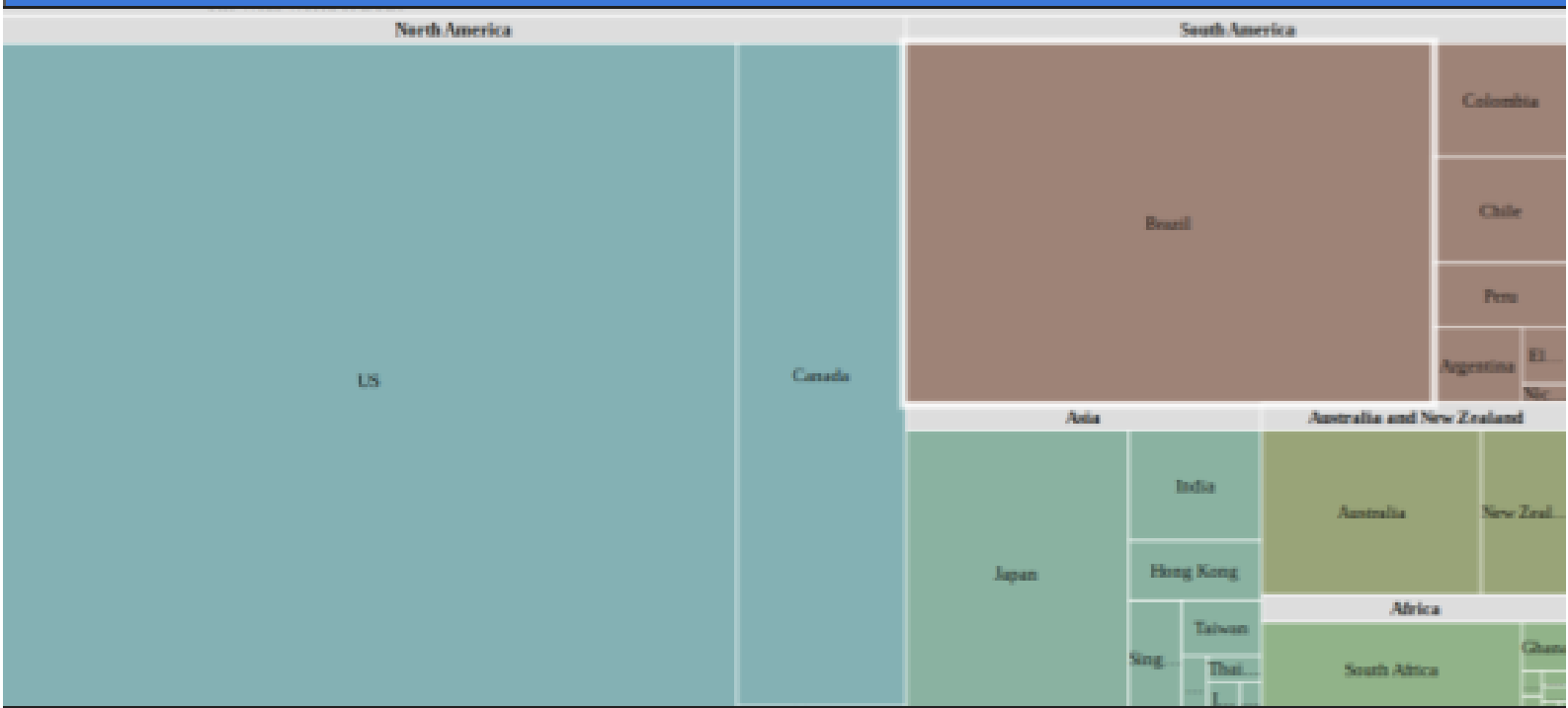


ETD 2024

27th International Symposium on Electronic Theses and Dissertations

*"Electronic Thesis and Dissertation Visibility at a Global
Scale"*

November 4-6, 2024
Livingstone · Zambia



Contents

About	3
27th International Symposium on Eletronic Theses and Dissertations	3
Local Conference Organizing committee	4
Volunteers	4
Programme committee	4
Partner Institutions and Sponsors	5
Gold, Silver and Bronze Sponsors	5
Partner Institutions/Organisations	5
List of Abstracts	6
Workshops	6
Full Papers	10
Posters	31
Invited Talks	48
List of Participants	50
Conference Proceedings	53
Posters	53

27th International Symposium on Electronic Theses and Dissertations

The 27th International Symposium on Electronic Theses and Dissertations (ETD 2024) aims to bring together global leaders and researchers working in the broad areas of digital libraries, institutional repositories, scholarly research and electronic theses and dissertations. The theme of ETD 2024 is "ETD Visibility at a Global Scale", and will explore innovative approaches that make use of Electronic Theses and Dissertations (ETDs), including the use of modern-day Artificial Intelligence techniques such as Large Language Models and exploration of advances that will result in increased visibility of ETDs at a Global Scale.

Thank you so very much for coming to Livingstone, Zambia to participate in the ETD 2024 conference events. We sincerely hope you enjoy the conference and your short stay in Livingstone, and, more importantly have productive and worthwhile discussions and meet new people!

Zikomo Kwambili!

Best Wishes
Lighton Phiri (Chair)

Local Conference Organizing committee

Lighton Phiri, UNZA	Denny Nsokolo, HEA	Francina Makondo, UNZA
Phyela Mbewe, UNZA	Thabiso Mwiinga, UNZA	Kaoma Daka, UNZA
Matildah Muchinga, LAMU	Stein Mkandawire, ZAMREN	Buumba Dubeka, ZCAS
Adrian Chisale, UNZA	Dokowe Tembo, ZCAS	Cecilia Kasonde, KNU
Habeenzu Mulunda, UNZA	Francis Kawesha, HEA	Fabian Kakana, UNZA
Brian Munkondya, MU	Elijah Chileshe, UNZA	Mpande Ntumbo, ZAMREN

Volunteers

Mubanga Chibesa, UNZA	Gift Muwele, UNZA	Lwiimwe Shansonga, UNZA
Albertina Mooka, UNZA	Christabel Kunda, UNZA	

Programme committee

Ana Pavani	Pontificia Universidade Católica do Rio de Janeiro
Charles Greenberg	NDLTD
Edward Fox	Virginia Tech
Gabriela Mejias	DataCite
Hussein Suleman	University of Cape Town
Iryna Kuchma	EIFL
Jian Wu	Old Dominion University
Joachim Schöpfel	University of Lille
Lazarus Matizirofa	University of the Witwatersrand
Lighton Phiri	University of Zambia
Maïté Roux	ABES
Mirjana Brkovic	University of Novi Sad
Nabi Hasan	Indian Institute of Technology Delhi
Ramesh Gaur	Indira Gandhi National Centre for the Arts
Shantashree Sengupta	P.E.S Modern College of Arts Science and Commerce
Tainá Batista De Assis	Brazilian Institute of Information in Science and Technology
William Ingram	Virginia Polytechnic Institute and State University

Partner Institutions and Sponsors

The ETD 2024 conference was organised by the Networked Digital Library of Theses and Dissertations (NDLTD) and, additionally hosted by The University of Zambia (UNZA) and co-hosted by The Higher Education Authority (HEA) of Zambia and Zambia Research and Education Network (ZAMREN).

Gold, Silver and Bronze Sponsors



Partner Institutions/Organisations



List of Abstracts

Workshops

ETDs 101: No Experience Required!

Charles Greenberg, Networked Digital Library of Theses& Dissertations, United States

Many universities do not yet participate in the global network of digitized student theses or dissertations. The Networked Digital Library of Theses and Dissertations (NDLTD) wants to help universities to get started by sharing experiences of those that started ETD programs from nothing. Among the topics to be covered are: (i) Lifecycle Management of ETDs; (ii) Institutional Partnerships; (iii) Technology–Institutional repositories and OAI-PMH; (iv) Intellectual property issues; (v) Building an institutional consensus and (vi) Regional, National, and International support.

Please bring questions for the NDLTD Board members and paper presenters.

DataCite Connect: Global visibility for your ETDs and beyond with PIDs

Gabriela Mejias, DataCite, Germany

PIDs and their associated metadata can be considered the building blocks of research infrastructure (Meadows et al., 2019). In addition, PIDs play a central role in the Open Science framework as they can increase transparency and recognition in research, and facilitate integration and interoperability, as well as making research FAIR (Wilkinson et al. 2016).

This half day workshop will explore how Persistent Identifiers (PIDs), particularly Digital Object Identifiers (DOIs), can significantly boost the global visibility of Electronic Theses and Dissertations (ETDs) and other research outputs. We'll cover the basics of PIDs, showing how DOIs improve visibility and discoverability for academic work. Through case studies, you'll see real-world examples of research organizations that have successfully integrated PIDs into their ETD repositories and other workflows. The session culminates in a breakout discussion to share best practices and explore implementation strategies. This workshop offers valuable insights for enhancing the visibility of your research through DataCite's PIDs, providing a unique opportunity to connect with the broader research community. Proposed activities:

- (20 min) Welcome and Ice breaker activity - (30 min) Introduction to Persistent Identifiers (PIDs) and how they support open research - (30 min) Discover the role of Digital Object Identifiers (DOIs) in making ETDs more visible and discoverable on a global scale. - (30 min) Improving ETD and other research workflows with PIDs. - (20 min) Coffee break - (40 min) Case studies featuring invited guest speakers from research organizations implementing PIDs in their ETDs, research data repositories (and other!) workflows - (50 min) Breakout session to discuss best practices and implementation workflows - (10 min) Adjourn

Details of similar previous workshops:

<https://datacite.org/event/datacite-connect-gothenburg-dataciteconnect23/>
<https://datacite.org/event/datacite-connect-buenos-aires-dataciteconnect23/> ETD & Open Science: Maximizing the Discoverability and Impact through Persistent Identifiers (PIDs) <https://etd2023.inflibnet.ac.in/programme.php>

Facilitators: Gabriela Mejias, Bosun Obileye (DataCite). Guest speakers from research institutions will be confirmed upon workshop acceptance notification.

Globalizing Knowledge: Leveraging Large Language Models to Enhance Accessibility of ETDs

Yinlin Chen, Virginia Tech, United States

Electronic Theses and Dissertations (ETDs) encapsulate significant research findings and innovative ideas but often have limited visibility and accessibility, particularly in regions and disciplines with restricted digital reach. This workshop introduces an LLM-based application using a Retrieval-Augmented Generation (RAG) architectural approach to address these challenges. By utilizing LLMs to translate and standardize ETD metadata and content into a user's native language, a unified vector database is established as a knowledge source for retrieving relevant information. This information is then supplied to the LLMs to generate comprehensive responses, enhancing searchability tailored to local or remote ETD collections.

This approach improves the indexing and discoverability of ETDs and ensures accessibility across linguistic boundaries. During the workshop, we will present the details of this system's components, illustrating the program workflow and the interaction dynamics between the query, retrieval, and response generation phases. Participants will learn how to integrate these technologies into their digital library systems and repositories, adapting them to various institutional needs to enhance their ETD collections' global visibility and utility.

