

RUFORUM Working Document Series (ISSN 1607-9345) No. 14 (1): 327-335.
Available from <http://repository.ruforum.org>

Research Application Summary

Implementing a digital repository to increase access to agricultural higher education research and learning in Africa: The Regional Universities Forum for Capacity Building in Agriculture experience

Nnam, J., Apio, J., Kuteesa, C., Kalungi, J. & Adipala, E.

Regional Universities Forum for Capacity Building in Agriculture (RUFORUM), P. O. Box 16811,
Kampala, Uganda

Corresponding author: j.nnam@ruforum.org

Abstract

Providing access to knowledge outputs is important to support development in Africa through strengthening learning, innovation, and avoiding duplication. However, although there has been an increase in the research output from the continent, a lot of it is still inaccessible. Institutional repositories have been identified as a feasible means of closing this access gap, which would help to showcase research and making it visible to search engines and thus easy for users to discover. RUFORUM's Theory of Change envisions the network becoming a continental reference point for agricultural research and higher education, a goal that will require increasing access to the network's rich knowledge base. The RUFORUM Institutional Repository, launched in 2013, provides a boost towards achieving this goal. Presently, the repository is an online storehouse for the collection, storage and retrieval of the network's knowledge output and holds close to 1,500 publications comprised of student theses and dissertations, journal publications, conference papers, project reports, as well as corporate and research outputs from the RUFORUM Secretariat. Three years after its launch, this paper documents reflections on the operations, challenges and lessons learned drawn from implementation of the repository and the way forward. The main challenges relate to the length of time taken to upload documents, acquisition of publications, and access restrictions instituted by publishers, including universities.

Key words: Access to knowledge, Institutional repositories, RUFORUM, sub-Saharan Africa, universities

Résumé

Le fait de fournir l'accès aux productions de la connaissance est important pour soutenir le développement en Afrique à travers le renforcement de l'apprentissage, l'innovation, et éviter la duplication. Cependant, bien qu'il y ait une augmentation de la production de la recherche du continent, beaucoup d'entre elle restent encore inaccessibles. Les banques de données institutionnelles ont été identifiées comme un possible moyen de combler cet écart

d'accès, ce qui permettrait de mettre en valeur la recherche et la rendre visible aux moteurs de recherche, et ainsi faciliter la découverte des utilisateurs. La théorie de RUFORUM du changement envisage de faire du réseau un point de référence continental pour la recherche agricole et l'enseignement supérieur, un objectif qui impliquera l'accroissement à l'accès à la riche base des connaissances du réseau. La base des données institutionnelles du RUFORUM, lancée en 2013, fournit un coup de pouce pour atteindre cet objectif. Présentement, la banque des données est un entrepôt en ligne pour la collecte, le stockage et la récupération de la production de connaissances du réseau, et détient près de 1.500 publications comprenant des thèses et des mémoires des étudiants, des publications dans des revues, des documents de conférence, des rapports de projet, ainsi que des productions d'entreprise et de la recherche du Secrétariat du RUFORUM. Trois ans après son lancement, ce papier garde la documentation des réflexions sur les opérations, les défis et les leçons apprises tirées de la mise en œuvre de la banque des données et de la voie à suivre. Les principaux défis ont trait à la longueur du temps nécessaire pour télécharger des documents, l'acquisition de publications, et les restrictions d'accès mis en place par les éditeurs, y compris les universités.

Mots clés: accès à la connaissance, Banque des données institutionnelles, RUFORUM, Afrique au Sud du Sahara, les universités

Introduction

Providing access to knowledge outputs from agricultural higher education research and experience in Africa is important if it is to support development of the continent through strengthening learning, innovation, and avoiding duplication. However, although there has been an increase in the research output from Africa over the past two decades, research on the continent contributes only 2.3% to the global output (Schemm, 2013). This signals that only a small portion of research in Africa is accessible. Universities for example, generate a lot of research on an annual basis in the form of student theses and dissertations, but most of these remain confined to university libraries or online databases which only students and staff can access.

In the face of this accessibility gap, institutional repositories have been identified as a viable solution. According to Gibbons (2009) an institutional repository helps to showcase the research at an institution and expose the metadata of repository contents to metadata harvesters such as scholarly search engines and general-purpose ones like Google and Yahoo. Through this exposure, scholarship, once relegated to the hard drives of personal computers, is now discovered, read, and cited because it was deposited into an institutional repository. Jain *et al.* (2009) agree with this assertion, stating that "maximising public access to research findings online, in turn maximises its visibility, usage and impact".

