ETD2011

A comparative overview of the development of the institutional repositories at the University of Cape Town and at the University of Pretoria

Annah Macha
MPhil Student in the Department of Library & Information Studies
Annah.Macha@uct.ac.za

Karin De Jager
Department of Library and Information Studies
Karin.DeJager@uct.ac.za

ABSTRACT

The growth of open access institutional repositories has been remarkable in South Africa. Increasing numbers of institutions in South Africa are starting to implement digital institutional repositories in order to collect, disseminate, manage, preserve, and index the research output created by their communities. For a university, this would include materials such as scholarly publications, before (pre-prints) and after (post-prints) undergoing peer review, and digital versions of theses and dissertations. The Repository might also include special materials such as digital assets generated by and donated to an institution, such as art and a wide range of other image collections. Once digital materials are accepted by an Institutional Repository, the Repository should be able to guarantee access and preservation of the digital data.

South Africa at present has twenty-three Institutional Repositories. The leading example is the University of Pretoria, which helped the Council of Science and Industrial Research to establish its repository in 2007. Other universities are benchmarking from it.This paper will compare the Institutional Repository at University of Pretoria with the development of new Institutional Repository at the University of Cape Town In Pretoria, it was found that for an establishment and implementation to be successful, it depends on a number of factors that will be considered and discussed. An analysis of interviews with key informants from the University of Cape Town will be used to assess the extent to which the Institutional Repository at University of Cape Town is dealing with the identified issues.

The paper will critically discuss the importance and benefits of having an Institutional Repository and explore what the University of Pretoria is doing to produce a successful repository.

Keywords

Digital libraries; institutional repository; open access; scholarly publishing in South Africa;

INTRODUCTION

The establishment and development of institutional repositories are gaining momentum in South Africa (SA). In many institutions, there are documents which need to be preserved, managed, as well as shared and that is one of the main reasons for establishing an institutional repository in academic libraries. Other reasons for the establishment of an institutional repository are because it would help to preserve the institution's intellectual property and increase the institution's visibility and prestige. It also acts as an advertisement for funding sources and industrial sponsors (Prosser, 2003:168). According to Prosser, an institutional repository provides "a central archive of a researcher's work, it increases its dissemination and potentially, its impact on the research community and it may act as researchers' Curriculum Vitae as all their output is gathered in one place" (2003: 168). Not only does the institution and the individual benefit from the establishment of an institutional repository, society may also benefit as it provides access to an institution's research and ensures its long term preservation (Prosser, 2003:168).

This paper will explore the developments towards the establishment and implementation of an institutional repository at the University of Cape Town (UCT) and will compare this with the institutional repository at the University of Pretoria (UP) which has established itself as a pioneer and a leader in the development and management of scholarly repositories in South

Africa (Gray: 2009).

OVERVIEW OF SOUTH AFRICAN INSTITUTIONAL REPOSITORIES

The inaugural meeting of the Academy of Science of South Africa (ASSAf) Journal Editors' Forum that took place in late July 2007, marked the beginning of South African initiatives towards open access movement (ASSAF website). A national research and development strategy for South Africa had been published in 2002 (Pienaar and Van Deventer: 2008), which invited all stakeholders to reconsider and look into the renewal of the information services sector. This brought about the South African Research Information Services (SARIS) Project which investigated ways in which South African research institutes and university libraries would access the world research literature. It was recommended that a framework for eResearch services to the entire South African research community should be created. Organisations such as Electronic Information for Libraries (eIFL) and the Mellon Foundation provided assistance to universities with funding for starting up institutional repository projects. eIFL is an independent foundation that strives to lead, negotiate, support and advocate for the wide availability of electronic resources to enable and sustain access to knowledge by library users in developing countries. These two organizations' initiative helped to establish of a number of university institutional repositories in South Africa.

TOWARDS DEVELOPING AN INSTITUTIONAL REPOSITORY AT THE UNIVERSITY OF CAPE TOWN

The Manuscripts and Archives Department of the University of Cape Town Libraries began digitizing selected material in 2001 (Dunlop and Hart: 2005). It was a small-scale project, based on the unique collection of records about the San people. It involved the digitization of 310 photographs taken by Dorothea Bleek between 1910 and the late 1920s. It is one of the most important collections at UCT which attracts researchers from all over the world. It has also been listed on UNESCO's Memory of the World register as documentary heritage of international importance and that is why the department started this digitization project (Dunlop and Hart: 2005).

During the project, the staff at Manuscripts and Archives followed international guidelines and benchmarks regarding capture parameters and type of metadata. They also attended workshops given by Digital Imaging South Africa (DISA) which were helpful in giving them an overview of the processes and of the practical implications of undertaking a digital imaging project (Dunlop and Hart: 2005).

Other digitization projects that followed were instigated by individual departments. Computer Science started in 2003 to digitize theses and dissertations by the students in their department. The LawSpace Digital Subject repository was started in 2005, digitizing theses and dissertations of the students in the faculty of Law. Currently there are four different departments with repositories in UCT:

1. Faculty of Law - UCT LawSpace
2. Department of Computer Science - UCT Computer Science Research Document Archive
3. Department of Manuscripts and of Archives in the Library
4. Open Educational Resources - UCT - Open Content is not a repository but a directory for educational materials. It is the web portal for accessing open teaching and learning content from UCT. It is produced by the Open Educational Resources [OER] project in the Centre for Educational Technology at UCT. The directory aims to display the teaching materials of UCT academics and to encourage the publication of open educational resources.

While investigating the already functional departmental digital repositories, it became clear that there were plans to bring together the four departments and that a new UCT Institutional Repository was envisaged in the library.

THE ESTABLISHMENT OF THE INSTITUTIONAL REPOSITORY IN THE UCT LIBRARIES

Interviews were carried out to understand why the digital initiatives at the UCT Libraries were conducted at a relatively small scale in the early 2000s. Two reasons that were identified were cost and staff resources. Firstly, it was realized that establishing an institutional repository needed a budget for staffing, and for hardware and software to start up the project. Secondly, appropriately trained members of staff were required.