Leveraging ORCID's Global Participation Program and Regional Consortium Approach to enhance Global ETD discoverability and reuse

Lombe Tembo, ORCID, Zambia

This workshop proposal aims to explore strategies for leveraging ORCID and Persistent Identifiers (PIDs) to enhance the discoverability and reusability of Electronic Theses and Dissertations (ETDs) on a global scale. ETDs represent a vital component of scholarly communication, providing valuable insights and contributions to research across various disciplines. However, their full potential often remains untapped due to challenges related to visibility, accessibility, and interoperability.

ORCID, providing an open registry and a unique identifier for researchers, offers a transformative solution to address these challenges by placing researchers at the center of scholarly communication. With over 6000 connected systems worldwide, ORCID provides a robust infrastructure to facilitate trusted data exchange through integrations with global scholarly systems and indexes. By adopting ORCID in ETD repositories, institutions can enhance the quality of data associated with ETDs, improve interoperability with other scholarly systems, and increase visibility and reuse on a global scale.

In addition to ORCID, other Persistent Identifiers (PIDs) play a crucial role in enhancing the discoverability of ETDs by providing unique and persistent links to scholarly output.

Through the implementation of PIDs in ETD repositories and connected systems, institutions can ensure that ETDs remain discoverable and accessible over time, thus maximizing their impact and visibility within the research community.

Furthermore, this workshop will explore the role of ORCID's Global Participation Program (GPP) and Regional Consortium initiatives to facilitate global communities of practice, fostering collaboration among diverse stakeholders that advance ETD management practices. By participating in the GPP and regional consortia, institutions can leverage collective expertise, resources, and networks to enhance the visibility and impact of ETDs on a global scale.

Through collaborative discussions, participants will gain actionable insights and strategies to effectively leverage ORCID and PIDs for ETD management and dissemination. By engaging with this workshop, participants will be empowered to maximize the impact of ETDs through PIDs within their institutions and contribute to the advancement of scholarly communication on a global scale.

This session can be part of a wider workshop relating to scholarly infrastructure and/or

PIDs, and can also be placed within the main conference presentation line-up.

Full Papers

Improving the Mkulima Repository Content: Utilizing Theses, Dissertations, and LLMs for Agricultural Knowledge Dissemination in Kiswahili

Joseph P Telemala, Sokoine University of Agriculture, Tanzania

The Sokoine National Agricultural Library (SNAL) at the Sokoine University of Agriculture (SUA) faces significant challenges in disseminating agricultural information to Swahili-speaking communities, as most research outputs are predominantly in English. This language barrier hinders the effective transmission of vital agricultural knowledge to key stakeholders in the agriculture-food value chain who use Kiswahili in their daily activities. To address this gap, SNAL established the Mkulima Collection and Repository, dedicated to collecting agricultural content in Kiswahili. Despite these efforts, the Swahili content in the repository remains limited.

This study seeks to enhance the Mkulima Repository by translating abstracts from English-language theses and dissertations using MarianMT, a machine translation (MT) model based on large language models (LLMs). The selected abstracts underwent pre-processing, machine translation, and subsequent quality assessment by multilingual experts.

Our findings reveal significant challenges in using LLMs like MarianMT for low-resource languages such as Kiswahili. While the MT system offers a rapid and scalable method for translating academic content, the accuracy and fluency of the translations were found to be suboptimal, as indicated by the evaluators. Common translation errors, particularly in agriculture-specific terminology and scientific names, highlight the limitations of current MT models in handling specialized agricultural content. These issues underscore the need for a more refined approach, including the development of a curated dataset of Swahili-English pairs that focus on agricultural jargon and the integration of a knowledge base to address the translation of scientific terms.

E-Theses and Dissertations in Zambia: A Case Study of Two Universities in Kabwe

Jive Lubbungu, Kwame Nkrumah University, Zambia

The Sokoine National Agricultural Library (SNAL) at the Sokoine University of Agriculture (SUA) faces significant challenges in disseminating agricultural information to Swahili-speaking communities, as most research outputs are predominantly in English. This language barrier hinders the effective transmission of vital agricultural knowledge to key stakeholders in the agriculture-food value chain who use Kiswahili in their daily activities. To address this gap, SNAL established the Mkulima Collection and Repository, dedicated to collecting agricultural content in Kiswahili. Despite these efforts, the Swahili content in the repository remains limited.

This study seeks to enhance the Mkulima Repository by translating abstracts from English-language theses and dissertations using MarianMT, a machine translation (MT) model based on large language models (LLMs). The selected abstracts underwent pre-processing, machine translation, and subsequent quality assessment by multilingual experts.

Our findings reveal significant challenges in using LLMs like MarianMT for low-resource languages such as Kiswahili. While the MT system offers a rapid and scalable method for translating academic content, the accuracy and fluency of the translations were found to be suboptimal, as indicated by the evaluators. Common translation errors, particularly in agriculture-specific terminology and scientific names, highlight the limitations of current MT models in handling specialized agricultural content. These issues underscore the need for a more refined approach, including the development of a curated dataset of Swahili-English pairs that focus on agricultural jargon and the integration of a knowledge base to address the translation of scientific terms.

Designing a plan for sharing ETD among the University Libraries in Bangladesh

Zillur Rahman, Ahsanullah University Of Science and Technology, Bangladesh

Purpose – ETDs have a great impact on exploring higher education in all higher educational institutions of the world. Sharing of ETDs among the universities can prevent repetition of research and guide new research direction. It assures in averting plagiarism, controlling theft and copyright of one's intellectual property. Hence, the purpose of this research is to design a plan for sharing ETD among the University Libraries in Bangladesh within the existing consortium and networking and resource sharing framework.

Design/methodology/approach – This study uses both qualitative and quantitative approaches along with review of related literature. This paper begins with an overview of consortium and networking and resource sharing framework currently being employed by the university libraries of Bangladesh in electronic and digital platforms. It moved into a discussion of necessity of consortium and networking and resource sharing among the libraries as a whole and particularly in higher education. It closes by stressing a suitable model plan for developing countries like Bangladesh. Data has been collected through survey questionnaire. The conceptual and textual information related to the present study have been collected from primary and secondary sources of information such as books, journals, magazines, newspapers, conference proceedings, official documents, and unpublished sources. Websites of the sampled University Libraries have also been used for collecting information. Literature has been reviewed and extraction has been presented in the form of figure and tables. After processing and analysis of the data, appropriate physical meaning and interpretation to the numerical results in real life given for each of the Table and Figures.

Findings – The study reveals that ETDs have a great role in disseminating knowledge among the academic communities. ETDs can be shared and impart in academic community locating any corner of the world. In such a paradox, no standard University Consortium has been built in Bangladesh since the independence in 1971. Several initiatives were taken by different organization including University Grants Commission (UGC), BANSDOC and universities individually but failed due to administrative skills and national guidelines. In case of university library consortium the situation is worse than other research organizations. There was no library consortium in Bangladesh before 2007. In 2006, an initiative was taken by the UGC of Bangladesh to form a Digital Resource Consortium for university libraries in Bangladesh for sharing integrated library resources including e-resources and computer database (Uddin & Chowdhury, 2006, p. 490-3, Uddin, 2009, p. 196). In 2007, one consortium formed in Bangladesh was named Bangladesh INASP-PERI Consortium (BIPC), presently LiCoB with the participation of major public and private universities and a few research institutions under the guidance

and supervision of Bangladesh Academy of Sciences (Uddin, 2009, p.196). Digital Archive on Agricultural Theses and Journal (DAATJ) found only established network in Bangladesh. In 2009 UGC has formed UDL consortium under HEQEP to provide electronic resources among the universities with the financial support of World Bank.

Research limitations/implications – There are around 172 universities in Bangladesh now. Data collections from these universities are important to come in any concrete conclusion.

The study depends on questionnaire on selected public and private universities only.

Practical implications – This study divulges a simple and cost effective Model Plan to implement within the existing framework of resource sharing to support in the digital environment. The research findings will direct the students of Library and Information Science, Faculty Members, and Policy Makers in further decision making and research.

Implementing Persistent Identifier Infrastructure for Effective Management of ETD Repositories: A Case Study from Chartered Institute of Personnel Management, Sri Lanka

Kamani Perera, Chartered Institute of Personnel Management, Sri Lanka

Introduction: Electronic Theses and Dissertations (ETDs) have become integral components of scholarly communication, offering valuable insights into academic research and contributing to the advancement of knowledge across diverse fields. As the digital landscape continues to evolve, the management and accessibility of ETDs present both opportunities and challenges for academic institutions, libraries, and researchers alike. In this context, it is important to establish a vibrant infrastructure for persistent identifiers (PIDs) to proper management of ETDs. Persistent identifiers are essential tools that enable the unique and unambiguous identification of digital objects, ensuring their long-term discoverability, accessibility, and citability. In the context of ETD repositories, implementing a comprehensive PID infrastructure is paramount to streamline access, enhance interoperability, and facilitate scholarly communication. In this research study, it is described the significance of implementing a persistent identifier infrastructure for the effective management of ETD repositories. By examining the role of PIDs in improving discoverability, ensuring integrity, and fostering collaboration within the scholarly community, this study seeks to provide insights into best practices and strategies for implementing PID systems tailored to the unique needs of ETD repositories.

Objectives 1.To enhance discoverability 2.To improve accessibility and long-term preservation 3. To foster interoperability and integration among different ETD repositories 4.To enable accurate and reliable citation of ETDs

Methodology: A comprehensive review of existing literature, scholarly articles, reports, and best practices related to persistent identifiers, electronic theses and dissertations, repository management, and digital preservation were conducted. This review provides a foundational understanding of the current state-of-the-art, challenges, and opportunities in implementing persistent identifier infrastructure for ETD repositories. Various persistent identifier systems and frameworks such as DOI (Digital Object Identifier) and ARK (Archival Resource Key) were evaluated to identify the most suitable system for the repository.

Results and Conclusion: Implementation of persistent identifier infrastructure led to a significant increase in the discoverability of ETDs within repository systems and external databases. By assigning unique identifiers to each document, ETDs became more easily searchable and identifiable, thereby improving their visibility among scholarly community. Further, it ensured the long-term accessibility and preservation of ETDs. Stable links provided by persistent identifiers enabled continued access to ETDs over time, even as

technologies evolved and platforms changed. In this context, ETDs remained accessible to future generations. It promoted collaboration and knowledge sharing across institutional boundaries, facilitating interdisciplinary research and fostering a more interconnected scholarly ecosystem. Persistent identifiers streamlined the citation process for ETDs, promoting accuracy, reliability, and proper attribution to authors. Stable references provided by persistent identifiers were easily included in scholarly publications, citations, and bibliographies, supporting transparency, reproducibility, and scholarly communication.

Conclusion: The implementation of persistent identifier infrastructure has proven to be instrumental in enhancing the management of ETD repositories. By providing unique PIDs for ETDs, this infrastructure has improved the discoverability, accessibility, interoperability, and citability of scholarly works, thereby facilitating broader dissemination of knowledge and fostering collaboration within the academic community. PIDs support the long-term preservation and accessibility of valuable research outputs for the benefit of future generations.

Empowering HRM Professionals: Advancing Research Culture with ETDs in The Chartered Institute of Personnel Management (CIPM), Sri Lanka

Kamani Perera, Chartered Institute Of Personnel Management, Sri Lanka

Introduction: The Chartered Institute of Personnel Management (CIPM) plays a key role in the professional development of Human Resource Management (HRM) practitioners.

