A Survey of Sustainable Curation in Research Repositories of Higher Education Institutions in Southern Africa

B. van Wyk
Department of Information Science, University of Pretoria, Private Bag X20, Hatfield 0043, South Africa
bvanwyk.uz@gmail.com

A.S.A. du Toit
Department of Information Science, University of Pretoria, Private Bag X20, Hatfield 0043, South Africa
adeline.dutoitconsulting@gmail.com

Abstract
The status and the prestige of higher education institutions depend on the quality, visibility and accessibility of their research. Globally, research indicates that valuable research output originates from both public and private higher education institutions, but the results of scholarship are often not archived and curated sustainably. Poor scholarship curation and lack of research visibility deter higher education institutions from taking their rightful place in higher education and higher education research communities. This article reports on investigations into digital scholarship curation trends in a purposefully selected target group of private and public higher education institutions in Southern Africa. Empirical research was triangulated with webometric analysis to derive solutions and best practices to ensure sustainable scholarship curation in institutional repositories. In all, 16 institutions were selected for the study. All the selected 16 institutions were subjected to webometric analysis but only 10 of the institutions completed the questionnaire. The study reveals a number of gaps affecting the effectiveness of institutional repositories in higher education institutions in Southern Africa. These gaps include true understanding of the nature and the importance of interoperability in open access. Also, collaboration within the higher education institutions, as well as external networks, is lacking. There is lack of awareness and knowledge regarding scholarship curation, and the value that web visibility holds for the entire institutions. The study recommends that institutions should include both social and technical aspects of scholarship curation.

Keywords: Digital Scholarship Curation, Sustainability, Open Access Institutional Repositories

Introduction
Higher education institutions are knowledge-intensive environments. Research and scholarship created in these institutions are institutional knowledge capital and must be managed as assets to give the institutions a competitive edge in research and academic stature. Knowledge capital is expected to be managed in a way that will ensure return on investment. Digital scholarship such as dissertations, theses, proceedings and publications form part of the knowledge capital created in higher education institutions. The curation of digital scholarship refers to the management, archiving and preservation of digital data over the lifecycle of the data (Yakel, 2007).

The digital curation of scholarship is expected to add value to existing knowledge and assist in creating new knowledge. Sustainability of digital collections and services, such as institutional repositories, is defined by Rieger (2011) as the ability to secure access to all resources needed to protect, maintain, develop and increase the value of a product's content and the service it has for the user there of. Anbu (2007) adds to this definition by stating...
that sustainability should include long-term preservation and curation of content and services in the institutional repository context of the definition. Sustainability is thus seen as surpassing mere successful implementation and content management of an institutional repository. Sustainability in institutional repositories and digital scholarship curation requires a socio-technical approach, where decision-makers need to realise its value and align technical and financial operations in support of scholarship curation (Rieger, 2011). Institutional repositories are expected to expand and develop to satisfy the environmental (academic) and socio-cultural (research cultural) needs of the higher education institution. The sustainability of institutional repositories poses challenges in institutions where the value of knowledge capital is not realised. Knowledge capital in the form of scholarship is expected to be purposefully and strategically supported by policies, processes and strategies on a high level of management. In some Southern African higher education institutions, especially private higher education institutions, sharing data in open access is slow.

There are 35 institutional repositories in Southern Africa registered on OpenDOAR (OpenDOAR 2016). Public higher education institutions have most of the registered institutional repositories in Southern Africa. The main problem that is addressed in this article is why the management of digital scholarship appears to be underdeveloped, in terms of lack of visibility, ranking and open access to research in South Africa. The article will explore how the application of information management and knowledge management principles should be applied in the sustainable curation of digital scholarship, which in turn will reverse the current state of affairs of low ranking educational institutions and poor access to scholarship.

African higher education institutions need to develop their own e-strategies to provide the framework needed to establish digital repositories. Thus, creating a mandate for African digital scholarship. Without the virtual research environment in an institution, the digital data curation cannot take place.

Digital Scholarship

Most higher education institutions in the developed world have fully incorporated and adapted to e-learning and digital scholarship. Lack of access to information and technology has a profound negative effect on the African digital scholarship. Mutula (2009) warns that Southern African higher education institutions that neglect to deploy e-learning and e-research in their institutions do it at their own peril. Collaborative research cannot take place without digital scholarship curation.