From around 2006, the Executive Director of UCT Libraries repeatedly requested the University to budget for the implementation of an institutional repository. The deans and other members of the executive were supportive, but according to the Executive Director, in a highly competitive fiscal environment, this request always came closely behind higher priorities. In September 2009 the Executive Director submitted the UCT Libraries' strategic plan and priorities for years 2010-2014 to the University Senate outlining in detail the plans of the libraries for the next five years. This document explains the impact that a Carnegie Grant had on the establishment of the UCT Institutional Repository.

In 2005, the Carnegie Corporation of New York had awarded the UCT Library, together with the libraries at the Universities of the Witwatersrand and Kwa-Zulu Natal, a grant amount of $2.5 million. This was for a three-year project directed at supporting research and library staff development at these institutions (Rapp: 2009). Among other projects this grant provided funds to develop a new Digital Initiatives Unit with up-to-date equipment and expert staffing. The Unit is in charge of spearheading the establishment of the Institutional Repository, which is intended to be launched in October 2011. The Digitization Unit, by establishing an institutional repository, is implementing two of the UCT Libraries strategic plans and priorities, stating that the Unit will:

> "Attract, preserve, digitize, and make available via a sophisticated web portal, key African archival and other resources. Digitize these resources as a contribution to African scholarship, as an incentive for digital collaboration on the continent, and as a showcase for UCT's research presence." (Rapp, 2009: 2)

The institutional repository will therefore:

a) " provide a web-based presence for UCT's collections and make them available to scholars in Africa and elsewhere,
b) enhance UCT's research presence on the web, and
c) enable better preservation of the primary physical resources." (Rapp, 2009: 6)

According to the UCT Libraries' strategic plan and priorities for the years 2010-2014, the Libraries will aim to:

> "Strengthen the Special Collections and Digitization Units to attract additional heritage collections and to enhance capacity to digitize special collections and research materials. In particular, focus on providing adequate and appropriate space and added curatorial expertise in specialized materials, such as film, video, photographs, and music." (Rapp, 2009: 6)

The Carnegie funding therefore played a major role in establishing the Institutional Repository at the University of Cape Town, enabling for example the purchase of equipment such as scanners and cameras and the DigiTool software.

The contents of the UCT repository at present is divided into three categories: *digital collections, finding aids and theses and dissertations*. It consists mainly of material from the old digitization projects that took place from 2001, combined with the new contents digitized by Digital Initiatives Unit since its inception to date.

Digital collections in the new institutional repository have a total of 1891materials as last updated on the 30th June 2011. It includes "Under exposed" materials which is an in-depth and special digital collection of photography, showcasing different twentieth and twenty-first century photographers' work, including both the pre and post-apartheid periods in South Africa. There is the "Varsity", which refers to the UCT student newspaper from the 1900s. "Sax appeal" covers show all the annual UCT Remember And Give [RAG] club Magazine covers from 1933 to date. The "Percival Kirby collection" from 1887-1970 is a large collection of images of musical instruments made and played by indigenous Southern African groups compiled by Professor Kirby. "UCT through the years" is another collection which dates back from 1900 to the present and is a collection of photographs from the 1900s of the Groote Schuur campus to the present UCT in 2011. "British concentration camps" is a collection about the concentration camps, which were formed by the British army to house the residents of the two Boer republics and the Orange Free State. "The San" photographs by Dorothea Bleek is a collection of 310 photographs from 1910-1920s. "Black Sash" is a collection of National Conference Papers on Black Sash minutes. "The Center for Popular Memory" has a collection of articles for center for popular memory and some on Islamic faith healing documents and interviews.

Finding aids consists of 866 materials, as last updated on the 30th June 2011.The collection is arranged in alphabetical order by categories. A "Finding Aid" is an inventory of an archival collection, and provides detailed information about the content, arrangement, and context of a collection (DigiTool website: 2011). These include; "South African College/UCT" which is a collection of documents and minutes of the history and events about UCT from the 1900 to-date. "Architectural Collections" consists of papers, architectural drawings; photograph albums; press-cutting books; pamphlets; pictures and sketches. "San people" is a collection of original documents on "Bushman" and "Bantu" philology and folklore. "Political Collection" contains documents and newspaper clippings on political issues in the 1940s about organizations and political parties, labour and trade unions and church unions and the formation of churches and associated politics.

Theses and dissertations is a collection of Masters and PhD student theses and dissertations. The UCT Institutional Repository at present has a total of 1099 theses and dissertations as last updated on 30 June 2011 from the six different faculties in UCT: Humanities, Science, Heath Sciences, Engineering, Commerce and Law dating from 2008 to 2010. The Institutional Repository can be found under the UCT library website, under the name DigiTool, the URL is http://www.digitool.uct.ac.za.

A DESCRIPTION OF UP ELECTRONIC THESIS AND DISSERTATIONS REPOSITORY

UP started making theses available online in the year 2000 as a pilot project (UPeTD website). Shortly afterwards the library was tasked with establishing an Institutional Repository for which the open access ETD-db software was used. At the end of 2002 the repository contained 39 theses and 26 dissertations and in 2003 a policy was adopted by Senate to make submission compulsory and a new server was bought. In 2004 UPeTD expanded in order to create a repository which was meant especially for research collections based on the success of UPeTD, which had provided PhD students according to Gray (2009) with research visibility and had contributed to their career success as their work was being read and downloaded.

The management of the UP Institutional Repository soon discovered that the Institutional Repository needed full time clerical staff. In 2006, UP established yet another Institutional Repository for the management and dissemination of digital research materials donated to or created by the community publications and for their special collections known as UPSpace. UP also has OpenUP which is a sub-collection of the larger UPSpace collection (Pienaar and Van Deventer: 2008). According to Olivier (2010), this "repository collection offers open access to full text of research articles published by staff, students and affiliates of the University of Pretoria." OpenUP houses e-prints, of peer reviewed and published research articles (UP Website). In total UP has two repositories, UPeTD and UPSpace while OpenUP is part of UPSpace (Olivier: 2010). Olivier (2010) further states that OpenUP offers authors two routes of publishing their research output which are; Open Access self-archiving which is known as the green route whereby authors submit their post-prints to the repository, and Open Access publishing which is known as the gold route whereby authors publish in open access journals.