To enhance this development, there is a growing recognition of the importance of fostering a culture of advanced research within the HRM community. This study focuses on the utilization of Electronic Theses and Dissertations (ETDs) as a means to empower

HRM professionals in CIPM, enabling them to engage in rigorous research practices. Objectives: •To assess the current state of research culture among HRM professionals in CIPM •To identify the potential benefits of ETDs in advancing the research culture, and •To provide recommendations for integrating ETDs effectively into the professional development framework in the field of HRM

Methodology: A mixed-methods approach was employed, starting with a comprehensive literature review to establish the foundation of research culture, the significance of ETDs, and their potential impact on HRM practitioners. This was followed by a survey distributed among a sample of HRM professionals within CIPM to gather insights into their awareness, usage, and perceptions of ETDs. Additionally, qualitative interviews were conducted with HRM leaders to delve deeper into the challenges and opportunities associated with integrating ETDs into professional development.

Results: The findings reveal a varied kind of utilization of ETDs among HRM professionals in CIPM. While there is a general awareness of ETDs, their full potential in enhancing research skills and knowledge dissemination is yet to be realized. Challenges such as limited access to ETD repositories and inadequate training opportunities were identified. However, respondents expressed enthusiasm for incorporating ETDs into their professional development, higher education activities recognizing the value they hold in expanding their research capabilities and staying updated with current HRM trends. Conclusions: In conclusion, this study highlights the importance of advancing research culture among HRM professionals in CIPM through the effective utilization of ETDs. Recommendations include developing tailored training programs on ETD usage, improving access to ETD repositories, and nurturing collaboration between HRM professionals to share research findings. By embracing ETDs, HRM professionals can not only enhance their individual skills but also contribute to the overall knowledge base within the field. This research serves as a call to action for CIPM and similar organizations to prioritize the integration of ETDs into their professional development strategies, ultimately empowering HRM professionals to excel in their roles and drive innovation in the field for the sustainable economy of the country.

Characterization of Scientific Production on Electronic Theses and Dissertations Based on a Bibliometric Analysis

Juliana Sousa, Instituto Brasileiro de Informação em Ciência e Tecnologia, Brazil

The text presents the importance of expanding the dissemination of electronic theses and dissertations (ETDs). Given this importance, the study proposes a bibliometric analysis of scientific production that focuses on ETDs. The time frame used for the analysis spans from 2002 to 2023. To assist in data retrieval, the Publish or Perish tool was employed. Consequently, 601 documents were identified, of which 61 were selected for detailed analysis. The results indicate a higher maturity in scientific publications in English, which constitute the majority of these publications. Given these findings, it is possible to infer that there are still gaps in the topic to be addressed, which could lead to greater incentives for further studies in this area.

ETD-MS v2.0: A Proposed Extended Standard for Metadata of Electronic Theses and Dissertations

Lamia Salsabil, Old Dominion University, USA

The growth of Electronic Theses and Dissertations (ETDs) in academic repositories requires comprehensive and robust schemas for compliance with the FAIR (Findability, Accessibility, Interoperability, and Reusability) principles. Dublin Core and ETD-MS v1.1 were established as the metadata standards for general scholarly documents and ETDs. We identified several gaps between the existing schemas and the need to represent ETDs comprehensively toward a better digital service. The content-level data, including objects comprising ETDs, become increasingly crucial to facilitate the development of machine learning models to mine scientific knowledge from ETDs, and scholarly big data services in general. By organizing content-level data into a new schema, this paper addresses the critical need for enhancing the expressiveness and depth of metadata for ETDs. The schema proposed includes a Core Component building on the existing ETD-MS v1.1 schema, and an Extended Component that captures objects, their provenance, and user interactions for ETDs. The schema covers 28 entities with a total of 160 metadata fields. To demonstrate applicability, we implemented the schema using MySQL and populated it with data derived from 1,000 ETDs collected from U.S. university libraries. This work provides a comprehensive and flexible approach that addresses the limitations of existing standards by enabling the description of content-level data, laying the groundwork for integrating advanced AI techniques into academic repositories.

Nurturing Advanced Research Culture among Medical Practitioners through ETDs: A case study from University of Kelaniya, Sri Lanka

Lanka Ranaweera, Faculty Of Medicine, University Of Kelaniya, Sri Lanka

Introduction: In the context of the Faculty of Medicine at the University of Kelaniya, Sri Lanka, electronic theses and dissertations (ETDs) play a pivotal role in fostering an advanced research culture among medical practitioners. This study delineates the multifaceted contributions of ETDs across various academic programs, including the Master of Public Health (MPH), BSc Degree in Speech and Hearing Sciences, BSc Degree in Occupational Therapy, and PhD dissertations from the Molecular Medicine Unit.

Purpose of the Present Study: The aims of the present study were to highlight the significance of ETDs in advancing medical research data and showcase the diverse applications of ETDs across different academic programs at the Faculty of Medicine.

Methodology: A comprehensive review of ETDs produced by students across various academic programs within the Faculty of Medicine at the University of Kelaniya was conducted. The review encompassed ETDs from the MPH program, BSc Degree programs in Speech and Hearing Sciences and Occupational Therapy, and PhD dissertations from the Molecular Medicine Unit. Data were analyzed to identify key themes, contributions, and impacts of ETDs on healthcare services and research activities. The Master of Public Health (MPH) program equips students with evidence-based approaches to promote health and prevent diseases, with ETDs contributing substantially to evidence-based public health interventions and policy formulation. PhD dissertations from the Molecular Medicine Unit provide consultancy services in Molecular Diagnosis and DNA typing and engage in research on infectious/vector-borne diseases and genetic diseases, thereby enhancing healthcare services in Sri Lanka. Importantly, it was noticed that the students of the BSc Degree program in Speech and Hearing Sciences produce compact discs (CDs) rather than ETDs. The BSc Degree program in Occupational Therapy commenced only two years ago and has not yet produced any theses.

Discussion: The findings from this research highlight the significant contributions of various academic programs to the field of public health and healthcare services in Sri Lanka. Each program plays a crucial role in advancing health-related knowledge, practices, and policies. The MPH program emphasizes evidence-based approaches to health promotion and disease prevention, equipping students with critical skills. Their ETDs contribute substantially to public health interventions and policy formulation, enhancing the effectiveness of interventions and policies at both local and national levels. PhD dissertations from the Molecular Medicine Unit illustrate the program's pivotal role

in advancing healthcare services through specialized research and consultancy, directly enhancing healthcare delivery. In contrast, the BSc Degree program in Speech and Hearing Sciences has been producing CDs rather than ETDs, hindering accessibility and dissemination of research findings. Transitioning to ETDs could facilitate interdisciplinary research and collaboration. The BSc Degree program in Occupational Therapy, established only two years ago, has not yet produced any theses. Adopting the practice of producing ETDs from the outset would ensure that student research contributes to the global body of knowledge in occupational therapy and aids in developing evidence-based practices. ETDs serve as catalysts in identifying potential opportunities for collaborative research endeavors, driving innovative solutions despite financial constraints.

Conclusion:It is imperative to encourage the initiation of ETD production for newly established degree programs that have not yet commenced research activities. Transitioning from CDs to ETDs is highly recommended to amplify the impact and accessibility of student research outputs. The widespread adoption of ETDs in all Faculty of Medicine programs is crucial for promoting interdisciplinary research, fostering partnerships, and innovating health care solution.

Future-Proofing Research by Long-term ETD Preservation: Challenges and Opportunities

Shahzeb Hasan., Akal University, India

This paper examines the challenges and opportunities in preserving Electronic Theses and Dissertations (ETDs). It highlights the significance of ETDs in academia and public knowledge while addressing technological hurdles such as digital obsolescence, data integrity, cybersecurity threats, and infrastructure limitations. The paper explores technological advancements like digital preservation tools, open-source platforms, cloud storage, and automation that can mitigate these challenges. Organizational and policy issues, including institutional policies, funding, legal considerations, and the need for collaboration, are also discussed. The study advocates for comprehensive preservation policies, strategic funding, enhanced legal frameworks, and strengthened institutional collaboration. It calls for a proactive approach to ensure ETDs' long-term accessibility and reliability, urging stakeholders to prioritize and invest in preservation efforts to safeguard these critical academic resources for future generations.

Do More Complete Dissertations' Metadata Get More Engagement?

Brenda Van Wyk, Iranian Research Institute For Information Science And Technology (IranDoc), Iran

This study investigates the role of metadata quality in Electronic Theses and Dissertations (ETDs), focusing on its completeness and its impact on discoverability and user engagement within institutional repositories (IRs). Using DSpace@MIT as a case study, the current research analyzed 22,276 doctoral dissertations to assess metadata completeness and its correlation with the number of views and downloads. Various metadata fields and usage statistics were extracted for detailed analysis. The study identified a moderate positive correlation between the number of unique metadata fields and both the Department Views Ratio (DVR) and Department Download Ratio (DDR), suggesting that enhanced metadata can improve the visibility and accessibility of dissertations. Additionally, the length of abstracts is positively correlated with engagement metrics. In contrast, title length does not significantly influence the visibility.

These findings showed the importance of high-quality metadata in enhancing the discoverability of ETDs.

Characterization of scientific production on electronic theses and dissertations based on a bibliometric analysis

Juliana Sousa, I

n

stituto Brasileiro de Informação em Ciência e Tecnologia, Brazil The text presents the importance of expanding the dissemination of electronic theses and dissertations (ETD). Given this importance, the study proposes a bibliometric analysis of scientific production that focuses on ETD. The time frame used for the analysis was made between 2002 and 2023. To assist in data recovery, the Publish or Perish tool was used. Therefore, the result was the identification of 601 documents, of which 61 were selected for analysis. There was a greater maturity in scientific publications in English, with the majority of publications. Given this, it is possible to infer that there are still gaps in the topic to be filled, which could lead to greater incentives for studies in the area.

Extracting and Registering References to Improve Scholarly Impact of ETDs

Mark E. Phillips, University of North Texas, USA

Background: The UNT Libraries began registering ETDs with Crossref to assign Digital Object Identifiers (DOIs) to the documents submitted in 2023. This decision was made to position our students' ETDs in the best light and to collaborate better in the scholarly communication landscape. Assigning DOIs to ETDs is not a new activity for many institutions and has been a major way of providing persistent linking to these resources and improving their discoverability. The assignment of DOIs introduces the option to register references to the scholarly resources cited by the work receiving the DOI in the Crossref platform. Doing so allows for Crossref to display interconnections of cited works and improves their scholarly impact. While assigning DOIs to ETDs is common, registering references for these resources is much less common. Currently, there are over 660,000 theses and dissertations registered with Crossref. Of these, just 1,558 (less than 0.24%) have references registered with their DOI, and those records come from only 19 institutions. While there have been many previous research projects into the automatic identification and extraction of references in scholarly publications—including ETDs—this research does not seem to have made its way into ETD workflows for most institutions tasked with their long-term hosting. One likely reason is that references in ETDs can take many formats and occur in different places within the document. These references may be arranged using different citation styles. In some cases, a single ETD could exhibit multiple reference sections with multiple reference formats (e.g., alphabetical vs. numbered) and multiple citation styles. This makes automated extraction of references challenging.