Digitised institutional repositories databases developed rapidly during the past ten years in most higher education institutions in the developed world (Smith, Barton and Branschofsky, 2003). Institutional repositories projects cannot develop in isolation and should support the aims and objectives of the educational institution as a whole. Digital scholarship is a networked, scholarly or academic environment extensively integrated with digital and information technologies in teaching and research (Mutula, 2010).

The whole of Africa still has only 5% of the global total of institutional repositories (OpenDOAR, 2015; OpenDOAR, 2016). The first developments towards electronic submission, storage and dissemination of theses and dissertations in Southern Africa date back to the early 1990s (Lor, 2005), followed by the establishment of the South African Research Information Services (SARIS) project which aimed at providing a framework for e-research services to all South African researchers (Van Deventer and Pienaar, 2008). Mutula (2008) laments the fact that African higher education institutions perform poorly in global web rankings because researchers publish in low impact journals with no internet links, and states that 80% of African higher education institutions suffer from no or poor internet connection.

Institutional Repositories and Open Access

Developments to promote access to research in the open access environment resulted in the creation of a number of treaties and agreements such as the Bethesda Open Access Statement (BOAI) in 2001 and the Berlin Declaration of 2003. The value of open access was communicated and encouraged and soon became the norm in institutional repositories. Awareness of the importance of open access research grew and gradually more institutions worldwide, and in Southern Africa, joined open access initiatives and movements by signing treaties and advocating open access. Recently, the value of
open access was communicated and encouraged and soon became the norm in institutional repositories.

Cullen and Chawner (2010) reported that institutional repositories were created with great initial enthusiasm, but it soon became just another task to be done. Generally, the focus in institutional repositories was on improving dissemination of digital scholarship and wider impact of research (Ball, 2010). Ball (2010) mentioned that institutional repositories were not initially tasked with preservation responsibilities, but as the content of repositories evolved to include more aspects of scholarship than just being a temporary storage until papers or research were officially published in mainstream publishing, this function became increasingly important. Digital curation and preservation need to be planned and managed with great care.

**Information and Knowledge Management in Digital Scholarship Curation**

Chaffey and Wood (2005) stressed that information and knowledge are increasingly valued as 'capital' in both business and higher education institutions. Rowley (2000) maintained that institutional knowledge must be embedded in knowledge management. Rowley (2000) gave these descriptions for the total knowledge existing in higher education institutions and not just scholarship. Scholarship and digital scholarship repositories are, however, seen as important subsets of the sum of all knowledge assets in higher education institutions.

The challenge of achieving sustainability lies not only in the institutional repository project itself and how information and data are managed, but also how the project relates to the bigger higher education institution's objectives. Sustainability of institutional repositories is dependent on how knowledge is seen, valued and managed on all higher education institution operational and decision-making levels. Effective sharing of knowledge created at higher education institutions remains a challenge. The role that institutional repositories should play in the management and curation of knowledge capital still needs to be formalised in policy and strategy by higher education institutions' decision-makers and governance processes. Jelavic (2011) posited that not only was knowledge management in institutions critical for success, but also that it should focus on the interrelatedness of the human element with the technical.

**Research Methodology**

The mixed methods research methodology of this study targets 16 purposely selected Southern African institutional repositories (IRs) as focus areas to observe their scholarship web presence and trends in scholarship curation. According to Best (2012), mixed methods research stems from pragmatism and is seen to strengthen the study by interrelating qualitative and quantitative methodologies (Best, 2012). Data analysis in mixed methods research allows for quantitative analysis of descriptive and inferential statistics.

Ranking Web of Universities was used to identify higher education institutions ranking below the top 500 global ranking universities and not included in the African top ten institutional repositories on Ranking Web of Repositories. The target group...
included 16 public and private higher education institutions in Southern Africa, offering postgraduate programmes and creating scholarly communication in the form of research (see Table 1). For the sake of confidentiality, the respondents in the public sector is referred to as A1-8, and in the private sector as B1-8). To further ensure anonymity rankings are supplied in intervals of 10 in Table 1 below. OpenDOAR and Web Ranking of Repositories are

Table 1. World and Sub-Saharan ranking of HEIs in the target group (Ranking Web Universities, May 2016)

