



Exploring the Workflow of Document Preparation in Digitizing Research Output: Librarians' Insights in Ahmadu Bello University, Zaria, Nigeria

¹Moruf Hawwau A., ²Enang Uduak U., ³Umoren Eboro E.

¹Department of Library and Information Science,
Federal University Dutsin-Ma, Katsina State

^{2&3}Department of Library and Information Science,
University of Uyo, Uyo State

ABSTRACT

Document preparation is the overall handling of text for processing which begins with obtaining selected documents to be digitized; registering to keep track of them; checking for suitability and prepare for digitization. This study was carried out to investigate the influence of document preparation on the visibility of research output at Ahmadu Bello University, Zaria, Nigeria. The study was guided by four specific objectives with their matching research questions. The design employed in this study was survey research design. Data were collected using structured questionnaire with Four-point Likert Scale. The data were analyzed using descriptive statistics of mean and standard deviation, while Student-t-test was used to test the hypotheses at 0.05 level of significance. The results revealed that the mean range for document preparation was 1.06 -3.68, with a standard deviation range of 0.42-1.31, indicating an influence on visibility of research output. Notably, the P-value (.0001) is less than the alpha level (.05), thus there is a significant influence of document preparation on visibility of research output in Ahmadu Bello University, Zaria. The result further shows that all the mean responses are above 3.0, indicating that the respondents all agreed that the listed 15 items and the identified 10 items constitute the challenges and strategies for enhancing effective digitization workflow management of research outputs respectively. The paper recommends that the preparation of documents and conversion to digital formats can be enhanced through the purchase and application of digital conversion machines and equipment.

ARTICLE INFO

Article History

Received: August, 2023

Received in revised form: August, 2023

Accepted: January, 2024

Published online: February, 2024

KEYWORDS

Digitization, Institution, Library, Research,
Publication, University

INTRODUCTION

The concept of research output is integral to the academic and scientific community, as it represents the culmination of scholarly inquiry and investigation. It is made up of two distinct words "research" and "output" – as a process and product. Davidavičienė (2018) defined research as studious inquiry or systematic investigation of some phenomenon or series of phenomena by experimental method in order to discover facts, to establish or reverse a theory, or to develop a plan of action based on the facts discovered.

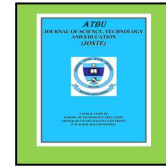
Universities are ranked on the basis of their researches/research output. It is through research output that new knowledge is generated, theories are tested and refined, and innovative solutions to complex problems are developed. The significance of research output extends beyond academic pursuits, as it has practical implications for addressing professional challenges, developing analytical tools, and informing decision-making processes (Mohammed et al., 2019). Moreover, Ugwuona and Dike (2015) emphasize the transformative potential of

Corresponding author: Moruf Hawwau A.

✉ hawwau.moruf@yahoo.com

Department of Library and Information Science, Federal University Dutsin-Ma, Katsina State.

© 2023. Faculty of Technology Education. ATBU Bauchi. All rights reserved



research output, emphasizing its role in translating knowledge into tangible technologies, interventions, and strategies that can benefit society. In today's digital age, research output is disseminated in various forms and formats, with traditional paper-based publications being supplemented and, in some cases, replaced by digital materials. This shift towards digital dissemination reflects the evolving nature of scholarly communication and the increasing accessibility of research outputs to a global audience.

Academic library is a platform that support the representation of dully articulated, understood and documented research outputs that sprout as a result of activities such as teaching and research engagements. Libraries in public universities have become the repertoire for research output management through their institutional repositories (Mohammed et al., 2019). The electronic version of research outputs forms part of the university library collections, hence, the need for their preservation. These resources enhance the institution's aim in visibility, increases its prestige and the reputation of its researchers. To systematically respond to the preservation of these scholarly resources, there is need for a holistic preservation technique. University libraries are saddled with the responsibility of preserving and disseminating these research outs for global visibility (Inuwa and Abrizah, 2018).