Objectives: This paper presents work by the UNT Libraries to build reference lists for ETDs created by students at UNT for two main purposes. First, these lists provide references to Crossref to support additional visibility and functionality related to DOIs created for ETDs in the UNT collections. Second, they provide a building block for local collection analysis regarding high-level usage of library resources purchased for our students and faculty.

Methods: In order to register our references with Crossref, we decided to create a human-curated workflow to extract text references from ETDs. First, we established a workflow with documentation for student assistants in the libraries to identify and copy references from the submitted PDF documents. Next, we adopted a simple text format to serialize each text reference from the ETD. We developed simple normalization and transformation scripts in Python to modify the copied text reference so that it is presented in a standard format. Finally, these text-based reference lists are added to a public Git repository on Github.com, where they can be used as a starting point for

different analyses and data mining.

Results: In 2023, 410 ETDs were added to the UNT Libraries collection, containing nearly 48,000 references. Once we had these references in a consistent format, we were able to leverage the Crossref API to register these references with our DOIs. The goal for including these references is to increase the scholarly impact of our students' theses and dissertations while also contributing to the global scholarly communication landscape in a new way.

Conclusion: This paper presents a background of automated methods for reference extraction in ETDs and also discusses the benefits and challenges of building reference lists for these documents in our collections. It presents the workflow that has been piloted at the UNT Libraries to extract and normalize references for the 410 ETDs submitted in 2023, along with the costs associated with human efforts to prepare these reference lists. The goal of the paper is to provide a clear set of steps and methods that other institutions could replicate without complicated technologies or processes.

Examining the Cultural and Institutional Factors Impacting ETD Visibility in Zambia: Policy and Practice Implications

Mpundu Chilonga, *The University of Zambia, Zambia*

This study delves into the cultural and institutional factors influencing the visibility of Electronic Theses and Dissertations (ETDs) in Zambia and their implications for policy and practice. Employing a mixed-methods research design comprising surveys, interviews, and document analysis, the research seeks to illuminate the multifaceted landscape surrounding ETD accessibility. The findings of this study unveil a nuanced understanding of the challenges and opportunities in the Zambian context. They reveal a complex interplay of cultural attitudes towards digital scholarship, institutional infrastructures, technological literacy, and policy frameworks shaping the accessibility of ETDs.

Furthermore, the study uncovers significant barriers hindering the dissemination and utilization of ETDs beyond academia. These barriers include limited digital infrastructure, inadequate institutional support, and prevailing cultural norms favouring traditional forms of knowledge dissemination. The implications of these findings underscore the need for targeted policy interventions and institutional reforms aimed at fostering a culture of open access and digital scholarship in Zambia. By addressing these cultural and institutional challenges, Zambia can harness the full potential of ETDs as valuable resources for education, research, and socioeconomic development. **KEYWORDS:** Digital Repositories, Resource accessibility, Research visibility

The changing landscape in research data management in Kenya's Universities: an analysis of development and implementation

Rotich, Kenneth K , Egerton university, Kenya

There is move towards Open Science with the goal of making research data accessible and reusable by the research community is a growing global trend. Research data is key pillar to Open Science. Research institutions are generating research data in huge magnitude which requires well planned and thought out strategies for ensuring that they are available when needed. This has resulted in implementation of research data management (RDM) strategies as critical component for any research organization for enhancement of quality research output. Development of policies and guidelines by Universities in Kenya is critical in implementation and provision of in Research data management services. Institutions in developed countries have already developed policies which require researchers to deposit their research data in open repositories. With the globalization this trend is finding its way to developing countries such as Kenya as evident in the development of research data management services in countries such as South Africa. In Kenya, there are few indications that point to existence of research data management services such as awareness training. There is evidently huge research activities in research institutions in Kenya, however management of research data that are acquired during the research processes seems to be largely ignored or given inadequate attention that it deserves. The impact of RDM on innovations has consequently not been documented. This begs the questions as to what extend are research data is managed in academic institutions in Kenya. These study will therefore unravel the state of development and implementation of RDM services in academic Libraries in Kenya. The analysis draws on a review of existing RDM policies at Kenyan universities, interviews with key stakeholders, and a survey of researcher awareness and practices. The findings will inform the development of a framework for robust RDM policy design and implementation in the Kenyan university context. This framework will contribute to improved research data management practices, fostering transparency, reproducibility, and ultimately, the impact of Kenyan research.

Unlocking the Potential of ETDs: Implementation of a Novel ETD Repository in Chartered Institute of Personnel Management in Sri Lanka

Kamani Perera, Chartered Institute of Personnel Management, Sri Lanka

Introduction:

Electronic Theses and Dissertations (ETDs) represent a vital and widely utilized open-access resource for scholars and researchers globally. Within higher education, these digital versions of theses and dissertations serve as crucial reservoirs of information and knowledge for forthcoming research activities. ETDs cover both electronic versions and traditional hard copies of scholarly works submitted by researchers to their respective universities or institutions. In brief, an ETD can be defined as any thesis or dissertation mainly submitted, archived, and disseminated in electronic format. The creation of ETD repositories represents a revolutionary advancement in utilizing Information and Communication Technology (ICT) to organize institutional research materials systematically, nurturing a creative platform to drive forthcoming research activities. However, in Sri Lanka, despite the growing significance of ETDs, there exists a gap in the availability and accessibility of these resources, particularly in the field of Human Resource Management (HRM). The implementation of a novel ETD repository tailored specifically for HRM research in Sri Lanka is necessary.

Objectives:

To implement a novel ETD repository to serve the scholarly community to conduct novel research. To increase the accessibility and visibility of HRM research outputs by providing a centralized platform. To enhance the research culture and academic reputation of Sri Lanka by promoting open access to high-quality HRM scholarship through the ETD repository.

Methods:

Implementing a new ETD repository has become a crucial task, and initiatives have already been taken to upload the selected ETDs from bachelor and master level scholars. DSpace software is being used to build up the repository according to customary needs. Copyright clearance has duly been taken to host the institutional ETDs on the platform. Open access ETDs, which were collected through web navigation, are uploaded to the repository based on the institutional curriculum. Initially, ETDs are uploaded to the repository in descending order starting from 2023, the latest. Anti-plagiarism software is being used to check for plagiarism issues in ETDs before uploading to the repository. Metadata can be considered an integral part of the ETD lifecycle. Thus, it is ensured to

adhere to internationally recognized metadata standards such as ETD-MS, established by the Networked Digital Library of Theses and Dissertations (NDLTD).

Results and Conclusion:

Quality ETDs are selected by the appointed academic committee. Moreover, proper guidance is provided to the newly enrolled students to encourage them to make use of the institutional repository and come up with innovative research titles for their academic assignments. Regular awareness programs are conducted to optimize the use of the ETD repository, and statistics are collected for annual evaluations; scholars can make comments for further improvement of the repository. This is the starting point of the scholarly journey by unlocking the potential of ETDs and inviting the new generation to conduct novel research using the ETD repository to find solutions for the burning questions in the field of HRM in the country. Moreover, this ETD repository links with Google Scholar and thereby generates Google Scholar Ranking for the research scholars who are the authors of the particular ETDs and Webometric Ranking for the hosting institute of the repository. It provides global visibility to HRM research. It preserves valuable resources and thereby reduces the dependence on physical copies and enables efficient search and retrieval of information. Further, ETDs contribute to environmental sustainability by decreasing paper consumption. In this context, ETD repositories create a dynamic and vibrant knowledge-sharing ecosystem that empowers the scholarly community to thrive in their careers and contribute meaningfully to building a vibrant research culture.

Landscape of Open Access Repositories with Special Reference to Electronic Theses and Dissertations (ETD) across SAARC and BRICS Nations: A Comparative Analysis

Vivek Ranjan, INFLIBNET, India

Background of the Study: Research outputs stored as softcopies and archived through open access are gaining momentum across the globe. Academic and research institutes, meticulously following research guidelines and policies, are systematically producing and storing research results in sophisticated ETDs in almost every nation. In the digital age, global visibility of research is crucial, with Electronic Theses and Dissertations (ETDs) playing a vital role. Open Access Repositories (OARs) have gained traction, led by Europe, North America, and increasingly by Asia. Key repositories like ProQuest Dissertations and Theses Global and DART-Europe are instrumental in this movement. Noteworthy ETD repositories in SAARC and BRICS nations include Shodhganga in India, maintained by the INFLIBNET Centre, the Digital Archive on Agricultural Theses and Journals in Bangladesh, the Pakistan Research Repository, and Biblioteca Digital Brasileira de Teses e Dissertações in Brazil. Starting a little later in 2010, India gradually created its repository and stands out for its substantial contributions to the ETD landscape. However, a few challenges remain in establishing comprehensive ETD repositories, given the diverse characteristics of these nations in terms of population size and higher education institutions. Structure and architecture of ETDs, submission workflows and authentication, metadata standards used, harvesting methods implemented, scalability, interoperability, and DRM issues are major concerns while comparing ETDs across the globe. In order to find commonalities and differences, a detailed analysis of ETD repositories across SAARC and BRICS nations is essential.

Objectives: 1. To compare open access ETD repositories in SAARC and BRICS countries, focusing on subject coverage and regional language contributions. 2. To investigate the software used for open access ETD repositories across SAARC and BRICS nations. 3. To analyze copyright policies relevant to open access ETD repositories across SAARC and BRICS nations. 4. To evaluate persistent metadata standards within the open access ETD repositories across SAARC and BRICS nations.

Methodology: The available open-access ETD repositories under Open ROAR/DOAR are examined to derive the sample size for analysis. The study employs a systematic approach for ETD data collection and analysis. It identifies active ETD repositories based on accessibility and content relevance. Information on software platforms, copyright policies, metadata standards, and subject coverage is gathered. Both thematic analysis and quantitative methods are used to identify patterns and trends.

Data Analysis: The analysis involves conducting thorough searches to identify active

ETD repositories, gathering repository information on software platforms, copyright policies, and metadata standards, categorizing ETDs based on subject classifications and language assessment, developing a search strategy for AI-related ETDs, and performing both thematic and quantitative analyses to discern patterns and trends.

Results: The analysis shows a varied landscape of ETD repositories across SAARC and BRICS nations. India's Shodhganga is a significant contributor, reflecting the country's leadership in this area. It also highlights differences in repository management software and metadata standards, with some countries adopting more advanced systems than others. Copyright policies vary, affecting the accessibility and use of ETDs.

Conclusion: The analysis underscores the diversity of open access ETD repositories across SAARC and BRICS nations, differing in subject coverage, language representation, software, copyright policies, metadata standards, and interoperability. By spotlighting regional languages and AI research, this study suggests a democratization of knowledge access and innovation.

A Pre-Processing Pipeline for Improved ETD Metadata Quality in Downstream Services

The University of Zambia, Zambia,

Introduction:

The Zambia National Electronic Theses and Dissertations (ZNETD) Portal serves as a crucial platform for accessing Electronic Theses and Dissertations (ETDs) from Higher Education Institutions (HEIs) in Zambia. The ZNETD functions by automatically harvesting ETD metadata from HEI Institutional Repositories (IRs) across Zambia. However, its usefulness is hindered by the inconsistently formatted and poor quality of metadata. This paper outlines the core challenges surrounding metadata quality in the ZNETD Portal and proposes the implementation of a pre-processing pipeline to address metadata quality issues.