<table>
<thead>
<tr>
<th>HEI</th>
<th>Type</th>
<th>Ranking Universities</th>
<th>World Ranking falling between Intervals of 500</th>
<th>Sub-Saharan Africa ranking falling between Intervals of 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>Public</td>
<td>Yes</td>
<td>3500-4000</td>
<td>50-60</td>
</tr>
<tr>
<td>A2</td>
<td>Public</td>
<td>Yes</td>
<td>3000-3500</td>
<td>40-50</td>
</tr>
<tr>
<td>A3</td>
<td>Public</td>
<td>Yes</td>
<td>500-1000</td>
<td>10-20</td>
</tr>
<tr>
<td>A4</td>
<td>Public</td>
<td>Yes</td>
<td>3000-3500</td>
<td>40-50</td>
</tr>
<tr>
<td>A5</td>
<td>Public</td>
<td>Yes</td>
<td>2500-3000</td>
<td>20-30</td>
</tr>
<tr>
<td>A6</td>
<td>Public</td>
<td>Yes</td>
<td>6500-7000</td>
<td>70-80</td>
</tr>
<tr>
<td>A7</td>
<td>Public</td>
<td>Yes</td>
<td>4000-4500</td>
<td>60-70</td>
</tr>
<tr>
<td>A8</td>
<td>Public</td>
<td>Yes</td>
<td>7000-7500</td>
<td>90-100</td>
</tr>
<tr>
<td>B1</td>
<td>Private</td>
<td>Yes</td>
<td>4500-5000</td>
<td>60-70</td>
</tr>
<tr>
<td>B2</td>
<td>Private</td>
<td>Yes</td>
<td>9500-10000</td>
<td>120-130</td>
</tr>
<tr>
<td>B3</td>
<td>Private</td>
<td>Yes</td>
<td>15000-15500</td>
<td>200-210</td>
</tr>
<tr>
<td>B4</td>
<td>Private</td>
<td>Yes</td>
<td>7500-8000</td>
<td>90-100</td>
</tr>
<tr>
<td>B5</td>
<td>Private</td>
<td>Yes</td>
<td>12500-13000</td>
<td>170-180</td>
</tr>
<tr>
<td>B6</td>
<td>Private</td>
<td>Yes</td>
<td>21500-22000</td>
<td>250-260</td>
</tr>
<tr>
<td>B7</td>
<td>Private</td>
<td>Yes</td>
<td>15500-16000</td>
<td>280-290</td>
</tr>
<tr>
<td>B8</td>
<td>Private</td>
<td>Yes</td>
<td>16500-17000</td>
<td>390-400</td>
</tr>
</tbody>
</table>

Catell and Fernberger, as cited in Jacobs (2010), researched the systematic use of bibliometrics and laid the foundation for further research. The mixed method used in this study included webometric analysis of the target's group web visibility and performance. Webometrics is a subset of bibliometrics. Bibliometrics is a scientific tool to measure research output (Jacobs, 2010). Jacobs reported that Eugene Garfield's Science Citation Index made analysis of research possible. There are three types of bibliometrics, namely descriptive, relational and evaluative bibliometrics. For the purpose of this study, evaluative bibliometrics is important, as it is a tool to assess the impact of scholarly work, as well as the quality of digital scholarly contributions to open access collections.
CURATION IN RESEARCH REPOSITORIES IN SOUTHERN AFRICA

authoritative examples and sources of reliable institutional repository statistics and performance monitoring worldwide. Webometrics analysis and institutional repository content analysis were used to gain deeper insight into the data collected from survey questionnaires. Analysing this data against webometric rankings gave insight into the inherent sustainability or lack thereof in the target group. For this study, quantitative data was collected from completed empirical survey questionnaires. Kim and Kuljis (2010) refer to content analysis as a useful qualitative methodology to examine web-based content, provided it is sampled and coded correctly.

Sixteen copies of the questionnaire designed for this study were sent out and 10, were received back. The feedback ratio on completed questionnaire was 62.5. The credibility of the research was measured by the Cronbach Alpha Coefficient and the scale employed was 0% to 100%, with a higher percentage indicating a higher credibility rating. An overall coefficient of 74.25% was calculated for the results obtained, and this is considered to be in the range of scores regarded as reliable.

Findings

The findings of this study were derived from webometric analysis and a survey based on the questionnaire distributed to the 16 institutions which was returned by 10 institutions.