In preserving and disseminating the research outputs, university libraries embark on digitization. Digitization is the process of converting information into a digital (i.e. computer readable) format, in which the information is organized into bits. Digitization refers to the process of converting analogue or physical data, such as paper documents or images, into digital formats that can be easily accessed, stored, and manipulated using computers and digital technologies (Dieffenbacher, 2023). The result is called digital representation or, more specifically, a digital image for the object, and digital form for the signal (Oyewole, 2020). Digitization can be performed on different types of information materials such as text (paper documents, articles, books, data etc.), images (moving or still images),

audio and video. The focus of this study is on paper and born digital research documents. Digitization has removed physical barriers to the discovery and use of rare and fragile resources and those recorded on difficult-to-access formats. Access to paper manuscripts and research documents has often been limited because of their value and/or fragile nature. Digitization not only expands the reach of these materials to researchers, students, and the general public, but in many cases it enhances the visual quality of faded and illegible documents (Zhou et al., 2023). Digitization as a practice improves efficiency, accessibility, and preservation of information content. Digitization therefore becomes a necessity to enhance visibility of scholarly research publications. To digitise these resources, a workflow must be followed which in other words is termed as digitization workflow.

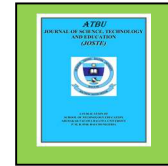
Digitization workflow entails all the steps involved in the process of converting non-digital research collections to a digital format and making them available online. The digitization process encompasses a range of procedures and technologies with widely varying implications and costs ranging from preparation and conversion to presentation and archiving (Prescott and Hughes, 2018). It is the process of digitizing information resources and requires careful justification, planning and systematic implementation. Digitization workflow is a step by step approach to explaining the processes involved in digitization of information resources. It is essential to study the digitization workflow as it helps to streamline the many processes and tasks involved in a digitization project, thereby facilitating consistency and reliable results, especially when dealing with a large volume of information resources to be digitized (Sharma et al., 2021). Researchers have identified various phases of digitization with each leading to the end result of preservation and access creation. Chisita et al. (2021) identified document digitization workflow to include but not limited to document selection, document preparation, scanning, editing, metadata and indexing, document validation, archiving and publishing.

Corresponding author: Moruf Hawwau A.

✉ hawwau.moruf@yahoo.com

Department of Library and Information Science, Federal University Dutsin-Ma, Katsina State.

© 2023. Faculty of Technology Education. ATBU Bauchi. All rights reserved



Document preparation is the starting point for digitization processing. It ensures checking the physical and visual condition of the material to remove any blemishes. The preparation of document is key to enable a flawless procedure that is devoid of office objects such as document clips, sticky notes, pins and spiral bind. This stage involves overall handling of the document in preparation for processing which begins with obtaining selected documents to be digitized; registering documents to keep track of them; checking document for suitability and prepare for digitization. Towle (2022) mentioned the considerations to determining document suitability for digitization to include but not limited to the format, the physical condition (damaged, fragile, curled etc.), and uniqueness non-degraded with sufficient clarity for detail. After the suitability of the documents has been determined, the document is prepared for scanning.

Most of the researches submitted to the university libraries nowadays are produced in hard copies where the soft copies are hardly available for use. It is also observed that most researchers prefer submitting published researches to their institutions. This thereby required the university libraries to convert and make them available through digitization. According to earlier studies, only a small fraction of Nigerian libraries has started the process of digitizing their holdings. There is still a dearth of information on the digitization workflow carried out by university libraries in Nigeria. Hence, the need to examine the influence of document preparation in digitization workflow on research outputs in Nigeria.

Purpose of the Study

The general objective of the study is to evaluate the impact of document preparation as a workflow process on the visibility of research outputs in Ahmadu Bello University, Zaria, Nigeria. The specific objectives are:

1. To determine the influence of document preparation in digitization on the visibility of research output in Ahmadu Bello University, Zaria.

2. To investigate the relationship between document preparation in digitization and the visibility of research output in Ahmadu Bello University, Zaria.
3. To find out the challenges involved in digitization of research output in Kashim Ibrahim Library, Ahmadu Bello University, Zaria.
4. To ascertain strategies for enhancing effective management of digitization of research output in Kashim Ibrahim Library, Ahmadu Bello University, Zaria.

Research Questions

This study was guided by the following research questions:

1. What is the influence of document preparation in digitization on the visibility of research outputs in Ahmadu Bello University, Zaria?
2. What is the relationship between document preparation in digitization and the visibility of research outputs in Ahmadu Bello University, Zaria?
3. What are the various challenges encountered in digitization of research output in Kashim Ibrahim Library, Ahmadu Bello University, Zaria?
4. What are the strategies that could be used to enhance effective digitization of research outputs in Kashim Ibrahim Library, Ahmadu Bello University, Zaria?

