorized for use by Strathmore University.

Non-exclusive Rights. Rights granted to the digital repository through this agreement are entirely non-exclusive. I understand that depositing the Work in the repository does not affect my rights to publish the Work elsewhere, either in present or future versions.

I agree that Strathmore University may electronically store, copy or translate the Work to any approved medium or format for the purpose of future preservation and accessibility. Strathmore University is not under any obligation to reproduce or display the Work in the same formats or resolutions in which it was originally deposited.

SU Digital Repository

I understand that work deposited in the digital repository will be accessible to a wide variety of people and institutions, including automated agents and search engines via the World Wide Web. I understand that once the Work is deposited, metadata may be incorporated into public access catalogues.

I agree as follows:

1.That I am the author or have the authority of the author/s to make this agreement and do hereby give Strathmore University the right to make the Work available in the way described above.

2.That I have exercised reasonable care to ensure that the Work is original, and to the best of my knowledge, does not breach any laws including those relating to defamation, libel and copyright.

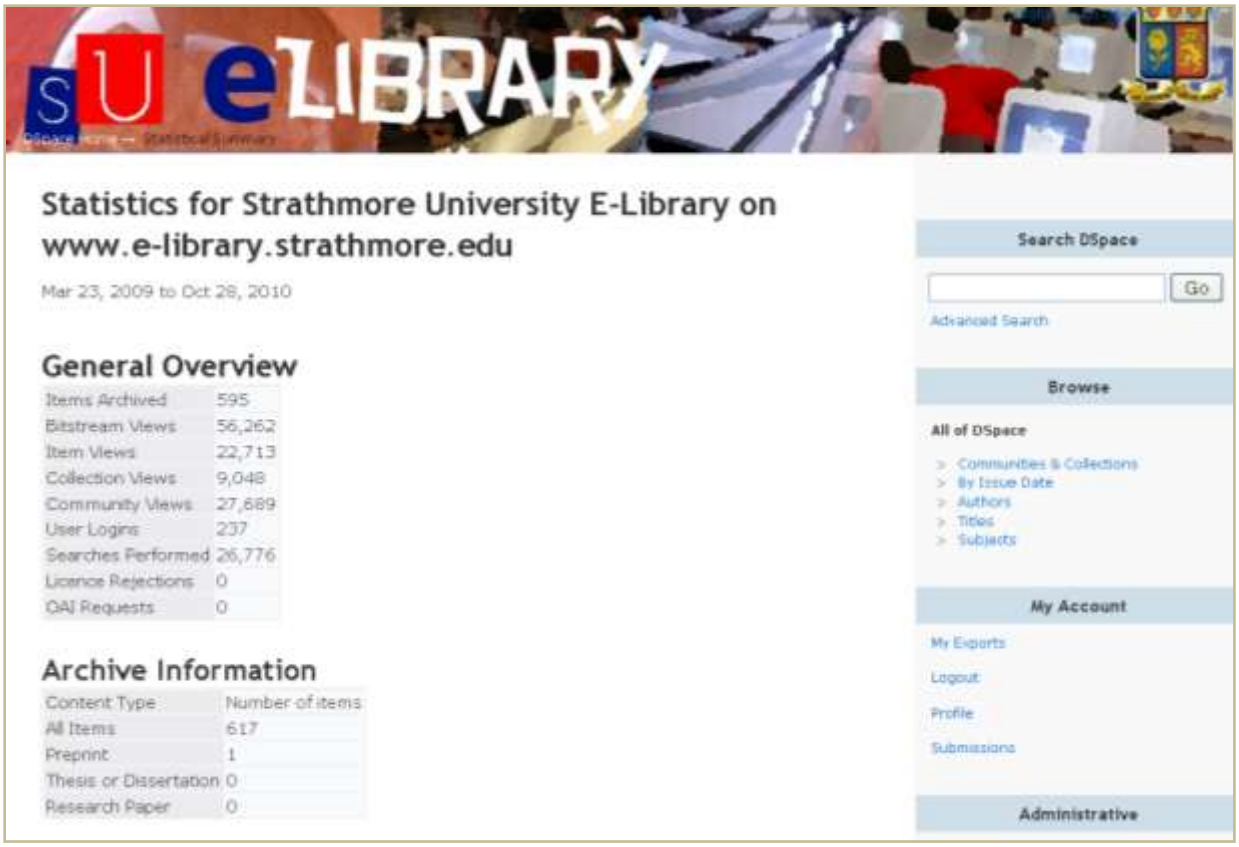
3.That I have, in instances where the intellectual property of other authors or copyright holders is included in the Work, gained explicit permission for the inclusion of that material in the Work, and in the electronic form of the Work as accessed through the open access digital repository, or that I have identified that material for which adequate permission has not been obtained and which will be inaccessible via the digital repository.

4.That Strathmore University does not hold any obligation to take legal action on behalf of the Depositor, or other rights holders, in the event of a breach of intellectual property rights, or any other right, in the material deposited.

5.That if, as a result of my having knowingly or recklessly given a false statement at points 1, 2 or 3 above, the University suffers loss, I will make good that loss and indemnify Strathmore University for all action, suits, proceedings, claims, demands and costs occasioned by the University in consequence of my false statement.

If you have questions regarding this license please contact the system administrators.