The Regional Universities Forum for Capacity Building in Agriculture (RUFORUM, see www.ruforum.org for details) Theory of Change envisions the network becoming a continental reference point for agricultural research and higher education (RUFORUM, 2015). Achieving this goal will involve, among other actions, ensuring access to the network's rich knowledge

base of both explicit and tacit knowledge through multichannel information dissemination and exchange. The RUFORUM Repository, Launched in 2013: is an online storehouse for the collection, storage and retrieval of the network's knowledge output and presently housing a collection of close to 1,500 publications. The collection includes: students' theses and dissertations on agriculture, rural development and related fields; journal articles, and corporate outputs from the RUFORUM Secretariat. Three years after the launch of the repository, this paper reflects on its implementation, challenges and lessons learnt that can benefit institutions seeking to implement a similar initiative.

Implementation of the current RUFORUM Repository

The collection. Presently, the RUFORUM repository has close to 1,500 publications. These are a mix of both academic and corporate publications from the network and represent varied document types including student theses, journal articles, policy briefs and extended abstracts. Most of the publications are in English, the official language of the network. A few publications, from Eduardo Mondlane University in Mozambique, are in Portuguese and abstracts of all papers submitted for the RUFORUM Biennial conferences are translated in French. Extended abstracts presented during RUFORUM's past biennial conferences, the network's largest knowledge sharing and networking event, represent the bulk of the collection as seen in Figure 1.

Technology platform. The repository is hosted on the AgriDrupal platform (www.agridrupal.org) preferred over the Dspace platform (www.dspace.org) that a number of universities have adopted. AgriDrupal is a derivative of the Drupal Content Management System, customised by the Food and Agricultural Organisation of the United Nations (FAO) for use by agricultural information management specialists. Like Dspace, the AgriDrupal platform is built on free and open source software meaning that the software is acquired at no cost and the source code is made fully available for customizing according to the needs of a user. AgriDrupal's unique selling points for RUFORUM were that it is already customized for agricultural information management, is simple to install and maintain, and allows advanced searching among other attributes (Levey, n.d.). AgriDrupal also integrates two controlled subject vocabulary indexes, Agris and Agrovoc, which help to ensure standardization and consistency in subject description. This vocabulary control helps to ensure that keywords used to describe documents match the commonly used search terms thus making it easier for users to find the documents. The predefined list of subject categories provided by Agris and Agrovoc's ability to scan a document and automatically assign it keywords also simplifies the work of the repository editors who may not be specialists in agriculture and therefore not be familiar with the standard vocabulary used to describe concepts in the subject.

Supported document formats. Most documents on the repository are stored in Portable Document Format (PDF) format to help maintain their integrity and make them smaller in size thus easier to store and share. Any publications in Microsoft Word document format are first converted to PDF format before uploading, except for Open Education Resources, which are required to remain in a format that is easy to edit. The repository also supports other file formats like images, spreadsheets, and presentations. Media like videos which