UP became a member of the international body the Networked Library of Theses and Dissertations (NDLTD), which provides access to all the world's theses and dissertations and which hold annual conferences. It also rewards eTDs that use technology in innovative ways. UP has contributed to nine workshops in South Africa, Botswana, Lesotho, Ghana and Ethiopia, sharing expertise and enthusiasm and helping colleagues to start their own operations. They assisted the Council of Scientific and Industrial Research [CSIR] to develop a digital repository (Pienaar and Van Deventer: 2007). UPeTD was also the subject of seven conference presentations at ETD, open scholarship and library quality conferences. This indicates that to a considerable extent it is the leading university in SA in this field (UPeTD website-celebration of 10-yrs).

A COMPARATIVE ANALYSIS OF THE UP AND UCT INSTITUTIONAL REPOSITORIE

A detailed comparative analysis of UCT and UP will be discussed. UP was chosen for the comparative analysis based on its well established repositories and their successful implementation. When planning and deciding to establish an institutional repository, certain decisions and minimum requirements need to be taken into account. The decisions include:

> "How will the repository effort be staffed and funded, who will manage it, what platforms will be used, what initial policies will be made known and what marketing techniques will be used? Essentially these are the minimum required to get a repository up and running." Nabe (2010:13)

Below are some of the factors that UP has identified which may be taken into consideration when setting up an institutional repository (Smith & Pienaar: 2007).

1. Identify important role players

In order to establish an institutional repository at UP, the major role players had to be identified. These included "the head of the Department of Information Science, subject librarians who were comfortable with new innovative tools, a metadata specialist, a digitization specialist and IT staff" (Smith & Pienaar: 2007).

At UCT the main role player is the Head of the Digitization Unit who spearheads the UCT Institutional Repository project. Her role is that of project manager thus the Institutional Repository champion, and various staff members, such as subject librarians and metadata specialists, support her.

2. Address issues of resources

At UP, issues of resources were addressed by conducting a survey of available funds and other resources. Findings from the original survey showed that they had a suitable server, they had some funds to start the project, they also had enthusiastic staff who believed in the project, and they had support from the library management, specifically the Assistant Director of eInformation Strategy & Knowledge Management. They therefore went ahead with the project.

At UCT Libraries, the project originated from a small-scale project in 2001, which digitized uniquely African material. For this project, UCT libraries had their own funds. A lack of resources delayed the implementation and establishment of an Institutional Repository at the time. In 2009, UCT Libraries obtained funding from the Carnegie Foundation to support research in the library, when they began to set up an Institutional Repository, which is still under way. The intention is that it

will be launched in September/early October 2011. According to Nabe (2010), most of the funding for institutional repositories comes from the institutions and in some cases institutional repositories have been successful in receiving grants from organizations. This is what happened with the UCT institutional repository establishment.

3. Evaluate software

Both open source software and proprietary software are available for institutional repositories. It is important for an institution to choose software based on its need. At UP, it was decided to choose open source software and open standards based on the UP IT framework. The open source software chosen was ETD-db software developed by researchers at Virginia Tech for their theses and dissertation repository (Pienaar and Van Deventer: 2008). In 2004, the UP digital repository project team evaluated, over a six-month period, several open source and commercial software platforms such as Greenstone, Innovative, Fedora, E-prints, DSpace, and I-Tor for their UPSpace repository as the eTD used ETD-db. DSpace was eventually chosen because it fitted the UP IT architecture and supported a distributed approach to an institutional digital repository (Pienaar and Van Deventer: 2008).

At UCT, an extensive evaluation of available software for both open source and proprietary software was conducted in 2009. The evaluation assessed what software other universities nationally and internationally had chosen for their institutional repositories. It was found that nationally, most institutions had chosen either DSpace, an open source solution, or EPrints, which is also open source software. Internationally, DSpace was the open source software package of choice, while ContentDM and Luna Imaging were the proprietary packages of choice.

UCT Digital Initiatives Unit conducted an evaluation of the four major software solutions, namely DSpace, EPrints, ContentDM and DigiTool. Omeka is another software package that was also evaluated but was discarded as it was not yet widely used. The findings were that DSpace and EPrints were good solutions for non-visual materials or text-based materials, but not for visual material, like images. It was also found out that it was possible to customize the look and feel of the interface on both packages, although there was yet no instance where this had worked effectively. ContentDm and DigiTool were very similar in functionality and price. DigiTool was finally selected, as it is a reputable, vendor-supported tool, adopted by many major universities and the UCT library had a long-standing relationship with Ex Libris, the company that created DigiTool. It was also thought that the library did not have the capacity or did not wish to develop local expertise to sustain and maintain the installation of DSpace and EPrints. DigiTool would integrate well with the UCT online catalogue and UCT portal, PRIMO by Ex Libris, the new search interface, to be launched in September/early October 2011. Ever since UCT Digital Initiatives Unit started using DigiTool, it has made major strides in content deposits.

In addition to the reasons above, DigiTool offered UCT ongoing technical support in the form of a contract with Ex Libris. After evaluating various institutional repository platform options it was found out that DSpace did not adequately address the visual aspects of the UCT Institutional Repository, as photographs and other visual material are not displayed in a satisfactory manner in open source software such as DSpace or EPrints.