Motivation:

There are a reported 10 IRs, out of which eight (8) of them are presently functional (Chisale & Phiri, 2023). The HEIs with functional repositories generally have different ingestions policies and practices, resulting in inconsistently formatted metadata. More importantly, existing studies have revealed that some HEIs lack IR policies (Kasonde Phiri, 2023).

Methodology:

The work presented in this paper involved two (2) primary objectives: * Empirical evaluation of inconsistencies with ETD metadata from HEI IRs—Metadata from HEIs with functional IRs was harvested using the Open Archives Initiative Protocol for Metadata Harvesting (Open Archives Initiative Protocol for Metadata Harvesting, n.d.). The metadata elements were then analysed related to the ETD-ms metadata standard.*

Design and implication of an ETD metadata preprocessing pipeline—Pre-processing pipeline scripts were implemented to format metadata consistently and, additionally, to address the issue of incomplete metadata.

Anticipated Results:

It is anticipated that the implementation of the pre-processing pipeline software tool will lead to a significant improvement in metadata quality within the Zambia National ETD Portal. Enhanced metadata consistency and completeness will facilitate more efficient resource discovery and retrieval, thereby enriching the user experience and maximising the utility of the platform. Chisale, A., Phiri, L. (2023, November 17). Towards Metadata Completeness in National ETD Portals for Improved Discoverability. <http://ir.inflibnet.ac.in/handle/1944/2446> Kasonde, C. C., Phiri, L. (2023). Assessing

and Promoting Metadata Quality for Electronic Theses and Dissertations in Institutional Repositories Using a Policy-Driven Approach. INFLIBNET Centre, Gandhinagar. Open Archives Initiative Protocol for Metadata Harvesting. (n.d.). Retrieved May 18, 2024, from <https://www.openarchives.org/pmh>

Determining the factors influencing the utilization of open source digital repository software in the preservation of ETDs in academic libraries in Bangladesh

% MDr. Md. Zillur Rahman , % Ahsanullah University of Science and Technology*

Problem/motivation/goal: This article aims to identify the variables that affect the use of open source digital repository software (DRS) for ETD preservation in academic libraries of Bangladesh. Additionally, the study examines the compliance difficulties with standards and protocols, such as metadata, WebDAV, OpenSearch, OpenURL, RSS, ATOM, OpenArchives Initiative (OAI)-PMH, OAI-ORE, SWORD, and WebDAV, for access, ingest, and export while using open source DRS. The link between technical and financial aspects of employing DRS is also covered in the study. **Methodology/approach:** The current investigation would employ a quantitative methodology. Academic librarians in Bangladesh will be given a structured questionnaire to complete in order to meet the study goals. The instrument will be shared on a number of forums, email groups, and mailing lists for libraries. There will be three portions to the instrument (A, B, and C). The respondents' demographic information is in Section A. The questions in Section B dealt with aspects of employing DRS to preserve ETDs. Questions about complaints about employing DRS will be found in Section C. The degree of each question will be determined using a 7-point Likert scale. Descriptive statistics will be used to examine the gathered data. A statistical significance test will be utilized in addition to descriptive statistics to examine the connection between technical and economic factors of using DRS. **Objectives** The main objective of the present study is to explore the factors and usage of open source DRS.

The present study attempts to ascertain the following objectives:

1. To determine the factor of using DRS in the preservation of ETD's.
 2. To know the complaints about using DRS for the preservation of ETDs.
- Research Questions**
1. What is the level of factors influencing the usage of open source DRS by academic libraries in Bangladesh?
 2. What is the relationship between the variables, multiple factors, and the usage of open source DRS?
 3. What are the complaint issues that are adhered to by the academic institutions in Bangladesh?

Anticipated results: Despite the advantages of open source DRS of capturing and ingest of ETDs, including metadata about materials, easy access to the ETDs both by listing and searching and long-term preservation of the ETDs, it has some drawback too which are- flat file and metadata structure, poor user interface, lack of scalability and extensibility, limited API, limited metadata features, limited reporting capabilities and lack of support for linked data. Overcoming all the barriers the use of open source DRS has been increasing day by day in Bangladesh. The majority of open-source DRS are

functional in nature, and their base URL is OAI-PMH. DSpace continues to be the program of choice for DRS content management. The majority of DRS do not have rules in place for content control. The majority of DRS do not provide usage data, and English continues to be a priority language for the material scattered among them. Web 2.0 tools are included in a considerable number of IRs, with RSS being the most used Web 2.0 tool. Sizable portions of IRs have not altered their user interface. The interface of the majority of IRs is bilingual. According to Crow (2002) and Rahman (2015), DRS in Bangladesh essentially began as a portion of the institution's digital repository, which is defined as "[...] a digital archive of the intellectual product created by the faculty, research staff, and students of an institution and accessible to end-users both within and outside of the institution, with few if any barriers to access." Open source software was chosen by academic libraries for their institutional repositories and comprehensive digital library platforms. In Bangladeshi libraries, computerization and digitalization have gained significant traction in the last twenty years, which is an amazing accomplishment (Rahman et al, 2015).

Posters

Creating an Electronic Thesis Database (ETD) using OMEKA. The Zimbabwe Open University experience

Nobbie Kandira, *Zimbabwe Open University, Zimbabwe*

The purpose of this study was to establish the implementation experiences of the Zimbabwe Open University librarians in creating an Electronic Thesis and Dissertation Database (ETD) using OMEKA over Dspace software. The study explored the successes and challenges that were faced when coming up with the new institutional repository. The objectives of the study were to: (i) Establish the benefits of using OMEKA over DSpace. (ii) Establish the challenges faced in implementing the new ETD database. (iii) Highlight the benefits accrued from establishing the ETD. The study was qualitative in nature and used the phenomenology paradigm. Data was collected from all Librarians and Library IT staff at the Zimbabwe Open University that were involved in the implementation of the new repository. An interview guide was drafted in order to collect responses from the target population. The study was significant in that it gave a comparison of DSpace and OMEKA as two open source software that were available on the market. Results of the study revealed that lack of trained staff was one of the major problems in creating the ETDs.

Exploring AI-Driven Strategies for Enhancing the Visibility of E-Theses in Shodhganga Repository

Nabajit Saikia, INFLIBNET

Shodhganga Repository, managed by the INFLIBNET Centre, serves as a national digital repository for Indian electronic theses and dissertations (ETDs). It plays a crucial role in the academic ecosystem by providing a platform for researchers to access an extensive collection of scholarly works, thereby contributing to academic and research excellence.

Despite its significance, Shodhganga faces several challenges in managing ETDs, including inconsistent metadata quality, manual entry errors, and a lack of comprehensive metadata for some records. With over 500,000 theses from various universities across India, continuously increasing annually and covering diverse subjects such as science, technology, humanities, and social sciences, the repository requires robust solutions to streamline its management processes and enhance user experience.

Artificial Intelligence (AI) has emerged as a transformative technology in various sectors, facilitating solutions that require human-like intelligence. This poster proposes the integration of AI in the creation and management of ETDs to address these challenges and improve ETD management in Shodhganga.

Objectives: Primary Goal: To provide a comprehensive flowchart and step-by-step implementation plan for enhancing ETD management in Shodhganga using AI, based on a literature review.

Specific Objectives: 1. Illustrate AI-driven solutions for improving metadata quality and consistency. 2. Outline steps to enhance search and retrieval functionalities. 3. Present a roadmap for increasing user engagement and satisfaction.

The implementation plan for these AI-driven solutions is divided into three phases: 1. Research and Development: A detailed literature review will identify best practices in AI-driven ETD management. Current management practices in Shodhganga will be analyzed to pinpoint areas for improvement. AI models for metadata creation, search enhancement, and recommendations will be developed and tested. 2. Pilot Testing: This phase will involve implementing these solutions on a subset of ETDs, collecting user feedback, and refining the algorithms based on initial results. 3. Full-Scale Deployment: The AI solutions will be rolled out across the entire Shodhganga repository, with continuous monitoring and iterative improvements to ensure optimal performance and adaptation to new data.

Expected Outcomes: 1. Integrating AI into Shodhganga's ETD management will streamline processes, reduce manual workload, and enhance efficiency. Improved search,

personalized recommendations, and better metadata will enhance user experience and discoverability. 2. Integration with external databases will broaden research access, making Shodhganga a more valuable resource for the academic community.

****Visual Elements for Poster**:** - Figure 1: Flowchart of AI-driven ETD Management Process. - Figure 2: Automated Metadata Extraction Flowchart. - Figure 3: Flowchart of Graph Machine Learning for Metadata Quality Enhancement. - Figure 4: Diagram Illustrating Semantic Search Implementation. - Figure 5: Diagram Depicting Personalized Recommendation System. - Figure 6: Diagram of Automated Content Tagging and Classification. - Figure 7: Metadata Standardization Process Diagram. - Figure 8: Integration of External Databases with Diagram Illustration. - Figure 9: Graph Showing the Expected Outcome for the Enhancement of Shodhganga.

References: Shodhganga, INFLIBNET Centre. Available at:
[<https://shodhganga.inflibnet.ac.in/>]

Design and Implementation of an Interoperable Zambia National Electronic Thesis and Dissertation Portal

Elijah Chileshe, *The University of Zambia, Zambia*

Introduction: Zambia boasts a significant number of Higher Learning Institutions (HEI) contributing to a vibrant academic landscape marked by prolific scholarly research output. The Higher Education Authority (HEA) reports a total of 123 HEIs as of 2024 (Higher Education Authority, 2018). A number of the 123 HEIs offer postgraduate programmes which, in part, will require students to produce a thesis or dissertation manuscripts.

Increasingly, such manuscripts are stored in electronic form, resulting in so-called Electronic Theses and Dissertations (ETDs). Existing literature reports that a total of 10 HEIs have implemented Institutional Repositories (IRs), with eight (8) of them still functional (Chisale & Phiri, 2023). Motivation: With the increase in the number of HEIs offering postgraduate programmes and, additionally, the rise in the adoption and implementation of IRs, discoverability of such scholarly outputs is compromised. In addition, assessment of the quality of ETDs by entities such as HEA is challenging. A potential solution to this problem is the implementation of a platform to provide centralised access to the ETDs.

Methodology: A national ETD portal has been implemented and primarily makes use of the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) to harvest ETD metadata from HEI IRs across Zambia. By leveraging OAI-PMH, the portal ensures efficient and standardized retrieval of metadata, enabling seamless integration of ETDs into its centralized repository. Additionally, ongoing monitoring and maintenance will be conducted to uphold the functionality and usability of the portal.

Results: It is expected that the centralised repository of ETDs will significantly improve accessibility to scholarly resources for researchers in Zambia. By providing a single platform for accessing ETDs from various institutions, researchers will be able to save time and effort in searching for relevant materials. It is expected that the portal will lead to increased visibility and recognition of research output from higher learning institutions in Zambia.