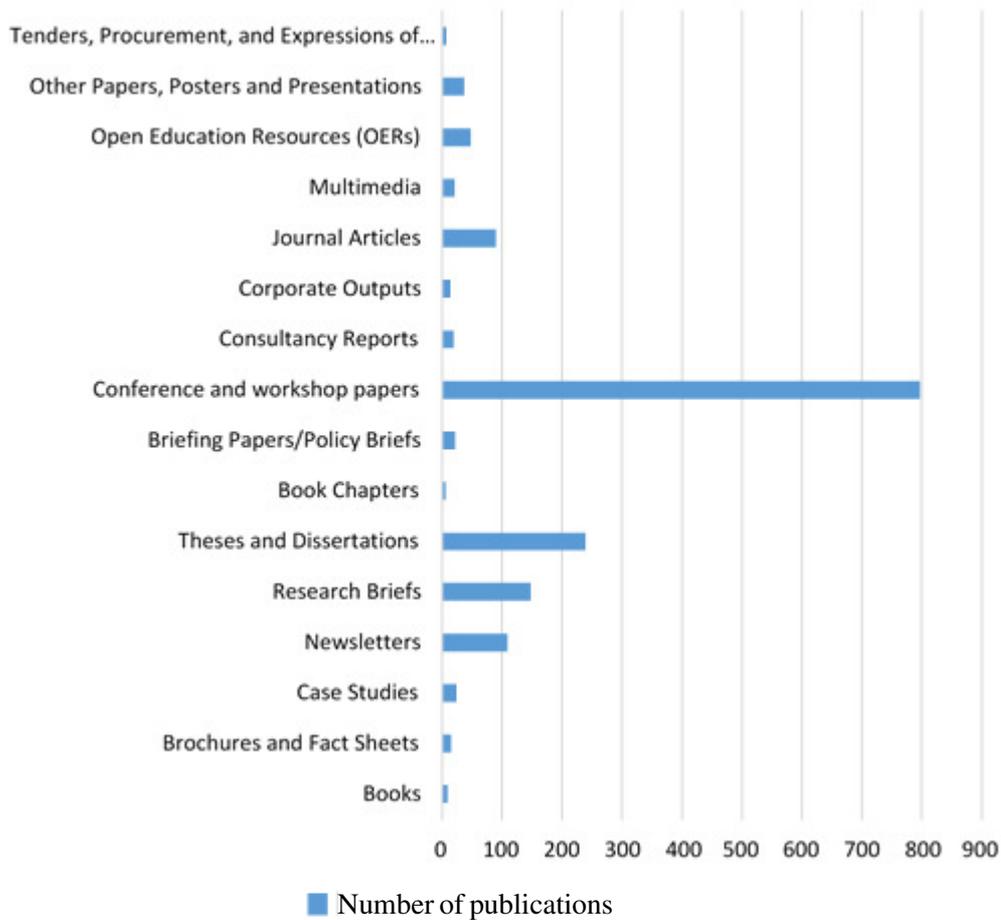


Figure 1. Distribution of publications in the RUFORUM repository

consume a lot of storage space are hosted on external websites like Youtube and then linked to the metadata description on the repository. Moreover, hosting multimedia on many of these dedicated platforms comes at no extra cost and increases their visibility.

Collection management. Acquisition of publications from outside the RUFORUM Secretariat is principally through two ways. Extended abstracts and scientific posters are acquired through responses to calls for submission of abstracts for the RUFORUM biennial conferences, which are held every other year. The second means of acquisition is through mandatory submission of student theses and other derivative publications, like journal articles, from RUFORUM funded research. RUFORUM alumni and other practitioners also send in their materials for sharing.

Updating the repository. Uploading of publications to the repository is done on a continuous basis, by an assigned staff member, the Knowledge Management Officer. The process of uploading publications is a manual one and rather lengthy as a publication has to be described to the fullest extent possible so as to facilitate retrieval by users with varied information

seeking strategies. On average, it takes about thirty minutes to upload a single publication. Part of this delay is due to the need to ensure authority control or consistency in the naming of authors, conferences and journals. For example, before entering author information for a document, the repository's author database has to be searched to check if a record for that author already exists. If it does, then the form of the author name in the database is adopted. If there is no record of the author in the author database, then their name and aliases have to first be added to the database. In the case of journal articles which sometimes have up to six authors, each author has to be searched and/or added separately which can take a considerable amount of time. During addition of author information, the full form of the author name is used if available. Among other advantages, authority control enables easy finding of all material by a given author, conference or journal.

The number of fields to be filled varies by the publication type, but can be as many as thirty for journal articles. Although it has been rather difficult to implement, a fully described publication derived from RUFORUM-funded research should have the standard bibliographic information as well as additional contextual information such as the associated research project, its funder, research supervisor and whether the grant is still ongoing or not. Including this additional information further lengthens the process of uploading documents, but has the advantage of enabling publication tracking by project as well as providing visibility to the project partners and lead researchers.

Access rights control. The majority of publications in the RUFORUM repository are publicly available under open access terms that permit free sharing, downloading and reuse. This philosophy is generally shared by RUFORUM's funding partners. However, there are instances where publications are copyrighted and permission has to be sought from authors or publishers for sharing and reuse. Such publications are made openly accessible when authors grant RUFORUM the right to archive them in the repository, which has happened in many cases. In a few cases, though, all forms of sharing and storage, even storage in retrieval systems in the author's institution is forbidden. Where such restrictions exist, RUFORUM only shares descriptive information about the document and an abstract.