It is important to note that DigiTool can support open access as it is an OAI - PMH (Open Archives Initiatives-Protocol for Metadata Harvesting) complaint enterprise solution for academic libraries and research institutions (Stevenson and Hodges: 2007). DigiTool was designed for creating, managing, preserving and sharing digital collections. It also supports Dublin core. The Ex Libris website states that "the open-access platform also facilitates the sharing of scholarly information in a library's user community by integrating digital collections with institutional portals, e-learning systems, and other library systems." (Ex Libris website: 2011). Even though it is proprietary software, it therefore meets the needs and requirements of open access. UCT Institutional Repository will be an open access initiative in spite of being based on proprietary software.

4. Policy for the Institutional Repository

The need for a policy to guide the operation of the repository is an important factor to be considered. Issues such as what to accept or decline; copyright issues; self or mediated archiving; submission and withdrawal policies; types of materials to accept; and other issues governing the operation of the institutional repository form part of the policy of an institutional repository (Asamoah-Hassan: 2010). These aspects would need to be discussed and agreed upon by the institutional repository project committees. The policy should clearly indicate whether the repository is an open access initiative, and whether it would include abstracts as well as full text publications.

The Deutsche Initiative for Networked Information (DINI) recommends that repositories should make policies and make them available to everyone specifying how they will be run and what their content will be (DINI: 2011). The policy must also specify the rights and obligations of repository operators and of authors/editors.

At UP the Institutional Repository is governed by a policy which has been approved by the University, and is managed by the Institutional Repository team consisting of the University Librarian, the Systems Librarian and the Institutional Repository Librarian (Smith & Pienaar: 2007). UCT has a policy for the submission of print and electronic theses; in addition to that it

has guidelines for students on how to submit their theses and dissertations, but not yet an overall institutional repository policy statement.

5. Restructuring the library to accommodate change

The UP library management accommodated the repository by generating new roles and responsibilities during the implementation of their digital repository. In 2007 they had approximately 18 subject librarians who were involved as collection administrators. A total of 12 cataloguers who were involved as metadata editors and nine faculty library leaders fulfilled the roles of collection managers (Smith & Pienaar: 2007). Involvement in repository activities are part of the official role description of their cataloguers and subject librarians, and their performance is also monitored on the performance management system. To remove any possible uncertainty amongst role players, each role was clearly defined and communicated, and certain responsibilities were assigned to these roles so that all were familiar with exactly what was expected of them (Smith & Pienaar: 2007).

UCT Library management is in the process of restructuring roles and responsibilities of its staff in order to accommodate the change that will be brought about by the establishment of the Institutional Repository. Existing staff members will be offered new roles and new staff will be hired for new roles.

6. Licensing

In order to ensure open access to its collection, Stevenson & Hodges (2007) state that an institutional repository needs to be registered with open access harvesters such as OAISTER, the Registry of Open Access Repositories [ROAR], Open DOAR and Google Scholar in order to be harvested by these search engines. UP is registered with the ROAR, openDOAR, Google and as a publisher at Google Scholar, and registered with the DSpace community. After setting up the Institutional Repository, the project manager in UCT would also need to register with open access harvesters such as those noted above.

Licensing is what the DINI refers to as *visibility of the service.* A repository must have an entire range of services, which are available via a website. The Institutional Repository's homepage must be accessible from a central location on the institution's homepage. According to the DINI criteria an Institutional Repository has to be visible not only to the immediate and individual user but also to external services such as search engines or other referencing services so that it can be visible to the whole world (2011). In addition to registering with the search engines it must also be registered and listed on the DINI website.

Even though a repository can be harvested by providing OAI-PMH interface, it is still important to register repositories as it is a useful means of promoting the visibility of repositories to service providers for harvesting. The Open Archives Initiative provides a service allowing repositories to be registered as a data provider in the OAI registry (JISC Repository.Net).

CRITERIA FOR A SUCCESSFUL REPOSITORY

The criteria for the success of an institutional repository would normally be taken from its statement of purpose. An institutional repository can only be said to be successful if it fulfills its stated purpose, which includes to "preserve information and make it available for a designated community" (Thibodeau: 2007, 1). Another important factor to note when judging the success of an institutional repository is the context, or purpose it serves and the environment in which it operates. Thibodeau (2007) is of the view that "the metrics for estimating success against the criteria must be formulated in light of the culture, constraints and opportunities existing in the environment" (2007:1). Below is a discussion of some of the criteria that have been identified with which to determine success of an institutional repository.

DINI and DRIVER (Digital Repository Infrastructure Vision for European Research) discussed other pointers and criteria for success of institutional repositories. DINI noted eight which are:*visibility of the service*; *policy*, s*upport of authors and publishers;legal aspects*; *information security*; *indexing and interfaces; access statistics;* and *long-term availability*. DRIVER proposes guidelines which focus on five issues: *collections, metadata; implementation of OAI-PMH, best practices;* and *vocabularies and semantics*. It is concerned with textual resources that can be harvested with OAI-PMH (DRIVER: 2008).This is because the major resource-type provided by digital repositories is text and the major approach for offering these textual resources is the Open-Archives-Initiative Protocol for Metadata-Harvesting. The DINI Certificate provides a solid basis for what to consider when operating a repository. Since DRIVER looks at repositories from the perspective of an aggregator, the DRIVER Guidelines do not cover the aspects described in the DINI-Certificate that are designed for guiding the overall local operation of a repository (DRIVER: 2008). Below is a discussion of some of the criteria that have been identified with which to determine success of an institutional repository some of which do have both the DINI criteria embedded in them and the DRIVER guidelines.

1. Content

The content of an institutional repository is the most important factor that has been cited by researchers to show the success of a repository (Organ and Deveson, 2010; Thomas, 2007; Markey et al., 2009). Contents must be secured in order to be considered successful by the organisations or funders and other stakeholders. Markey et al. (2009) are of the view that not many researchers yet agree on what criteria to use to judge a successful repository but what seems to be clear is that content recruitment is an important factor to show whether the institutional repository is successful. Content recruitment is key because it literally is the core of the institutional repository. A critical mass of material is necessary to generate both additional content and end-user activity. Value-added services in the institutional repository are also seen as an important part of success. These include everything from full text retrieval to preservation (Markey et al.: 2009).