Chisale, A., & Phiri, L. (2023, November 17). Towards Metadata Completeness in National ETD Portals for Improved Discoverability. 26th International Symposium on Electronic Theses and Dissertations. ETD 2023, Gujarat, India.
<http://docs.ndltd.org/metadata/etd2023/9/index.html>

Higher Education Authority. (2018, January 4). Higher Education Authority. Higher Education Authority - Ensuring Quality in Higher Education; Higher Education Authority.
<https://hea.org.zm>

Automatic Electronic Thesis and Dissertation Guideline Verification For Consistently Formatted Manuscripts

Mubanga C. Chibesa,

The University Of Zambia

Introduction: Higher Education Institutions worldwide enforce guidelines and academic approaches to ensure scholarly integrity and adherence to academic standards (Razi et al., 2019). The University of Zambia is not an exception. Like most HEIs, it offers training to postgraduate students, and one of the key aspects of postgraduate training is producing an Electronic Thesis and Dissertation manuscript. The Directorate of Research and Graduate Studies (DRGS) at the University of Zambia provides guidelines that stipulate how ETDs should be formatted. However, the process of checking for conformance is a manual and tedious procedure, resulting in the submission of inconsistently formatted manuscripts in the Institutional Repository (IR). To address this challenge, our project seeks to implement a tool that will automate the process of checking ETD compliance against established postgraduate guidelines. The tool will leverage data mining techniques to perform these tasks. More specifically, Document Layout Analysis (DLA) (Binmakhashen and Mahmoud, 2019) will be the core approach used in the implementation. The tool will flag portions of ETD manuscripts that do not conform to established guidelines. Hence, this will help resolve the inconsistencies in the format of submitted manuscripts.

Methodology: Using a mixed methods approach, document analysis will be employed to understand the postgraduate guidelines stipulated in the "Regulations and Guidelines for Postgraduate Studies" document; content analysis will be used on randomly sampled ETDs to experimentally determine the extent of the problem. Finally, a DLA Natural Language Processing model will be developed and evaluated using standard DLA metrics such as Structure Similarity Index and Intersection over Union.

Results: This study is part of ongoing work aimed at developing an automated tool that will verify ETD manuscripts' compliance with postgraduate guidelines. Upon successful completion of this project, we anticipate a reduced workload associated with the manual checking of manuscript conformity to postgraduate guidelines, freeing up time and resources to handle other academic tasks like focusing more on the content of the manuscripts. Not only that, academic integrity will consequently improve through the standardization of ETD formats, adding to the university's reputation and credibility.

In conclusion, the development of an ETD automatic guideline verification tool presents an opportunity to enhance efficiency and promote consistency in the quality of ETDs while alleviating the challenges faced in the manual checking process, consequently reducing the workload for students, supervisors, and examiners alike.

References: - Binmakhashen, G. M., and Mahmoud, S. A. (2019). Document Layout Analysis: A Comprehensive Survey. *ACM Comput. Surv.*, 52(6), 1–36. - Mishra, B. K., and Kumar, R. (2020). Natural Language Processing in Artificial Intelligence. - Razi, S., Glendinning, I., and Foltýnek, T. (2019). Towards Consistency and Transparency in Academic Integrity. Peter Lang GmbH, Internationaler Verlag Der Wissenschaften.

Enhancing Electronic Thesis and Dissertation Visibility: A Focus on Institutional Repository Platforms and Inherent Challenges in the Nigerian Context

David, Okhakhu O, The University Of Zambia

Higher Education Institutions (HEIs) regularly publish manuscripts of academic research that provide useful insights into social, economic, and technological issues affecting society (Phiri and M'sendo, n.d.). In Electronic Theses and Dissertations (ETDs), researchers often focus on impact-driven research with the potential to inform policy direction. However, due to the large size of these ETD manuscripts, important stakeholders, such as local mainstream media outlets, find it difficult to synthesize the content of these manuscripts. A potential solution would be to leverage advances in Artificial Intelligence by using Natural Language Processing (NLP) techniques to summarize the ETDs. As Allahyari defines it, automatic text summarization is the task of producing a concise and fluent summary while preserving key information content and overall meaning. Text summarization techniques can generate snippets of scholarly research output that are concise and easily understood by non-technical persons (Allahyari et al., 2017; Ingram et al., n.d.).

Existing literature broadly categorizes automatic summarization into two classes: abstractive summarization and extractive summarization. This study seeks to understand the challenges with synthesizing ETDs and to design and implement software tools to automatically summarize ETDs, focusing on mining text data and creating tools that allow automated summarization and text modification.

Specifically, the objectives of this study are: 1. To determine how frequently research findings are reported in mainstream media. 2. To investigate challenges in synthesizing long documents (ETDs). 3. To implement summarization models for summarizing ETDs for public consumption. 4. To evaluate the summarization models.

The proposed research methodology is as follows: - Frequency of Publishing Research Findings: Manual and automatic content analysis of existing media publications will be conducted to determine the frequency of publishing research findings. - Challenges Synthesizing ETDs: Interviews will be conducted with purposively sampled journalists from randomly selected media outlets. - Design and Implementation of ETD Summarization Models: Classic text summarization techniques, such as abstractive and extractive summarization, will be employed to build the ETD summarization models. Additionally, publicly available ETDs from HEIs in Zambia will be used to construct a dataset for the study. - Evaluation of ETD Summarization Models: Standard evaluation metrics for text summarization, such as ROUGE (Johnson, n.d.), BLEU (Allahyari et al., 2017; Johnson, n.d.), BERTScore (Bhandari et al., 2020; Zhang et al., n.d.), and

METEOR (Ermakova et al., 2019; Johnson, n.d.), will be used to assess the effectiveness of the ETD summarization models. Additionally, human evaluation will be employed to determine the perceived usefulness of the ETD summarization models.

This study could provide valuable insights into the challenges of enabling the general public to engage with ETDs and how the application of NLP and automatic text summarization can help overcome some of these challenges.

Total Quality Management in ETD Repository at Chartered Institute Management (CIPM)

Kamani Perera , Chartered Institute of Personnel Management

Introduction: As the landscape of education and professional development evolves, the use of Electronic Theses and Dissertations (ETDs) has become integral to knowledge dissemination and research advancement. The Chartered Institute of Personnel Management (CIPM) recognizes the significance of its ETD repository as a valuable resource for HRM professionals. This study focuses on implementing Total Quality Management (TQM) principles in the ETD repository at CIPM to ensure efficiency, accessibility, and the highest standards of content.

Objectives: 1.To identify the areas for improvement through TQM principles 2.To assess the impact of these enhancements on user experience and knowledge dissemination 3.To make recommendations to improve the quality of ETD repositories in order to provide effective service to the CIPM stakeholders

Methodology: A comprehensive approach was employed, starting with an analysis of the existing ETD repository at CIPM to identify strengths and areas needing improvement. TQM principles were then applied, including process mapping, stakeholder engagement, and continuous improvement strategies. A survey was conducted among HRM professionals and researchers utilizing the ETD repository to gather feedback on usability, content relevance, and accessibility. Additionally, interviews were conducted with repository administrators and stakeholders to gain insights into challenges and opportunities for enhancement. Results: The assessment of the ETD repository at CIPM revealed strengths in terms of content variety and relevance to HRM professionals. However, areas for improvement were identified, including navigation complexity, search functionality, and metadata consistency. Through the application of TQM principles, enhancements were implemented, such as improved metadata tagging, streamlined search features, and user-friendly interface updates. The survey results indicated a positive reception to these enhancements, with users reporting increased satisfaction with the repository's usability and content accessibility.

Conclusions: In conclusion, this study highlights the importance of implementing TQM principles in the ETD repository at CIPM to ensure it remains a valuable and user-centric resource for HRM professionals. The enhancements made through this process have resulted in a more efficient and user-friendly repository, aligning with the institute's commitment to excellence in knowledge dissemination. Recommendations include regular audits of the repository, ongoing stakeholder engagement, and continuous improvement initiatives to uphold TQM standards. By embracing TQM in the ETD repository, CIPM can continue to support the research and professional development needs of HRM

professionals, contributing to the advancement of the field. This research serves as a model for other organizations seeking to optimize their ETD repositories through TQM principles, ultimately fostering a culture of quality and innovation in knowledge management within the HRM community.

Enhancing Electronic Thesis and Dissertation Visibility: A Focus on Institutional Repository Platforms and Inherent Challenges in the Nigerian Context

David, Okhakhu O , Lead City University, Ibadan

Electronic Theses and Dissertations (ETDs) are an essential part of academic work since they provide insightful analysis and significant contributions to the international academic community. Platforms for institutional repositories (IRs) are vital conduits for the international distribution and availability of ETDs. However, there is a notable disparity in the visibility and accessibility of ETDs due to particular difficulties with the implementation and efficacy of IR platforms in Nigerian institutions. With a focus on the difficulties Nigerian institutions confront, this paper examines how institutional repository platforms might improve the visibility of ETDs globally. The study emphasises the transformative potential of IR platforms in boosting the impact and dissemination of ETDs through a comprehensive evaluation of the literature. It also emphasises how critical it is to fill up the knowledge vacuum in the literature about the application and efficacy of IR platforms in Nigerian academic institutions. This study used a mixed-methods approach, combining qualitative investigation through stakeholder surveys and interviews with Nigerian institutions' stakeholders with quantitative analysis of IR availability and usage data. The quantitative analysis looks at IR availability, citation counts, download trends, and ETD submission rates on IR platforms to get a sense of how visible Nigerian ETDs are right now. Qualitative approaches explore the particular obstacles Nigerian institutions have in efficiently leveraging IR platforms, such as infrastructure limits, institutional support, cultural considerations, and knowledge gaps. Compiling quantitative and qualitative data allows for a more thorough understanding of the intricate relationships surrounding ETD deployment and visibility in Nigerian context. The study exposes inherent challenges to the efficient use of IR platforms. These results demonstrate how urgently customised interventions are needed to solve these issues and raise Nigerian ETDs' profile internationally. The study makes further recommendations based on its findings on the use of IR platforms in Nigerian institutions. These recommendations include making investments in technology infrastructure, starting capacity-building programmes, proposing a cultural mind shift, advocating for open access laws in policy, and encouraging stakeholder collaboration. Through implementation of these suggestions, Nigerian establishments can surmount current obstacles and utilise IR channels to augment the prominence and influence of their ETDs on the international academic scene. This study clarifies the unique difficulties encountered by Nigerian institutions while also enhancing knowledge of the function of institutional repository platforms in increasing ETD visibility. Through the implementation of focused interventions and the resolution of these issues, stakeholders can strive to promote increased inclusivity and involvement in the worldwide ecosystem of academic communication. **Keywords:** Nigerian Electronic Theses and Dissertations (ETDs),

Institutional Repository Challenges, Global Research Visibility Gap, Open Access Scholarship, Digital Divide Impact, Nigeria Universities.