Dissemination. Once uploaded, the latest addition to the repository is displayed on the homepage of the platform and instantly accessible to the public. Publications can be found through search engines like Google, or specialized indexing services like CAB Abstracts (see www.cabi.org for details). As they are published under open access terms, they can be downloaded, printed and re-shared. Of recent, RUFORUM uses its corporate communication channels like its online weekly digest, the RUFORUM Weekly, a publication that reaches over 8,000 subscribers, to share information on the latest publications in the repository. This has proven to be an effective means of current awareness on new additions to the repository. However, other automated methods like the use of RSS feeds and social media to inform on new acquisitions are also being explored.

Monitoring usage. There are many possibilities for analyzing the repository collection and use including the number of times a document is read. A shortcoming of the readership statistics, however, is that they are not unique, meaning that if one person reads a document

multiple times, each time is counted as a unique instance. Such usage statistics help to inform on user interest and information seeking behavior that can help improve service provision.

Challenges of implementing the repository

- 1. Creating an authority list of author names.** In order for users to easily find all publications by a particular author, it is important that each author is identified by a unique name which is used consistently. However, it is common that authors write their names in different styles across publications. For example, while an author may write their first middle and surnames in full on a thesis, they may initial or completely leave out their middle name in a journal paper. It therefore becomes difficult to establish whether J. Smith represents John Smith, James Smith, or John Mark Smith.
- 2. Long uploading time.** Cataloguing or describing publications as fully as possible, while ensuring authority control, enables users to find them and quickly make a decision on whether to retrieve the full publication. However, thirty minutes to upload each publication makes the book time intensive. The 1,500 publications currently in the repository, represent about 750 hours or close to five months of staff time.
- 3. Building the collection.** Submitting publications from RUFORUM-funded research is a requirement for RUFORUM grantees. However, while it is easier to solicit and obtain student theses which they submit either to their faculties, university libraries and/or schools of postgraduate studies at the end of their courses, acquisition of journal publications is more difficult as sometimes students publish after graduation. Also, some journals place restrictions on access and dissemination. Increasingly though, universities are encouraging students to publish before graduation, and in some cases like the taught PhD in Agricultural and Rural Innovation Systems at Makerere University, journal articles constitute chapters of the student's thesis. Such practices which encourage publication during the course make tracking and acquisition of journal publications much easier.
- 4. Restricted access to some publications.** RUFORUM generally adopts and encourages an open access policy for publication and makes publications on the repository fully available for download. However, occasionally, publications from RUFORUM funded research, especially journal articles and books, are published under restricted terms and RUFORUM can only share meta-data and the abstract information, but not provide full access to the publication.
- 5. Limitations of Agrovoc and Agris vocabularies.** Subject keywords are important in retrieval of publications whether through website browsing or internet search engines. However, use of pre-determined keywords, while easing the task of document description may limit their accurate description. For example, "in Agris subject categories", popular new terminologies such as sustainable development, agribusiness, and conservation agriculture are lacking and may have to be broadly classified as "agricultural research" which is rather vague. With Agrovoc which scans and assigns keywords to documents,

sometimes assigned words are not accurate, yet they cannot be edited. Although these challenges are partially addressed through presence of an open field for a user to enter subject keywords, the chosen words may not be the popular search terms used.

- 6. Low searchability.** Searchability is linked to access, as information that cannot be found is almost as good as that which does not exist. Publications in the repository should be visible to search engines as not everyone will access them through its direct website address (www://repository.ruforum.org). It thus becomes important for the repository to be optimised for readability and indexing by search engines like Google and Google Scholar. This has posed a challenge so far, although efforts are currently underway to address it.

Lessons learned from implementing an institutional repository

Issue	Lessons
Technology platforms	<ol style="list-style-type: none">i. Choose open source software for more flexibility in customizing the platform to one's requirements. Agridrupal provides the special advantage of being customized for agricultural information management systemsii. Store heavy multimedia files on free dedicated platforms e.g. Youtube, to save on storage space in the repositoryiii. To speed up uploading time, if possible, create an offline mirror of the repository where documents can be uploaded and synced later with the online version to update it.
Collection building	Have a plan for mobilizing content or one may end up with an empty platform. Reflect on the information flow process from the centres where the knowledge is generated to the desk where it is uploaded. This will help in exposing possible bottlenecks to be addressed, for example restrictive access
Human resource	<ol style="list-style-type: none">i. To prevent a backlog in document uploads, assigned the task to dedicated staff and source for help when support is needed to upload a big batch. Interns are very useful in this regard and may be drawn from any discipline as long as they have good general knowledge and are well oriented.ii. If there is more than one person managing the repository, each should have a login account and access rights that match their responsibilities. There are different levels of administrator rights the lowest being an editor who can upload documents, but not delete them or change system settings and the highest being the super administrator who can overwrite the system.