Data Analysis of Ranking Web of

Repositories

Fifty-two Sub-Saharan institutional repositories were registered on Ranking Web of Repositories. Nine of the top ten repositories are situated in South Africa. The top ten are all from public universities. The top ten institutional repositories were explicitly excluded from this study, as the assumption based on their ranking and OpenDOAR profiles is that they are well funded, planned and managed. The ranking of top institutional repositories correlates with the ranking of top universities. Nineteen institutional repositories on Ranking Web of Repositories are registered in South Africa, two in Namibia, one in Botswana and four in Zimbabwe. The higher education institutions selected for this study were all ranked on the Ranking Web of Universities site, but the question was whether they were ranked and correlated as the top 10 higher education institutional repositories. Figure 1 presents a comparison of ranking positions of respondents. Five of the sixteen institutional repositories chosen for this study were ranked on Ranking Web of Repositories. Only one private higher education in this target group institution's repository was ranked. In all, six repositories in the target group were ranked, comprising 37.5% of the target population. Of all the institutional repositories in the target group, 62.5% were not ranked on Ranking Web of Repositories, indication poor web visibility to research.

Figure 1. Public versus Private HEIs on Ranking Web of Universities (2016)
Seven of the sixteen institutional repositories in this study were registered on OpenDOAR. Only one private higher education institution in this study was registered on OpenDOAR. Although respondents indicated their participation and appreciation of scholarship in open access, content analysis on OpenDOAR reveals that only three of all the higher education institutions supplied metadata standards information. The absence of reputable academic harvesters has a seriously negative impact on web visibility and is one of the reasons for low ranking and low impact. Findings indicate that respondents in this study were not OAI-compliant. The OAI Protocol for Metadata Harvesting (OAI-PMH) is a machine-to-machine interface provided by most repository software platforms (OpenDOAR 2016). They all indicated and supplied the metadata standards and re-use policies (OpenDOAR 2016). OpenDOAR offers clear guidance on how these policies can be added and also explains the benefits they had for increased web visibility.

Data Analysis Based on Questionnaire

Nature of Scholarship Production and Curation

All of the ten respondents indicated that both postgraduate students and academic staff members produced scholarship and communicated this scholarship through academic research platforms and publications. All the public higher education participants had digital repositories for showcasing their scholarship.

Strategies for Sustainable Curation of Scholarship and Research Output

Only five of the higher education institutions had a research strategy, IT strategy and an open access strategy. Nine indicated that they did not have a knowledge management strategy in place. This corresponds with the study by Blackman and Kennedy (2009), stating that higher education institutions are generally slow to take up knowledge management strategies, despite the potential benefits. Chakravarty and Wasan (2015) warned that where the institutional repositories performance was too low, policies and strategies should be reviewed to increase the volume and quality, making information management strategies a critical component of sustainable developments of institutional repositories (Chakravarty and Wasan, 2015. The results of this study indicate that Southern African higher education institutions are not yet on par with global trends.

Institutional Governance and Scholarship

Having strategies in place does not ensure best practice. Policy and procedure documents should be aligned with all higher education institutions' strategies. The nature of policies affecting scholarship and research output and communication shows that only five of the respondents, stated that institutional repository policies were in place. Only two had an open access policy in place. According to the answers, no one institution had a research information management strategy policy in place, indicating that the institutional repositories in the target group were not staying abreast of innovations.

Scholarship Curation, Policy and Procedure

Four of the respondents indicated that their library committee was the only governance body making decisions on institutional repository policies. Tian, Nakamori and Wierzbicky found in their 2009 study at a Japanese university that the biggest stumbling block in establishing knowledge management for the enhancement of research knowledge creation lies in the lack of higher education institution governance recognition, as well as their understanding and support in scholarship curation (Tian, Nakamori and Wierzbicky, 2009).

Value, Trust and Quality of Scholarship Curation

Seven of the respondents indicated that they were informed about all research related to digital projects in their respective higher education institutions. Eight of respondents were of the opinion that digital curation in institutional repositories should be a centralised function in the higher education institutions. Six of the respondents reported that their higher education institutions supported and funded research production. Four were of the opinion that research was secondary to teaching and learning at the higher education institution. This corresponds with
A study done in 2014 in Malaysian private universities by Thuraisingam et al. (2014), where they found that the research culture was not well established, and research and knowledge creation were indeed secondary to teaching and learning.

**Institutional Repository Relevance in Higher Education Institutions**

Eight of the respondents answered in the affirmative and indicated that their management and governance structures were informed about scholarship collections. On the question whether budgeting and separate funding for institutional repositories were in place, six of the respondents indicated that there were no separate budgets.