The contents of an institutional repository should include both born-digital and the older repurposed digital materials, and Westell (2006) argues that there should be more "born digital" materials in an Institutional Repository than repurposed materials. According to Westell (2006: 216), "the larger the critical mass of documents in an Institutional Repository, the more it will facilitate output measures." This means that if a repository has a lot of content then it will stand a higher chance of being visible to the outside community and the more it will be used, hence its citation rate will go up. *Long term availability* is a further criterion found in the DINI certificate and it can be discussed or embedded under the content factor; how the repository intends preserving these contents for future use. A DINI minimum requirement is that a document should at least be archived in a repository for no less than five years.

According to Olivier (2010) UPeTD has a total of 5693 materials. UCT currently has 4786 as updated on 30 June 2011 even though it has not yet been officially launched (Digitool website).

2. Use

Use is the second factor that has been identified as a success factor and can be divided into three sub-categories: *number of users*; *type of content used* and *nature of use* (Harnad and McGovern: 2009). The most widely used method to measure the use of a repository is by webometrics. According to Organ (2006) and Thomas (2007) "webometrics is the statistical technique of counting the gross number of items in the live repository and to measure retrieval for counting hits and downloads with the option to sort by country." Webometrics show how many hits have been made from the repository and how many articles have been downloaded. The more the downloads, the clearer it becomes that the institutional repository is being used. To the researchers whose papers are downloaded, it may mean that they are likely to receive more citations. Westell (2006) notes that statistics of this sort contribute to institutional research outcomes and are powerful data to be used for individual career progress.

As research outputs are increasingly exposed to the web via search engines such as Google, organisations will be looking at and making use of institutional repository statistical packages. According to Organ (2006), high citation rates of researchers' work in various institutions make vice-chancellors and CEOs accept these new kinds of research infrastructure, with the promise of improving and enhancing the reputation of their institution and research staff. In addition, bringing together input data, usage and citation analysis is helpful in giving the picture of how effective a repository is. These data give evidence that the institutional repository is being used by a variety of users and provide benchmarks of use and growth over time for institutional repository funders and university administrators (Westell: 2006).

According to DINI, the factor of use, which is measured by webometrics, is similar to its criterion *access statistics*, which looks at access statistics of documents and their publication rate for the qualitative, quantitative and technological evaluation of the repository (DINI: 2011). Therefore, every repository must log statistical data and this can be achieved by a tracking system of downloads and citations. This is similar to Bibliometrics and in websites it is called webometrics.

At UP Electronic Theses and Dissertations, 4317706 files were downloaded in 2009 and regular messages were received from users testifying to the usefulness of the system (UPeTD website). A survey done by Hammes (2007) also indicated that the ETD is appreciated by its alumni PhD students who were experiencing rewards in their careers because the ETD had made their theses visible. At UCT, an interview with the LawSpace repository manager stated that the Faculty of Law receives emails from people outside the country who use and appreciate the LawSpace repository, and ask why recent student theses have not been uploaded yet. According to the Head of the Digitization Unit, the UCT Institutional Repository intends to track its use through the daily downloads and hits.

3. Submission

The rate and number of submissions which lead to growth of the repository are the third factor to consider in judging if the

repository is a success. These are referred to as the "repository deposit activity measures" (Thomas: 2007). This means that "sustainable repositories are only possible with sustainable deposit profiles through active community engagement" (Carr and Brody, 2007: 18). A repository should have regular deposits that are broken down to daily deposits so that they may be spread across the institution and in that way the Institutional Repository is sustained. Institutional Repositories should avoid occasional huge deposits, as Thomas (2007), who quoted Carr and Brody (2007) proposed that "regular but less high volume inputs are a preferred way to sustain Institutional Repositories, describing it as the difference between the "gappy" and continuous deposit profile." Carr and Brody (2007) propose the use of the Registry Service ROAR (Registry of Open Access Repositories) to examine the performance of repositories according to the daily activity of any of its repositories. This factor of submission of deposits can also be sub-divided into categories:

> *Number of submissions* refers to the digital content that is deposited in the institutional repository. It can be a high or low number of submissions.
> *Frequency of submissions* refers to deposits being made in a continuously. As noted above it is recommended that submissions should be continuous over time not just large deposits occasionally.
> *Type of submitter* refers to the people who deposit contents in the Institutional Repository. They range from staff, to graduate students and undergraduate students. Researchers advocate for a broad representation of constituents in the institutional repository in a university; all departments should submit content.
> *Participation of key stakeholders* refers to participants as university administrators and funders, where they too submit contents in the Institutional Repository (Harnad and McGovern. 2009).

At UP, a policy was adopted by Senate in 2003 to make submission of research output compulsory. The UP Senate voted for a mandate for deposits and this is a sign that there is top-level support for the institutional repository establishment (Gray: 2009). At UCT the Library policy deals with submission of electronic theses and dissertations and digital collections. It is compulsory for students to hand in a digital copy of their thesis in the form of a CD ROM so that they could be uploaded into DigiTool. A policy regarding submission and deposits into the UCT institutional repository will still have to be formulated, approved and then passed on to the UCT Senate for final approval before the institutional repository can use it.

4. Advocacy

Advocacy is the fourth success factor to be considered. Advocacy deals with building an informed awareness for which a core message and ethos are essential. In order to ensure that the message of Open Access and the benefits of establishing and managing institutional repositories is spread as effectively as possible, staff members need to create a communication plan for the advocacy campaign. A number of different advocacy strategies can be used, including top-down and bottom-up, blanket and targeted approaches (DRIVER: 2008). This can be done by using informational web sites, promotional brochures and presentations to departments. Westell (2006) calls this promotion; he argues that a repository must be promoted to faculty staff in order for them to change their attitudes towards open access publishing, as it may be a new concept to them. Promotion is an important factor that can help populate the institutional repository once academics understand the changing culture of scholarly communications.