Hypothesis

There is no significant relationship between document preparation in digitization and the visibility of research output in Ahmadu Bello University, Zaria.

LITERATURE REVIEW

Han and Wan (2018) carried out a study on preparing a text document using PDF/A which involve real-life experience of converting TIFFs/JPEG2000s to PDF/A, and back, along with image metadata. The study found that using PDF/A gives institutions advantages of the newly preferred master file format for digitization of text



documents over TIFF/JPEG2000. The two studies share the same variable which is document preparation, however, they differ in terms of scope, objectives and methodology.

Ong et al. (2023) studied the digitization process of natural history collections. This study revealed the importance of document preparation in pre-production, production and post-production phases in digitization. It also proved that inadequate document preparation in post-production phase influences visibility of such document. The two studies share the same variable but different research methodologies.

Matthias et al. (2019) carried out a study on conversion of data formats which are not machine readable and consist of many different document elements such as tables, figures or text sections. The study concluded that documents with low precision and high recall lead to difficulty of identification, extraction and visualization as this affects retrievability and analysis of such documents. The study is related to the current study as it revealed the precision level that can influence visualisation in visibility. Thus, different approaches would assist in dealing with document preparation process for information extraction, information retrieval as well as content analysis of unstructured data. Both studies are related in their aim to enhance accessibility through document preparation however, their research approaches differ.

Ramdhani et al. (2021) examined the performance of Optical Character Recognition (OCR) engines in information extraction by comparing three OCR engines with the highest performance (Foxit, PDF2GO, Tesseract) over 8,562 government human resources documents within six document categories, two document structures, and four measurements. The study confirms the similarities between document preparation for scanning and visibility as inappropriate text extraction in data mining and curatorial works affect visibility and accessibility sometimes as a result of low precision and recall. Though the preceding study is experimental, it relates to the current study in term of variable but differs in methodology.

Rafiq et al. (2018) studied the barriers to digitization in university libraries of Pakistan. The finding revealed the barriers which were hindering the digitization activities in university libraries of Pakistan to include lack of technical-know-how of digitization and its requirements among library professionals. The study established that there is a relationship between technical processes involved in preparation of source document which impedes the digitization efforts and thereby affects visibility of digitised scholarly resources. The similarity is that the prior study explored technical process in digitization as a single variable while the current study identifies document preparation as a technical process as well as relates the effect on visibility. Equally, the scope of both studies cover university libraries, but differ in the area of study.

Javed and Shafait (2017) carried out an experimental research on real-time document localization in natural images using recursive application of a Convolutional Neural Networks (CNNs) for seamless integration of information from digital and paper documents. The results are comparable to state of the art on simple backgrounds and improve the state of the art to 94% from the previous 86% on the complex background. The prior study is experimental while the current study is descriptive. They therefore differ in terms of number of variables and methodology.

Coulbourne et al. (2022) conducted a case study research on the preparation of brittle newspapers for digitization. The study concluded that the process enables to store stabilized original newspapers in perpetuity in an organized and space efficient manner. Though the two studies relate in dealing with the same variable which is document preparation in digitization, while the current study differs by examining the influence of document preparation on visibility.

METHODOLOGY

Descriptive survey design was employed in this study. Questionnaire, as instrument was used to elicit information on the dependent and independent variables as well as answer the research question and test the

Corresponding author: Moruf Hawwau A.

✉ hawwau.moruf@yahoo.com

Department of Library and Information Science, Federal University Dutsin-Ma, Katsina State.

© 2023. Faculty of Technology Education. ATBU Bauchi. All rights reserved



hypotheses. Four-point Likert scale was employed in order to understand respondents' opinion in respect to the extent of impact without being neutral on a specific item. A total of 58 selected librarians with digitization experience served as the target population within Kashim Ibrahim Library, Ahmadu Bello University, Zaria. A total enumeration sampling method was employed, since the population size was manageable.

Descriptive statistics was used to answer the research questions. Student-t-test was used to test the hypotheses at 0.05 level of significance.

RESULTS

Research Question 1: What is the influence of document preparation in digitization on the visibility of research output in Ahmadu Bello University, Zaria?