Enhancing Access to Scholarly Knowledge: Strategies for Promoting Open Access ETDs in Sri Lanka

Kamani Perera, Chartered Institute of Personnel Management

Introduction: With the emergence of digital technology and the advancement of open access initiatives, the accessibility and dissemination of scholarly knowledge have undergone a significant transformation. This has been evident in Sri Lanka, as in many other parts of the world. There is a pressing need to increase access to scholarly knowledge and promote the adoption of open access principles, particularly in the realm of ETDs. ETDs play a dynamic role as repositories of treasured research output, representing the high point of academic activities. Though there are a significant number of scholarly outputs produced by universities and research institutions in Sri Lanka, access to this knowledge remains limited. Traditional bottlenecks, such as subscription fees and restricted access models, prevent the broader dissemination of research findings, hindering the possible impact of Sri Lankan scholarship on a global scale. At this juncture, the elevation of open access ETD repositories is of paramount importance to Sri Lanka. Open access facilitates the unrestricted online availability of scholarly literature, enabling anyone, anywhere, to access, read, download, and distribute research outputs freely.

Objectives: • To assess the current scenario in Sri Lanka • To develop strategies for promoting open access ETDs • To promote collaboration and capacity building

Methodology: A comprehensive review of existing literature, scholarly articles, reports, and best practices related to open access ETDs was conducted. In-depth reviews were conducted on existing policies and guidelines related to ETDs, copyright, open access, and scholarly communication at the national and institutional levels in Sri Lanka.

Results: Key barriers to the accessibility and dissemination of ETDs in Sri Lanka, including limited awareness of open access principles, inadequate technical infrastructure, and legal and copyright concerns, were identified. It is revealed that there are institutions that have already implemented open access policies and others that need support and guidance to implement the ETD repositories. Moreover, it is important to promote open access strategies such as policy development, repository infrastructure, metadata standards, and capacity building and collaboration among universities, libraries, government agencies, and international partners to support the implementation of open access ETD initiatives. Awareness programs and knowledge exchange are needed on a regular basis to build the capacity of researchers, librarians, and administrators in managing and promoting open access ETDs in both countries.

Conclusion: Open access ETDs have the potential to significantly improve the accessibility and visibility of Sri Lankan scholarly knowledge, promoting collaboration,

innovation, and knowledge exchange both domestically and internationally. Addressing bottlenecks such as technical infrastructure, legal concerns, and awareness gaps is crucial for the successful implementation of open access ETD initiatives. In the same vein, collaboration among stakeholders, including universities, libraries, government agencies, and international partners, is essential for advancing open access ETDs in Sri Lanka. It is necessary to promote a culture of open access and guarantee the sustainability of open access ETD initiatives in Sri Lanka as a developing nation.

Assessing the Efficiency of Data Science Programs in Enhancing Big Data Analysis Skills among Health Libraries and Information Scientists

Rajabu Simba, Ministry of Health

****Introduction****: The healthcare sector has seen a significant shift in data analysis methodologies due to rapid technological advancements. Data Science programs have emerged as a promising way to enhance the capabilities of professionals in health libraries and Information Science fields in managing and analyzing large datasets, known as Big Data. However, there is a research gap in assessing the effectiveness of these programs in equipping Health Libraries and Information Scientists with the necessary skills for proficient Big Data analysis. This study aims to bridge this gap by evaluating the efficacy of Data Science programs tailored for Health Libraries and Information Science professionals. The research aims to provide insights into the impact of these educational initiatives on skill development and proficiency in Big Data analysis, contributing to knowledge advancement in this evolving field.

****Aim****: To assess and evaluate the effectiveness of Data Science programs in improving the capacity and proficiency of professionals working in health libraries and Information Science. The study seeks to ascertain whether these programs adequately equip participants with the necessary skills and knowledge required for proficiently handling and analyzing Big Data in healthcare contexts.

****Method****: The study used a quantitative approach to gather data from 150 participants in health libraries and Information Science fields who completed Data Science programs in Tanzania. The survey evaluated their perceptions, skill sets, and confidence in executing Big Data analysis post-program completion. Statistical tools like descriptive statistics were used to analyze the collected data, providing insights into the effectiveness of Data Science programs in enhancing Big Data analysis skills.

****Results****: The study aimed to assess the efficacy of Data Science programs through quantitative analysis of participant responses. Among the surveyed professionals who completed the program (n=150), 82

****Conclusion****: The survey analysis shows that Data Science programs significantly improve professional skills in Big Data analysis, enhancing competence and confidence. These findings have implications for health libraries and Information Science, fostering a more agile workforce capable of utilizing advanced data analysis techniques for informed healthcare decision-making. However, limitations like survey bias and study scope need to be addressed for further improvement.

The University of Johannesburg's journey to enhance the content of the Institutional Repository (IR) and improve discoverability

Mutali M Lithole, University of Johannesburg

Institutional Repositories has been well established in most academic libraries over the past few decades. As technology developed in all spheres of life, the same development could be seen in software and technology used in the IR environment. We are all aware of software systems such as DSpace, EPrints, and Fedora, to name but a few (Castagne, 2013). The UJ library established their IR in 2008 and currently consists of research output generated by the university which includes journal articles, conference proceedings, books, and book chapters. In the last 14 years, it has been a journey during which the library used and experimented with several systems to enhance the content of the IR and to increase visibility. These systems include: DSpace, Vital, and most recently, Esploro. The purpose of this paper is to share our journey and experiences with these systems and software used over the past few years. The presentation will start with a short history of our IR journey. This will be followed by a more technical discussion, not only on the systems used to store and access the content but also the system we use to collate the relevant content. Towards the end of the presentation, the focus will be on the latest system UJ implemented, namely Exlibris Esploro, which allows us to maximize the impact of our institutional research using intelligent data capturing, thereby reducing the workload of the IR staff and also allowing increased visibility of the content as well as the authors who produced the content.

Keywords: Open access, institutional repositories, Esploro, University of Johannesburg.

Further reading:

<https://www.usetda.org/resources/institutional-repository-software/>

The Conversion of Printed Theses and Dissertations into Digital Formats: A Case Study of the University of South Africa Library

Mduduzi Ntetha, UNISA

Printed theses and dissertations in academic libraries are mainly accessible by patrons of that particular library. Converting printed copies of theses and dissertations into digital format can help eliminate the problem of losing these materials due to their age and the length of time they have been on the shelves, and also make visibility of the university's research output a reality. This study explores the process of converting printed materials into digital format in the UNISA Institutional Repository (UNISAIR) as a subsidiary of an academic library—an open access digital source that hosts the research and intellectual output produced by members of the UNISA community, including electronic theses and dissertations (ETDs). The UNISAIR is globally accessible via various internet browsers and search engines, advancing UNISA's research output and enhancing the visibility of African scholarship pertaining to ETDs.

The study employs observation as the main technique for gathering data, with an emphasis on the digitization process. This method enables a thorough and nuanced comprehension of the digitization phenomena as it occurs in real time. The key findings indicate that the process of converting printed copies of theses and dissertations into digital format can help eliminate the problem of losing these materials due to their age and the length of time they have been on the shelves. On the other hand, the process serves as a positive drive to attain the targets of Sustainable Development Goal 4 by providing ease of use for the international community who may want to access these materials.

In conclusion, the originality of the study provides valuable insights into the process of converting printed materials into digital format in the UNISAIR, highlighting the potential benefits of this process for both preservation and accessibility. Specifically, the study has shown that digitizing printed copies of theses and dissertations plays a vital role in eliminating the risk of losing these materials due to age and shelf time while also enhancing the visibility of the university's research output and African scholarship. Based on the findings of the study, it is recommended that academic libraries consider implementing digitization programs for their printed collections, particularly for materials that are at risk of being lost due to age or other factors. This can help ensure the long-term preservation of these materials while also making them more accessible to a wider audience.

Keywords: Institutional Repository, ETD, open access, academic library

Global Visibility of National ETD Repositories of G20 Countries: Comparative Studies with Respect to NDLTD's Meta Repository

Sukanta Kumar Patra, Vidyasagar College for Women

ETDs are primary information sources that originate from doctoral theses or dissertations submitted to universities for the doctoral award. The FAIR ETDs are those electronic theses and dissertations that are findable, accessible, interoperable, and reusable (FAIR).

The study covers ETD initiatives by G20 member countries, comprising 19 countries (Argentina, Australia, Brazil, Canada, China, France, Germany, India, Indonesia, Italy, Japan, Republic of Korea, Mexico, Russia, Saudi Arabia, South Africa, Turkey, United Kingdom, and United States) and the European Union. The G20 member countries represent around 85

Managing ETDs in Rajshahi University Central Library: Problems and Possibilities

M.A.M. Mominur Rahman, University of Rajshahi, India

Access to Electronic Theses and Dissertations (ETDs) is one of the recently introduced facilities in academic libraries. Obviously, there must be a digital library system to manage the ETDs. Rajshahi University Central Library (RUCL) introduced a digital system in the year 2016. It has been trying to establish an ETDs repository and its system since then. The ETDs repository in RUCL is at its primary stage now. It is managed by DSpace software and is developing day by day. However, it is simultaneously facing some problems. In this article, we will discuss the problems and possibilities of ETD management in RUCL.

Keywords: ETDs Service, Digital Libraries, Institutional Repositories, Theses, Dissertation.

Invited Talks

Resilience and ETD Repositories in Poor Countries

Hussein Suleman, University of Cape Town, South Africa

ETD repositories are quite common in countries and universities around the world and are often among the first digital repositories to be created. In poor countries, however, ETD repositories face many challenges because of the interplay between socio-economic and technical factors. Some past African ETD repositories no longer exist, with the theses they contained being no longer accessible online. This talk will focus on approaches to mitigate this risk by designing repositories that are resilient and robust in the face of disaster. Many years of experiments with principled design and simple architectures has demonstrated that the inherent complexity in modern repository tools can be avoided altogether. These alternative technical solutions can address how we build systems to safeguard our intellectual assets in the future, while also informing how the whole world designs resilient systems.

Enhancing Global Visibility and Academic Impact through Times Higher Education Rankings

¹ **Mosebjadi Petje**,

¹Southern Africa, Times Higher Education

The pursuit of global visibility and academic impact is more crucial than ever for universities in an increasingly interconnected world. Times Higher Education (THE) rankings provide institutions with a valuable framework to measure and enhance their standing on the world stage. This presentation introduces the unique value of THE rankings, focusing on how they offer insights that help universities attract top talent, foster research partnerships, and achieve greater institutional visibility. THE's rankings span various categories, each designed to meet distinct institutional goals. The World University Rankings assess universities globally on critical performance indicators, helping institutions benchmark themselves against leading universities. The Impact Rankings, aligned with the UN Sustainable Development Goals, highlight an institution's societal impact and sustainability efforts. We also have regional rankings that cater specifically to regional aspirations, supporting African universities in gaining international recognition. Participation in these rankings not only brings visibility but also provides universities with data-driven insights that highlight their strengths and areas for development. This data can be a powerful tool for strategic planning and identifying growth opportunities.

Additionally, Times Higher Education offers consultancy and data solutions that assist universities in achieving their unique goals. By engaging with THE rankings, universities in Zambia and across Africa can leverage global benchmarking to amplify their academic influence, foster new partnerships, and elevate their global standing. I invite institutions here to consider how these insights and rankings can serve as a pathway to enhancing their global academic impact.