APPENDIX 5: USAGE STATISTICS



APPENDIX 6: REPOSITORY POLICY

DIGITAL REPOSITORY POLICY

Contents

1. What is SU-Portal

SU-Portal is the digital Institutional Repository of Strathmore University.
An intuitional repository is a set of services that an institution offers to the members of its community for the management and dissemination of its digital materials created by the institution and its community members. It captures and preserves the intellectual output of a single or multi-university.
SU Portal therefore aims to collate into one place the research , scholarship and Learning materials of members of its University community.
This short policy outlines issues surrounding the administration of the SU Portal.

2. Repository Structure

Open Source software shall be adopted for structuring of the repository. D-Space is the software of choice for the following reasons:
i) It has the largest community of users and developers worldwide
ii) It is completely customizable to fit our needs
iii) It is used by educational, government, private and commercial institutions
iv) It can manage and preserve all types of digital content

3. Submission Policy : Who can submit?

The repository will accept contributions only from bonafide members of Strathmore University.
Authors may only submit their own work. (where they own the copyright). All submissions will have to be vetted by designated faculty representatives to confirm appropriateness for deposit to the repository. The Librarian-in-charge will further confirm completeness, formats and appropriateness of item in line with the Library's Collection Development Policy.
The validity and authenticity of the content of submissions is the sole responsibility of the submitter/author as is any copyright violation. SU-Portal only serves a repository and bears no responsibility whatsoever for any copyright violations arising from a submitter.
If proof of copyright violation is validated, an item will be withdrawn from the repository.
Submission can be done at any one time, but availability will only be after vetting by faculty representatives. Where publisher's embargos are in force, they will be respected.

4. Submission Procedure

An item can be deposited to the repository at any time by a bonafide Strathmore member.
Such a member will have to register in order to submit. Submissions will then go to the appropriate faculty representative for vetting. After faculty approval, they will then be submitted to the librarian in-charge who will edit metadata and check for appropriateness of formats. After this, the item will then be accessible to the respective public.

5. Repository Contents Policy : What can be submitted?

The repository shall accept submissions in the following categories (within copyright limits)
• Thesis and Dissertations
• Learning Objects (Past Papers, Cases, Lecture Notes and Presentations,

• Conference Proceedings
• Journal Articles
• Book Chapters
• Inaugural lectures
• Any other appropriate academic content. The items to be deposited could include text, photos, audio, video and datasets. For long term
storage, material will be stored in Acrobat Adobe's Portable Document Format (PDF). JPG will be the preferred format for photos. Audio and video contributions are often large in size and will thus undergo additional vetting before posting them to contributions.

6. Metadata and Full Data Policy : How will the item be organized?
All materials deposited in the repository shall be tagged with metadata. This will include at least the author and title of work. The Dublin Core Metadata Standard shall be used to describe collections.
Meta data may be entered by the submitter of the item.
The librarian-in-charge can modify meta data of items.
Meta data will be accessible to all users in the intranet and/or internet.
Full Text
Depending on the category of collections and respective embargoes and library policy, some full text items in the repository will be available only within the intranet while some will be available on the internet.
Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. All users will be expected to use materials in the depository within copyright limits. Strathmore University Library will therefore not be liable for any copyright violation arising from use of items in the repository. SU Portal is not a publisher; it is merely the online archive that improves visibility of research and preserves digital contents.

7. Repository Use - statistics
Statistics on use of the repository will be availed. Additional modules that track details of use shall be availed as need arise.

8. Preservation Policy
Items in SU portal will be retained indefinitely. SU Portal will try to ensure continued readability and accessibility wherever possible.
Items will not be removed from the repository unless :
- There is proven copyright violation or plagiarism- There are legal requirements and proven violations
- National security is at stake
- The research contravenes Strathmore University philosophies.
In case of withdrawal from the above reasons, the entire item will be removed from SU Portal (meta data and full text)
Changes to deposited items will not normally be permitted. If necessary, an updated version may be deposited. Errata and corrigenda lists may be included with the original record if there is need for minor changes.

APPENDIX 7: STRATHMORE UNIVERSITY OPEN ACCESS POLICY

Strathmore University is committed to disseminating the fruits of its research and scholarship as widely as possible. In keeping with that commitment, the University adopts the following policy: Each University member grants to the Vice Chancellor and Academic council of Strathmore University permission to make available his or her scholarly articles and to exercise the copyright in those articles. More specifically, each Faculty member grants to the Vice Chancellor and Academic council of Strathmore University a nonexclusive, irrevocable, worldwide license to exercise any and all rights under copyright relating to each of his or her scholarly articles, in any medium, and to authorize others to do the same, provided that the artides are not sold for a profit. The policy will apply to all scholarly artides authored or co-authored while the person is a member of the University except for any articles completed before the adoption of this policy and any artides for which the Faculty member entered into an incompatible licensing or assignment agreement before the adoption of this policy. The Vice Chancellor or the Vice Chancellor's designate will waive application of the policy to a particular article upon written request by a Faculty member explaining the need.

Each Faculty member will provide an electronic copy of the final version of the article at no charge to the appropriate representative of the Vice Chancellor's Office in an appropriate format (such as PDF) specified by the Vice Chancellor's Office no later than the date of its publication. The Vice Chancellor's Office may make the article available to the public in an open access repository.

The Office of the Director of research will be responsible for interpreting this policy, resolving disputes concerning its interpretation and application, and recommending changes to the University from time to time. The policy will be reviewed after three years and a report presented to the Academic Council.