- iii. Institutions without robust IT departments with adequate infrastructure to support hosting and maintenance of the repository are better off hiring the system developers on a retainer contract to manage the day-to-day technical issues. This would be cheaper and more efficient than running the services in-house which would require purchase and maintenance of expensive equipment and hiring of staff
- Authority lists
- i. Agree on a standard format for author names, early on, to avoid multiple entries of the same author under multiple forms of the same name.
 - ii. To supplement controlled subject vocabularies, have an additional open field for subject description where one can enter keywords not available in the vocabulary list.
 - iii. The author database that is inherently linked to publications in the repository presents an opportunity to identify experts in different fields
- Promoting sharing and use
- i. Use institutional communication channels for example newsletters, flyers and social media channels to notify users about new publications in the repository
 - ii. Enrich institutional reports and communication products with links to the relevant resources in the repository
 - iii. Integrate social media on the platform to make it easy for users to re- share resources
 - iv. Integrate RSS feeds to automatically share updates on latest resources added
 - v. When sharing a document, reference the page link rather than the direct document link. That way, in case the document is replaced, users can still find the new version using the same link
 - vi. Make the repository search engine friendly as most publications are accessed through search engines.
 - vii. Negotiate and clarify access policies with project funders at the start of any organisational project or programme and make sure all beneficiaries are aware of their obligations. It helps to include these obligations in grantee award letters or contracts
- Monitoring and learning
- i. Incorporate and make use of analytical reports on the collection status, access and use patterns of the repository. Statistics and trends in the number of publications, publishers, topics covered, most read resources and the entry point to the repository (whether through direct link or search engines) help to inform on user behavior and identify areas for improvement

- ii. As well as collecting data on total number of reads for a document, for a more complete picture, collect information on how many of these are unique

Conclusion and recommendations

The RUFORUM Institutional Repository is a cornerstone in achieving the network's goal of becoming a regional reference point on agricultural research and higher education. Three years after commencing its implementation, the repository has become firmly integrated in the Network's knowledge management strategy and achieved commendable strides in maintaining a stable platform, building a formidable collection of resources, promoting open access publication and increasing readership. The lessons and experiences are useful to guide similar institutions on good practices and issues to be addressed in order to enhance access to research. For RUFORUM, these lessons and experiences are the springboard for it to leap to the next step in transforming the repository from an archive of the network's scholarly and practice-based publications into a disciplinary repository, archiving scholarly publications from university led agricultural research and on higher education in Africa. No longer just a storehouse for publications resulting from research that it funds, the RUFORUM Institutional Repository should become a gateway to relevant research from a wider range of stakeholders and sources on the continent.

Acknowledgement

This paper is a contribution to the 2016 Fifth African Higher Education Week and RUFORUM Biennial Conference.

References

- Gibbons, S. 2009. Benefits of an institutional repository. *Library Technology Reports* 40 (4): 11-16.
- Jain, P., Bentley, G. and Oladiran, M.T. 2009. The role of institutional repository in digital scholarly communications. *Proceedings of African Digital Scholarship & Curation*. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.517.8448&rep=rep1&type=pdf>
- Levey, L.A. (n.d.). Maximizing RUFORUM's knowledge management and outreach: building an institutional repository and enhancing the effectiveness of the RUFORUM website. A consultancy report for the Bill and Melinda Gates Foundation.
- RUFORUM. 2015. Coming of Age: RUFORUM Strategic Business Plan 2015-2020. Unpublished.
- Schemm, Y. 2013. Behind the data: Africa doubles research output over past decade, moves towards a knowledge-based economy. *Research Trends: Issue 35 December 2013*. Retrieved from <https://www.researchtrends.com/issue-35-december-2013/africa-doubles-research-output/>