Constant development and innovation are requirements for success and development. Respondents indicated that new developments such as RIMS (Research Information Management System) and digital scholarship collections were jointly planned and managed. Only two indicated that these innovations were happening. Rieger (2011) stressed the importance of constant innovation and alignment with institutional developments as a critical factor in the sustainability of institutional repositories. Six of the respondents stated that their institutional repositories were well known in their institutions and research community. Eight of the respondents indicated that their scholarship collections were visible on their websites. However, content analysis on OpenDOAR indicates that even though scholarship is available on the websites, web visibility is compromised where open access harvesting and interoperability standards are not adhered to and implemented (OpenDOAR, 2016).

To a question on whether regular calls for participation and contribution of research output for submission to the repositories were made, four of the respondents answered that proactive efforts were made to populate their institutional repositories. This leaves six of the respondents open to random and inconsistent contributions by researchers' and students' scholarship to be curated in an organised and controlled way. Five of the repositories had a long-term preservation plan in place, but four had no preservation plans in place. After successful implementation, successful performance monitoring of institutional repositories growth and usage is cardinal for successful management of institutional repositories. Two of the respondents indicated that they were aware that their IR was ranked on Ranking Web of Repositories. Three respondents indicated that they were not ranked, and another three were not sure. Five of the respondents indicated that there had been clear development of their institutional repository. Four respondents indicated that there were no plans for maintenance and development. Despite low rankings, limited web visibility and lack of innovation, eight of the respondents felt that institutional repository managers were suitably skilled. Six of the respondents were using an open source software package to run their institutional repositories. Seven of the respondents indicated that their software had been upgraded during the past three years.

**Conclusion**

The main aim of this article, and the study, was to evaluate trends in digital scholarship curation in a purposefully selected target group. Participants in this target group were chosen for their existing web visibility and level of scholarship creation. Despite the fact that higher education institutions are knowledge-intensive institutions, where new knowledge is constantly created, researchers agree that knowledge management in higher education institutions in the form of knowledge management strategies, policies or even knowledge management awareness and conceptualisation is, surprisingly, rudimentary in most higher education institutions.

The empirical study reveals a number of gaps affecting the effectiveness of institutional repositories in higher education institutions in Southern Africa. Gaps were identified in terms of a true understanding of the nature and the importance of interoperability in open access. Collaboration within the higher education institution, as well as external networks, is lacking. Although respondents were of the opinion that institutional repository staff were well qualified, and that their higher education institutions were supportive and knowledgeable about open access, triangulation with webometric analysis indicated the presence of factors that had a negative impact on sustainability of the institutional repository. When triangulating the findings of the questionnaire survey results with recent statistics obtained from the reputable web directory, OpenDOAR, all indications are that the institutional repositories may be at peril, as serious sustainability threats surfaced.
This research explored how information management and knowledge management principles could improve the archiving, preservation and curation of digital scholarship, ultimately to enhance access to valuable research produced in Southern African higher education institutions. The research revealed that there is still insufficient understanding and support of scholarship curation at governance level. The study revealed serious gaps in the understanding of open access and application of open access protocols and standards.

There is lack of awareness and knowledge regarding scholarship curation, and the value that web visibility holds for the entire institutions.

**Recommendations**

The importance of research visibility is not realised by many higher education institutions. The sustainable management of scholarship in digital open access repositories must be prioritised. Higher education institutions' rankings, as well as repository rankings, need to be reported to decision-makers and their performance monitored. Knowledge management for sustainability needs to begin at a statutory decision-making level, where the institutional repository is formally recognised and incorporated into higher education institution's governance processes. Based on the research, it is clear that a suitable definition for a sustainability domain must include both social and technical aspects of scholarship curation.

**References**


Brenda van Wyk is the Head of Library and Information Services at the Independent Institute of Education, an internationally accredited private higher education institution. She holds a Master's degree in Information and Knowledge Management from the University of Johannesburg. She has been a part-time lecturer at University of South Africa and the University of Zululand. She is currently a PhD student in Information Sciences at the University of Pretoria, South Africa.

A.S. A. du Toit is Professor and Director of the Centre for Information and Knowledge Management at the University of Johannesburg. She also acts as consultant to various private enterprises.
Copyright of African Journal of Library, Archives & Information Science is the property of African Journal of Library, Archives & Information Science and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder’s express written permission. However, users may print, download, or email articles for individual use.