The literature shows that the project known as Securing a Hybrid Environment for Research Preservation and Access (SHERPA), which was set up in the United Kingdom (UK) mainly to inspire change, after helping UK universities to establish and implement institutional repositories, then ventured in advocacy projects. This was because many academic libraries in the UK had implemented institutional repositories and it was time to advocate for their use. Markland and Brophy (2005) are of the view that for an institutional repository to be successful, cultural change must be achieved through advocacy or "getting the right message to the right people with the tone and content varied by audience" (Johnson, 2007: 23). Though the institutional repository may have funding, staffing, fundamental values and policies, it still needs advocacy. SHERPA observed that neglecting advocacy resulted in repository decline and stagnation but, with the right level of engagement across the institution, advocacy became the key to success.

Johnson (2008: 24) is of the view that "advocacy is a route to achieve the crucial goal, real cultural change." This requires widespread engagement with academics, but time and effort are also needed. With regard to advocacy, UP formed a campus committee whose mandate was to create a framework within which the university policy, and guidelines and resources were made accessible to students, teaching staff and administrators. In order to make the university community of UP aware of the existence of the ETD, a web site was created, brochures produced, and regular anti-plagiarism campaigns organized and held (UPeTD website). UCT has not yet reached this stage; the institutional repository is still in its infancy and it is suggested that marketing plans will be put in place and executed in preparation for the launch.

5. Support

Support is the fifth factor to consider in judging the success of an institutional repository; without support from the university community and the stakeholders then the institutional repository will never be successful. Support comes in three forms, which are:

> *Constituent support* – The community that the institutional repository serves should support it. Content depositors should be willing to deposit their work without being persuaded first. Users too should express their satisfaction with the institutional repository by giving institutional repository feedback through blogs or emails.
> *Financial support* – is important for the setting up and maintenance cost while the institutional repository is running. The institution can provide ongoing, and preferably increasing support or the institutional repository may have funding from outside international organizations.
> *Technical support* – when setting up repositories it requires technical help and support. Thus "the initial set up of an institutional repository requires a mixture of library and IT skills, and these are not always in place" (Markland and Brophy: 2005, 13). There should be adequate support for and interest in the development and enhancement of the institutional repository infrastructure, the necessary software and tools are required. (Harnad and McGovern: 2009). Staff needs to be trained in order for them to work effectively and efficiently.

The Senate in UP voted to mandate for deposits and that is a sign that there is top-level support for the establishment of institutional repository (Gray: 2009). This is a sign of constituent support the very significant support that UPeTD gets from the University Senate. This means the institutional repository largely relies on mandates. This was supported by Harnad and McGovern (2009: 28) who stated that "mandates provide the only means for achieving a successful institutional repository," as UP also demonstrated.

At UCT in 2009 the library obtained funding from Carnegie Corporation to support research initiatives. This funding helped UCT to start setting up the institutional repository and to purchase the DigiTool software.

6. Influence

Influence refers to providing assistance to other institutions in establishing their own institutional repositories. This is similar to what Thibodeau (2006) calls collaboration, and it is about how the institutional repository associates with others in the country, region and in the world. Repositories are encouraged to work in collaboration with others to demonstrate that they are successful. If an institutional repository works in isolation, it might signify failure because it does not share and transfer knowledge. A more collaborative arrangement would exist where separate institutions decide to work together to leverage each other's strengths and where they form or join consortia for such purposes (Thibodeau: 2006).

This is already happening at UP which, is actively involved in helping other institutions set up their institutional repositories. It assisted CSIR and expressed interest in working with and helping the Southern African Regional Universities Association (SARUA) representatives who visited them in 2007. Though the UCT institutional repository has not yet been launched, it is already very active in supporting other institutions who wish to learn about the theory and practice of digitization. The Head of the Digitization Unit has delivered workshops to librarians and archivists; for example the Kimberley Digitization Workshop and the Digitization Workshop for Library of Parliament librarians. Here in UCT she is teaching digitization to Post Graduate Diploma and Honours students in Library and Information Studies. The UCT institutional repository staff have also advised various departments on campus, on metadata standards and the implementation of good metadata practice.

7. Interoperability

Lastly, an institutional repository should be interoperable, which refers to the capability of a computer hardware or software system to communicate and work effectively with another systems in the exchange of data, usually a system of a different type, designed and produced by a different vendor (Reitz: 2006). Interoperability can be achieved by being OAIster compatible; this means that institutional repositories must be accepted by standards provided by the Open Access Initiatives Protocol for Metadata Harvesting (OAI-PMH). These standards create interoperability, connecting distant institutional repository content through search engine capabilities. Without adhering to these standards, content searchability is limited and only accessible to community members who are locally aware of specific content. Westel (2006) states that interoperability indicates openness on behalf of the institution or the library which is willing to contribute to national and international scholarship.

The importance of interoperability is that OAI and other search engines such as Google Scholar can harvest their contents and in that it is advantageous to producers of repository contents thus researchers and authors as it exposes scholarship in a new way and puts it in an international context. The repository may be the only vehicle to distribute this material (Westel: 2006). A major success factor therefore is to meet the universal standards of OAI-PMH in order to make content more

accessible (Carpenter: 2008).

According to Nabe (2010:30), metadata also plays a role in the systems interoperability. This means that for digital repositories to achieve interoperability and to exchange digital objects between them, they firstly need to provide data and/or metadata to their partners which they can understand. "To achieve this, the parties need to agree on the structure, the syntax and the semantics of the data and/or metadata objects, and they need to either comply with common formats or to provide a means for metadata mapping and/or for object conversion" (Aschenbrenner & Kaiser, 2005: 25). The common format for repositories is the one proposed by OAI which is the Open Archives Initiative-Protocol for Metadata Harvesting (OAI-PMH) and for OAI compliance a repository must expose and allow dissemination of unqualified Dublin Core Metadata. The Dublin Core (DC) metadata schema is defined as the minimum standard to allow interoperability between repositories. This schema contains 15 elements defined by the Dublin Core Metadata Initiative (Aschenbrenner & Kaiser: 2005). There are three types of metadata, namely descriptive, administrative and structural. Descriptive metadata has basic information about a resource such as creator's name, title, subject content, description and publisher. The way this information is presented depends on the metadata standards being used. Dublin Core is compliant with OAI - PMH harvesting and all institutional repository solutions which use Dublin Core as their metadata standard are compatible with OAI. Metadata is also important in supporting repository management such as administration and preservation (Aschenbrenner & Kaiser: 2005).

