Effective Ingestion of Digital Objects in Institutional Repositories Using Subject Repositories

Moonga HABUKALI · Mathews MBEWE · Nozyenzi MWALE · Mathews MWEWA · Nandi SIKAZINDU · Lighton PHIRI

The University of Zambia
Zambia
Who Am I? (1/2)
• Currently a faculty staff at The University of Zambia
• Previously a graduate student at The University of Cape Town
• Founding member of The DataLab Research Group, comprising of academics and students, working in three main areas:
  – Data Mining
  – Digital Libraries
  – Technology-Enhanced Learning

Research

Members of the DataLab group conducted research in the following broad areas:

Data Mining
With the proliferation of data, the field of Data Mining has gained rapid popularity. Data Mining tools large datasets by making use of statistical and machine learning techniques. Our current focus involves leveraging machine learning techniques to facilitate efficient and effective educational domains—two areas that are of significance in the so-called developing world.

Digital Libraries
The field of Digital Libraries (DLs) generally involves the study of digital collections of information services used to retrieve data from the collections. DLs are in effect information systems that are objects, manage the digital objects and, facilitate access to the digital objects. Our focus in the field of DLs, as a research group, mostly involves experimenting with techniques and effective access to digital objects stored in DLs.

http://datalab.unza.zm
Online visibility of research output produced in the so-called Global South is noticeably low [...]  
- Recent studies show that Africa generates less than 1% of the world’s research
• Online visibility of research output produced in the so-called Global South is noticeably low [...]  
  – Recent studies show that Africa generates less than 1% of the world’s research
• Disproportionate online representation is apparent in most academic databases and is especially concerning [...]
• Disproportionate online representation is apparent in most academic databases and is especially concerning [...]

https://oatd.org
Motivation, Problem Area (3/5)

Research Visibility in the Global South: Towards Increased Online Visibility of Scholarly Research Output in Zambia

Scholarly research and publication form an integral part of the core functions of Higher Education Institutions (HEIs). It is generally standard practice for HEIs to deposit scholarly output into publicly accessible Institutional Repositories (IRs). While Zambia has seen a rise in the number of HEIs, with a total of six Public HEIs and 60 Private HEIs, there is little online visibility of scholarly output generated by these HEIs. A bibliometric analysis, focused on electronic theses and dissertations (ETDs), was conducted by harvesting scholarly publications from HEIs IRs, in order to demonstrate the low online visibility of scholarly research output in Zambia. We also outline technological initiatives, by using case examples from The University of Zambia, that can be employed to potentially increase the online visibility of HEIs scholarly output. Specifically, we illustrate how subject repositories and downstream aggregate services can be utilised to increase the visibility of scholarly output. The study shows that only two HEIs have established IRs, with noticeably low scholarly publications by academic staff. Furthermore, there is a noticeably long delay between the publication date of the ETDs and the ingestion date into the IRs. In addition, while not comprehensive, the proposed initiatives demonstrate technological initiatives that could be employed to increase the visibility of scholarly research output.

URI
http://dspace.unza.zm/handle/123456789/5723

http://dspace.unza.zm/handle/123456789/5723
Research Visibility in the Global South: Towards Increased Online Visibility of Scholarly Research Output in Zambia

Scholarly research and publication form an integral part of the core functions of Higher Education Institutions (HEIs). It is now generally standard practice for HEIs to deposit their publications and research data in Institutional Repositories (IRs). While Zambia has a total of six Public HEIs and 60 Private HEIs, a bibliometric analysis of the online visibility of all publications generated by these HEIs, revealed that very little of the scholarly output generated by these HEIs is available online. A bibliometric analysis of the institutional repositories (ETDs), was conducted by the University of Zambia, in order to demonstrate the low online visibility of the scholarly output of these HEIs. We also outline technological initiatives, that can be employed to improve the online visibility of HEIs scholarly output. Specifically, we illustrate how institutional repositories and aggregate services can be utilised to increase the online visibility of HEIs publications. The analysis shows that only two HEIs have established institutional repositories and are currently publishing publications online. Furthermore, there is a correlation between the publication date of the ETDs and the ingestion date of the ETDs in the institutional repository. The analysis is intensive, the proposed initiatives demonstrate how institutional repositories can be employed to increase the visibility of scholarly research.
Open Access Electronic Publishing for Increased Online Visibility: Tooling Challenges and Potential Solutions

Lighton Phiri <lighton.phiri@unza.zm>
Department of Library and Information Science
University of Zambia

ZCAS Colloquium on “Sharing Knowledge and Best Practices”
Motivation, Problem Area (4/5)

Open Access Electronic Publishing for Increased Online Visibility:

Towards Open Access Institutional Repositories in Zambian HEIs

Abel M’kulama <abel.mkulama@unza.zm>
Lighton Phiri <lighton.phiri@unza.zm>

Department of Library and Information Science
University of Zambia
Automatic classification of digital objects for improved metadata quality of electronic theses and dissertations in institutional repositories

Lighton Phiri

https://doi.org/10.1504/IJMSO.2020.112804

Published online 26 January 2021

Abstract

Higher education institutions typically employ Institutional Repositories (IRs) in order to curate and make available Electronic Theses and Dissertations (ETDs). While most of these IRs are implemented with self-archiving functionalities, self-archiving practices are still a challenge. This arguably leads to inconsistencies in
Motivation, Problem Area (5/5)

- Ingesting digital objects into The UNZA’s IR is challenging
  - Ingestion is time consuming due to the complex ingestion workflows
  - There is a backlog of digital content not yet ingested
• Ingesting digital objects into The UNZA’s IR is challenging
  – Ingestion is time consuming due to the complex ingestion workflows
  – There is a backlog of digital content not yet ingested
Research Objectives

• The main objective of this study was to investigate the feasibility of using subject institutional repositories for ingesting digital objects
  – To what extent is scholarly research output visible via the institutional repository?
  – What is the feasibility of implementing a subject repository as part of an institutional repository ingestion workflow?
  – What is the perceived usability of a subject institutional repository?