List of Participants

Delegate Name	Organisation	Country
Andy Neale	Clarivate	United Kingdom
Austin Mclean	ProQuest, a part of Clarivate	United States
Benaiah Akombwa	University of Zambia	Zambia
Bertrand Thomas	Abes	France
Birbal Boniface Musoba	Higher Education Authority	Zambia
Boniface Banda	The University of Zambia	Zambia
Brenda Van Wyk	University of Pretoria	South Africa
Cecilia Kasonde	Kwame Nkrumah University	Zambia
Charles Greenberg	Networked Digital Library of Theses and Dissertations	United States
Denny Nsokolo	Higher Education Authority	Zambia
Dokowe Tembo	ZCAS University	Zambia
Douglas Mwale	St Bonaventure University	Zambia
Dr Beatrice May Banda	Chipata College of Nursing and Midwifery	Zambia
Edward Fox	Networked Digital Library of Theses and Dissertations (NDLTD)	United States
Efraim Dumeni	Namibia University of Science and Technology	Namibia
Elijah Chileshe	The University of Zambia	Zambia
Elizabeth Namonje	Higher Education Authority	Zambia
Ernest Banda	Zynle Technologies Limited	Zambia
Fabian Kakana	University of Zambia	Zambia
Fitzwell Kabunda	Bridging Gap Solution Limited	Zambia
Francina Makondo	University of Zambia Library	Zambia
Francis Kawesha	Higher Education Authority	Zambia
Gabriela Mejias	DataCite	Germany
Gift Muwele	The University of Zambia	Zambia
Godfrey Marera	St. Bonaventure University	Zambia
Hope Chinganya	CIDRZ	Zambia
Hussein Suleman	University of Cape Town	South Africa
Johanssen Obanda	Crossref	Kenya
Joseph Walusiku	Muyunda Centre for Infectious Disease Research in Zambia	Zambia

Joseph Telemala	Sokoine University of Agriculture	Tanzania
Josiline Chigwada	University of South Africa	South Africa
Kabuku Mwandamena	Chipata College of Nursing and Mid-wifery	Zambia
Kabwe Kasoma	CIDRZ	Zambia
Kamani Perera	Chartered Institute of Personnel Management	Sri Lanka
Kaoma Lamba Daka	University of Zambia	Zambia
Kazhila C. Chinsemu	Higher Education Authority	Zambia
Kelvin Chibomba (Phd)	Information and Communications zuniversity	Zambia
Kudakwashe Siziva	Datacite	South Africa
Lamia Salsabil	Old Dominion University	United States
Lazarus Matizirofa Banda	University of the Witwatersrand, Johannesburg	South Africa
Lewis Cheelo	Zambia Air Force (ZAF)	Zambia
Lighton Phiri	University of Zambia	Zambia
Lombe Tembo	ORCID	Zambia
Manoj Kumar K	INFLIBNET Centre	India
Mark Phillips	University of North Texas	United States
Martin Mushumba	Higher Education Authority	Zambia
Martin Nkumbula	Zynle Technologies Limited	Zambia
Marvin Kabubi	Information and Communications University	Zambia
Matamela Michael Mulaudzi	University of the Witwatersrand, Johannesburg	South Africa
Matildah Mercy Muchinga	Lusaka Apex Medical University	Zambia
Mduduzi Aubrey Ntetha	University of South Africa	South Africa
Melusi Ncube	Bridging Gap Solutions	Zambia
Mercy Banda	St. Bonaventure University	Zambia
Mosebjadi (Mose) Petje	Times Higher Education	South Africa
Moses Makungu	Zambia Air Force Center of Advance Learning	Zambia
Moses Mupeta	Information and communications university	Zambia
Mpundu Chilonga	KWAME NKRUMAH UNIVERSITY	Zambia
Mubanga Chibesa	The University Of Zambia	Zambia
Olatunbosun Obileye	DataCite	Nigeria
Olivier Cian	ABES	France

Phyela Mbewe	University of Zambia	Zambia
Shahzeb Hasan	Akal University	India
Thabiso Mwiinga-Mwanachingwala	University of Zambia	Zambia
Vivek Ranjan	Silver Oak University	India
Wendel Fabian Chinsamy	DataCite	South Africa
Yinlin Chen	Virginia Tech	United States
Zachary Zulu	THE UNIVERSITY OF ZAMBIA	Zambia

Conference Proceedings

Posters

Automatic Electronic Thesis and Dissertation Guideline Verification For Consistently Formatted Manuscripts

**Mubanga C. Chibesa¹, Albertina Mooka¹, Gift Muwele¹, Lwiime Shansonga¹,
Lighton Phiri¹**

¹ The University Of Zambia

Abstract:

Higher Education Institutions worldwide enforce guidelines and academic approaches to ensure scholarly integrity and adherence to academic standards (Razi et al., 2019). The University of Zambia, is not an exception. Just like most HEIs it offers training to postgraduate students and one of the key aspects of postgraduate training is producing an Electronic Thesis and Dissertation manuscript. The Directorate of Research and Graduate Studies (DRGS) provides guidelines which stipulate how ETD's should be formatted. Successful graduation is dependent on postgraduate students' manuscripts of which their conformance to the guidelines is a key aspect. Examiners and other relevant authorities are expected to verify and check if postgraduate students' manuscripts conform to the guidelines. However, the process of checking for conformance is a manual and tedious procedure, resulting in submission of inconsistently formatted manuscripts. Although this research seeks to ascertain the exact reasons why this is so, it is apparent that the primary task for the examiners is approving the content written in these manuscripts, and the secondary task being making sure that each and every manuscript is in conformance to the postgraduate guidelines and this is arguably an intense exercise. To address this challenge our project seeks to implement a tool that will automate the process of checking ETD's compliance against established postgraduate guidelines. The tool will leverage data mining techniques to perform these tasks. More specifically, document layout analysis (DLA) (Binmakhashen & Mahmoud, 2019). The tool will flag off portions of ETD manuscripts that do not conform to established guidelines. Hence, this will help resolve the inconsistencies in the format of submitted manuscripts.

Methodology:

We will employ a mixed method approach combining both qualitative and quantitative methods for data collection and analysis. Primary data is being collected through structured interviews and questionnaires, while secondary data involves document analysis of postgraduate guidelines and archived historical ETDs. Convenient sampling and purposive sampling will be used to select participants, the selection is based on judgement and their willingness to participate in the study. The study is targeted on postgraduate examiners, alumni and current postgraduate students. Document analysis will be employed to understand the postgraduate guidelines stipulated in the (The University of

Zambia, 2015) document and the identified guidelines are to be categorised into themes and a checklist will be prepared which will be used as a measurement instrument and each guideline, will be set as a question and each question will be given a measurement scale of 1-5; content analysis will be used on randomly sampled ETDs in order to experimentally determine the extent of the problem using Park's metrics of metadata quality (accuracy, consistency and completeness) and, finally, a DLA Natural Language Processing model will be implemented and will be evaluated using standard DLA metrics such as Structure Similarity Index and Intersection over Union. This DLA pipeline is hinged on Artificial intelligence and Natural language processing(NLP)(Mishra & Kumar, 2020). We also leveraged an open source package called Deepdoctetion for the implementation of the software tool. For the development process, Agile methodologies, specifically scrum, will be used for software development, with document layout techniques applied to enhance the Deepdoctetion package in the process. We also accessed the accuracy and usability of the tool through a benchmark dataset composed of historical ETDs archived on The

University of Zambia institutional repository. Measurement instruments include structured interview guides and a checklist to evaluate the threshold of compliance of historical ETDs. Preliminary Findings and Discussion Empirical Analysis of Historical Archived ETDs The analysis of historical archived manuscripts against postgraduate guidelines revealed that most manuscripts did not adhere to the guidelines. The preliminary findings indicate that the majority of the students from all the schools at the institution performed poorly on sections such as the copyright declaration page which had the lowest compliance, with much lower scores in accuracy (2.227), consistency (2.194), and completeness (2.037). This suggests that this section of the manuscripts frequently falls short of meeting the required standards, highlighting a critical area for improvement.

This is followed by the declaration page with the second lowest scores in accuracy (3.042), consistency (2.681) and completeness (2.835). The compliance to miscellaneous requirements also had some low scores in accuracy (3.774), consistency (3.760) and completeness (3.841). Overall middle-ranking sections include the title Page with the following compliance scores in accuracy (4.078), consistency (3.861) and completeness (4.023). Whereas compliance on the certificate of approval indicates the following scores in accuracy (4.293), consistency (4.304) and completeness (4.322).

Similarly, the end matter section with the following compliance scores in accuracy (4.581), consistency (4.471) and completeness (4.551). Lastly, compliance to pagination had the following scores in accuracy (4.634), consistency (4.629) and completeness (4.623). However, while these sections generally show good compliance, the moderate degree of variation suggests that there are some inconsistencies in adherence to the guidelines, and targeted improvements could enhance their overall performance. Overall, the "Order Of Text" demonstrates exceptional compliance with nearly perfect scores in accuracy (4.969), consistency (5) and completeness (5). Similarly, compliance to the prescribed length of an ETD indicates remarkable scores in accuracy (5), consistency (5) and completeness (5). Survey responses for current postgraduate students According to the responses from current postgraduate students, the most common challenges in

ensuring correct document formatting in accordance with the universities guidelines included difficulties with citation management, generating the correct table of contents, and formatting tables and figures. Out of 22 respondents, seven (7) highlighted that they find difficulties in adjusting their manuscripts to the prescribed margins, another seven (7) highlighted difficulties when working on tables and figures. Ten respondents stated that the most challenging aspect is with regards to generating references and citations. Five respondents also highlighted that following the prescribed table of contents is challenging for them. Of the 22 respondents, 50% stated that they were not very familiar with the postgraduate guidelines of the institution while 6

Survey results indicate that postgraduate alumni faced challenges with formatting their manuscripts, particularly in adding tables, maps, figures (5 respondents), managing citations (4 respondents), and inserting footnotes (2 respondents). These difficulties align with the low scoring of miscellaneous guidelines (covering citations, figures, and tables) in historical ETD analyses.

Examiners also reported challenges in manually verifying manuscript compliance. With a high student-to-examiner ratio in some schools, the workload can be overwhelming, limiting their ability to be thorough. Due to time constraints, examiners often review up to 50 manuscripts per sitting to provide timely feedback and avoid delaying student research. Furthermore, examiners stated that they find themselves either resending the guidelines to the students or scheduling short seminars mid-research to re-educate the students about the guidelines when they notice a lot of inconsistencies in the submitted manuscripts. Conclusion The use of DLA methods to create a pipeline that will curtail the challenge of ensuring and checking compliance of ETDs to postgraduate guidelines at the university of Zambia has proven to be the most efficient solution. The development of an ETD automatic guideline verification tool presents an opportunity to enhance efficiency as well as promote consistency in the quality of ETDs while alleviating the challenges faced in the process of manually checking for compliance consequently reducing the workload for students, supervisors and examiners alike.

References:

- Mishra, B. K., & Kumar, R. (2020). Natural Language Processing in Artificial Intelligence. Razi, S., Glendinning, I., & Foltýnek, T. (2019). Towards Consistency and Transparency in Academic Integrity. Peter Lang GmbH, Internationaler Verlag Der Wissenschaften. Binmakhashen, G. M., & Mahmoud, S. A. (2019). Document Layout Analysis: A Comprehensive Survey. ACM Comput. Surv., 52(6), 1–36. The University of Zambia. (2015). Directorate of Research and Graduate Studies: Postgraduate regulations.

width=!,height=!,pages=1, width=