OAI-PMH is one of the mechanisms used to achieve interoperability between digital repositories (JISC Repository.Net). It provides a system to facilitate the harvesting, sharing and discovery of distributed resources. This allows materials within repositories to be accessed by a greater number of users via external services. In addition, data harvested via OAI-PMH is now being used for a range of other repository applications such as reporting, enhanced user interfaces for direct searching of local repositories, and assisting with the ingest of data into other systems (JISC Repository.Net).

A CRITICAL SUMMARY DISCUSSING UP AS A BENCHMARK AND THE WAY FORWARD FOR UCT

From the comparison between institutional repositories at UP and UCT it is clear that although both institutions have repositories, they are not the same. Firstly, one can look at the differences between how UP and UCT's repositories started. UPs first repository was mainly to house electronic theses and dissertations (ETD) and later UP established two other repositories for research output for the institutions' academic community, UPSpace and Open UP respectively. UCT's repository does not primarily focus on electronic theses and dissertations, but it is for all the institutional digital objects including research output, such as research papers, data sets, images, digital collections, and also electronic theses and dissertations. At the moment UCT institutional repository mainly focuses on special collections and heritage collections found in South Africa. It has one of its image collections listed under UNESCO'S Memory of the World register as documentary heritage of international importance. The UCT Libraries strategic plan for the years 2010-2014 also shows that the repository's main focus is on:

> "Strengthening the Special Collections and Digitization Units to attract additional heritage collections and to enhance capacity to digitize special collections and research materials. In particular, focus on providing adequate and appropriate space and added curatorial expertise in specialized materials, such as film, video, photographs, and music."

Secondly, one can also look at the different types of software that each repository chose. UP choose open source while UCT choose proprietary software mainly because it was looking at its contents which will be dominated by images, so they wanted software that displays images clearly, which proprietary provided, while open source software could not. Photographs and other visual material are not displayed in a satisfactory manner in open source software such as DSpace or EPrints. DSpace, which does not provide the same visual impact as products as DigiTool or ContentDM. The main difference is that on open source software one does not get a visual overview of the photographic collections, as one would get with proprietary software.

Thirdly, UP can be taken as a benchmark by other institutions including UCT because of the success of its electronic theses and dissertations repository. It among the early institutions in South Africa to set up a repository and helped those institutions that needed help. It set itself as a good example when it came to establishing institutional repository in the country and held workshops and seminars to influence other research institutions to follow in the movement of open access and scholarly publishing.

CONCLUSION

This paper gave a comparative overview of the establishment of the UCT institutional repository and that of UP, highlighting the factors to consider when establishing an institutional repository. It showed that UCT started its repository after UP and the advantage is that it is now able to benchmark from the most well established institutional repository in SA.

Criteria for a successful institutional repository have also been discussed in order to show what needs to be done for an institutional repository to be successful and valuable to its stakeholders and all its users. It has been shown that institutional repositories can play an important part in a university as they collect and store all the university's research output, thus helping with its preservation and archiving. Institutional repositories are also important in enhancing the visibility of an institution in Africa and the world.

REFERENCES

1. Asamoah-Hassan, H. 2010. Alternative scholarly communication: management issues in a Ghanaian university. *Library Management*. 31 (6): 420-426. [Online]. Available: www.emeraldinsight.com/journals.htm?articleid=1864903&show=pdf [2011, May 5].
2. Aschenbrenner, A & Kaiser, M. 2005. *White Paper on Digital Repositories.* reUSE Digital Master Files of Printed Material. eContent Project No. 11173. Online. Available: http://www.uibk.ac.at/reuse/docs/reuse-d11_whitepaper_10.pdf (2011, August 3)
3. ASSAF website: http://www.assaf.org.za/
4. Carpenter, Nicole. 2008. Tune it up: creating and maintaining the institutional repository revolution. *Open and Libraries Class Journal*. 1: 1 [Online]. Available: http://www.infosherpas.com/ojs/index.php/openandlibraries/article/view/13/23 [2011, May 8]'
5. Carr, L. and Brody, T. 2007. Size isn't everything: sustainable repositories as evidenced by sustainable deposit profiles. *D-Lib Magazine*. 13(7/8): 1-19. [Online]. Available: http://www.dlib.org/dlib/july07/carr/07carr.html [2011, March 24].
6. DigiTool. Available: http://www.digitool.uct.ac.za
7. DRIVER. 2008. DRIVER-Digital Repository Infrastructure Vision for European Research. Information for institutional staff. [Online]. Available: http://www.driver-support.eu/personnel.html [2011, May 5].
8. Dunlop, J and Hart, L. 2005. Digitization projects at the University of Cape Town Libraries. [Online]. Available: www.innovation.ukzn.ac.za/InnovationPdfs/No30pp32-42Dunlop&Hart.pdf [2011, May 5].
9. Gray, Eve. 2009. University of Pretoria adopts open access mandate. [Online]. Available: http://blogs.uct.ac.za/blog/gray-area/2009/05/25/a-mandate-at-the-university-of-pretoria [2011, May 3].
10. Hammes, Monica. 2006. Institutional self-archiving (ISA): a multi-faceted start-up strategy to populate the institutional repositories of the University of Pretoria. *Poster at Open Scholarship*. 18-20 October 2006. University of Glasgow. [Online]. Available: http://repository.up.ac.za/upspace/handle/2263/1113 [2011, March 21].
11. Harnad, S and McGovern, N. 2009. Institutional repository success is dependent upon mandates. *Bulletin of the American Society for Information Science and Technology*. April/May 2009 35(4): 1-5. [Online]. Available: http://onlinelibrary.wiley.com/doi/10.1002/bult.2009.1720350410/pdf [2011, March 27].
12. JISC Repository Net. Open Archives Initiative-Protocol for Metadata Harvesting (OAI-PMH)*www.rsp.ac.uk/documents/briefing-papers/technical-%20**OAI PMH**.pdf* (2011, August 02)
13. Johnson, G. J. 2008. In the kingdom of the blind: successfully implementing institutional repositories in the United Kingdom and the Sherpa partnership experience. *New Review of Academic Librarianship*.13(1-2): 13-33. [Online]. Available: http://www.informaworld.com/smpp/content~db=all~content=a792024743pdf [2011, April 11].
14. Markey, K., Yakel, E., Rieh, S.H., Jean, B. and Yao, X. 2009. Secrets of success: identifying success factors in institutional repositories. University of Michigan School of Information. [Online]. Available: http://smartech.gatech.edu/bitstream/handle/1853/28419/118-449-1-PB.pdf;jsessionid=574A28A20DC3A9502E60A7D4E3F65187.smart1?sequence=2 [2011, March 10].
15. Markland, M. and Brophy, P. 2005. SHERPA project evaluation: final report. Manchester: Centre for Research in Library & Information Management. http://www.sherpa.ac.uk/documents/SHERPA_Evaluation.pdf [2011, March 17].
16. Moahi, K. 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 2009. [Online]. Available: http://wiredspace.wits.ac.za/bitstream/handle/10539/8950/35%20Moahi.pdf?sequence=1[2011, March 10].
17. Nabe, Jonathan A., 2010. *Starting, strengthening, and managing institutional repositories: a how-to-do-it- manual*. New York: Neal-Schuman Publishers