• Main contributions of this work
  – Evaluating IR ingestion practices, using a content analysis approach of IR digital objects
  – Results from usability studies conducted to determine the outcome of using subject repositories
  – Demonstration of the feasibility and efficacy of using subject repositories
The study employed a mixed methods approach that involved undertaking a situational analysis, in order to understand the current state of the problem and, implementation and subsequent evaluation of a subject institutional repository

1. Content analysis of repository objects
2. Interview sessions aimed at understanding the ingestion workflow
3. Implementation of subject institutional repository
4. Usability evaluation of subject institutional repository
Research Approach (2/4)

- Content analysis
  - Dublin core descriptive metadata harvested using the OAI-PMH protocol
  - Metadata elements analysed
- Interviews with staff responsible for ingestion of digital objects
Research Approach (3/4)

Library and Information Science Document Archive
The Department of Library and Information Science document archive stores gray literature—technical reports, undergraduate research project reports and past examination papers—generated in the department.

Communities in Archive
Select a community to browse its collections.
- Scholarly Research Outputs
- Undergraduate Project Reports

Recently Added
Assessing the awareness and perceptions that researchers and lecturers have towards self-archiving at the University of Zambia
Lungu, Paul; Kaluwa, siiluta; Kapembwa, Musonda; Wamuwi, Liywali; Mumba, Chato; Chitengi, Chinyama (2018)

Barriers Associated with Open Access Publishing
Lubasi, Freeborn; Ndebele, Norah Nokuthula; Mayaka, Dora Everyn (University of Zambia, 2018)

Information needs and information seeking behaviour of pregnant women on antenatal in Lusaka Zambia: A case study of Mtendere health Centre
Mulenga, Brian; Imasiku, Margaret; Chanda, Vanessa; Musonda, Konsole; kabwe, Chansa (University of Zambia, 2018)

http://lis.unza.zm/archive
• Subject repository implemented using DSpace, an open source repository software platform

http://lis.unza.zm/archive
**Research Approach (4/4)**

- **Subject repository integrated with institution-wide repository using OAI-PMH and OAI-ORE protocols**
  - Two-way synchronization of digital objects
Major Outcomes/Results (2/4)

Total Digital Objects

<table>
<thead>
<tr>
<th>Theses and Dissertations (ETDs)</th>
<th>Pre-prints</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education (ETDs)</td>
<td></td>
</tr>
<tr>
<td>Medicine (ETD)</td>
<td></td>
</tr>
<tr>
<td>Humanities and Social Sciences (ETDs)</td>
<td></td>
</tr>
<tr>
<td>Law (ETDs)</td>
<td></td>
</tr>
<tr>
<td>Natural Sciences (ETDs)</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td></td>
</tr>
<tr>
<td>Agricultura l Sciences (ETDs)</td>
<td></td>
</tr>
<tr>
<td>Institute of Distance Education (ETDs)</td>
<td></td>
</tr>
<tr>
<td>Engineering (ETDs)</td>
<td></td>
</tr>
<tr>
<td>Mines (ETDs)</td>
<td></td>
</tr>
<tr>
<td>Library (ETDs)</td>
<td></td>
</tr>
<tr>
<td>Veterinary Medicine (ETDs)</td>
<td></td>
</tr>
</tbody>
</table>

Institute of IST-Africa 2021
Copyright 2021 Authors
Major Outcomes/Results (2/4)

- 523 publications against ~854 faculty staff
- ONLY 476 (55.7%) faculty staff with online publications
Major Outcomes/Results (3/4)

- Ingestion of digital content is not done in a timely manner, compromising the recency of digital content
- Backlog of legacy content not yet uploaded
Major Outcomes/Results (4/4)

- Overall, acceptable SUS score

![Bar chart showing SUS scores across different bands]

- <=40: 1 participant
- 41-50: 5 participants
- 51-60: 5 participants
- 61-70: 17 participants
- 71-80: 4 participants
- 81-90: 3 participants
- 91-100: 1 participant
Major Outcomes/Results (4/4)

- Overall, acceptable SUS score supported by reasonably “Good” adjective rating
Conclusion and Outlook

• A case study aimed at assessing the feasibility of using subject IRs for facilitating effective ingestion of digital content
  – Empirical analysis of digital objects ingested into The UNZA IR demonstrates the extent of these challenges.
  – Implementation and usability evaluation of the subject IR demonstrates the feasibility of implementing usable subject IRs
• The obvious benefits of subject IRs is that self-archiving is decentralised, making it relatively easier for research output to be made available online
  – HEIs are better placed to ensure that research output is made available in IRs in a timely manner
Bibliography

Who Am I?

lighton.phiri@unza.zm
http://datalab.unza.zm
http://lis.unza.zm/~lightonphiri