18. Olivier, Elsabe. 2007. The efficacy of institutional repositories: reflections on the development of a personalised collection on UPSpace . *Perspectives in Education*. 25(1), [Online]. Available: http://ol.up.ac.za/upspace/bitstream/2263/2277/1/Olivier_Efficacy(2007).pdf [2011, March 27].
19. Olivier, Elsabe. 2010. O*pen* access and the evolving scholarly. Communication Environment Workshop. September 1-2 2010. https://repository.up.ac.za/upspace/bitstream/2263/14788/1/Olivier_Electronic%282010%29.pdf [2011, May 14]
20. Organ, M and Lance Deveson, L. 2010.Two models for institutional repository success. [Online]. Available: http://www.google.co.za/#sclient=psy&hl=en&site=&source=hp&q=Organ%2C+M+%26+Lance+Deveson%2C+L.+2010.Two+Models+for+Institutional+Repository+Success&aq=f&aqi=&aql=&oq=&pbx=1&fp=76f37548cca82320 [2011. February. 15].
21. Organ, Michael. K. 2006. Download statistics - what do they tell us? The example of research online. The Open Access Institutional Repository at the University of Wollongong. Australia 2006. *D-Lib Magazine* 12 (11).[Online]. Available: http://ro.uow.edu.au/asdpapers/44 [2011, March 14].
22. Pienaar, H. and Van Deventer, M. 2007. Capturing knowledge in institutional repositories … playing leapfrog with giraffes. *Paper presented at the Knowledge Management preconference workshop at the World Library and Information Congress* (WLIC): 73rd IFLA General Conference and Council.17 August 2007. https://www.up.ac.za/dspace/handle/2263/3441. [2011, March 22].
23. Pienaar, H., Hammes, M. and Theo, B. 2006. Making UP research visible on the web. September 2006. Pretoria*:* CITTE http://repository.up.ac.za/upspace/bitstream/2263/822/3/webcitte.pdf [2011, March 11].
24. Prosser, C.D. 2003. Institutional repositories and open access: the future of scholarly communication. *Information Services and Use*. 23(2-3): 167-170. [Online]. Available: http://iospress.metapress.com/content/b0t9y69afjhylw9y/fulltext.pdf [2011. February. 15].
25. Reitz, Joan. M. 2006. ODLIS: Online dictionary for library and information science. [Online]. Available: http://lu.com/odlis/odlis_e.cfm [2011, April 26].
26. Smith, I. and Pienaar, H. 2007. "Adapt or die":energizing library staff and academics through the development of digital repositories. Paper presented at the *World Library and Information Congress* (WLIC): 73rd IFLA General Conference and Council, 21 August 2007. Knowledge Management Section.[Online]. Available: https://www.up.ac.za/dspace/handle/2263/3325 [2011, March 24].
27. Stevenson, V and Hodges, S. 2008. Setting up a university digital repository: experience with DigiTool. 24(1):48-50. *Emerald Group Publishing Limited*. [Online]. Available: http://www.emeraldinsight.com/journals.htm?articleid=1674217&show... [2011, May 6].
28. Thibodeau, Kenneth. 2007. If you build it, will it fly? Criteria for success in a digital repository. *Journal of Digital Information.* 8: 2.[Online]. Available: http://journals.tdl.org/jodi/article/view/197/174 [2011, April 24].
29. Thomas, Gwenda. 2007. Evaluating the impact of the institutional repository, or positioning innovation between a rock and a hard place. *New Review of Information Networking.*13(2). 133-146. DOI: 10.1080/13614570802105992 http://dx.doi.org/10.1080/13614570802105992 [24 March 2011].
30. UPeTD celebrates 10years of success. http://www.ais.up.ac.za/openup/docs/UPeTD10.pdf.
31. Westell, Mary. 2006. Institutional repositories in Canada institutional repositories: proposed indicators of success. *Emerald Group Publishing Limited. Library Hi Tech*. 24(2): 211-226. DOI 10.1108/0737883061066958. [Online]. Available: http://www.emeraldinsight.com/journals.htm?issn=0737-8831&volume=24&issue=2 [2011, April 20].