Table 1: Mean and Standard Deviation of Respondents on the influence of document preparation on visibility of research outputs in Ahmadu Bello University, Zaria

S/N	To what extent do the following document preparation techniques influence visibility of research outputs	Mean	Std. Dev	Remarks
1	Identifying documents with special handling or treatment before digitization	3.02	1.02	Great Extent
2	Sorting documents based on academic importance	3.10	0.95	Great Extent
3	Checking document for suitability	3.18	0.87	Great Extent
4	Organizing documents based on type, size and format	3.07	1.08	Great Extent
5	Identifying conservation and preservation treatments for selected document	3.61	0.54	Very Great Extent
6	Identifying appropriate cleaning methods for each document	2.82	1.29	Moderate Extent
7	Repairing torn pages of documents	2.28	1.31	Less Extent
8	Removing staples and paperclips	3.64	0.76	Very Great Extent
9	Stabilizing loose bindings	3.68	0.47	Very Great Extent
10	Aligning documents accurately on the digitization equipment	1.06	0.42	Less Extent
11	Re-ordering and organizing born-digital records	1.84	1.24	Less Extent
12	Automatic identification and extraction of data visualizations using algorithms.	1.93	1.24	Less Extent
Pooled Mean and Std. Dev.		2.77	0.32	Moderate Extent

Table 1 gives the summary of the mean and item analysis of respondents on influence of document preparation techniques on visibility of research output in Ahmadu Bello University, Zaria. The result shows that the mean range is 1.06 - 3.68. The standard deviation range is 0.42-1.31. This indicates that the means are dispersed. Items 1-5 and 7-8 have mean responses above 3.0, indicating that the respondents agreed that these items have a positive influence on visibility of

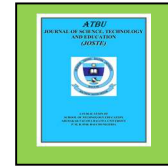
research output. Items 6 and 7 have mean responses 2.82 and 2.88 respectively. This indicates that there is less practice of identifying appropriate cleaning methods for documents and repairing of torn pages. However, items 11 and 12 have mean responses of 1.84 and 1.93 respectively, indicating that the respondents disagreed on re-ordering digital records and the use of automatic identification and extraction of data visualization using algorithms. The Pooled

Corresponding author: Moruf Hawwau A.

✉ hawwau.moruf@yahoo.com

Department of Library and Information Science, Federal University Dutsin-Ma, Katsina State.

© 2023. Faculty of Technology Education. ATBU Bauchi. All rights reserved



mean is 2.77, indicating that the respondents agreed to a moderate extent that document preparation techniques has a positive influence on visibility of scholarly research publications.

Research Question 2: What is the relationship between document preparation in digitization and

the visibility of research output in Ahmadu Bello University, Zaria?

Ho₁ There is no significant influence document preparation in digitization and the visibility of research output in Ahmadu Bello University, Zaria.

Table 2: Summary of t-test analysis for significant influence of document preparation on visibility of research output in Ahmadu Bello University, Zaria

Variables	Mean	Std. Dev	t	P-value	Decision
Document preparation	2.77	0.32	11.91	.0001	Ho ₁ Rejected
Visibility of research output	3.58	0.78			

Table 2 presents the summary of the t-test analysis. The result shows that the t-value is 11.91. The probability value at .05 alpha level is .0001. Since the P-value is less than the alpha level, the result is statistically significant and the null hypothesis rejected. Thus, there is a significant influence of document preparation on

visibility of research output in Ahmadu Bello University, Zaria.

Research Question 3: What are the various challenges encountered in the digitization of research output in Kashim Ibrahim Library, Ahmadu Bello University, Zaria?

Table 3: Summary of mean of respondents on the challenges encountered in the digitization of research output in Kashim Ibrahim Library, Ahmadu Bello University, Zaria

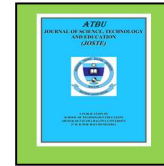
S/N	How do you agree to the following statements as challenges encountered in the digitization of research out?	Mean	Std. D	Remarks
1	Selecting collections to be digitized	3.04	1.25	Agreed
2	Absence of digitization workflow blue-print	3.18	1.17	Agreed
3	Recording or taking inventory of selected materials	3.64	0.76	Strongly Agreed
4	Creating metadata and indexing for digitized documents	3.68	0.47	Strongly Agreed
5	Scanning activities	3.14	1.30	Agreed
6	Uploading files formats	3.22	1.19	Agreed
7	Insufficient equipment for digitization	3.13	1.30	Agreed
8	Lack of technical support (hardware maintenance, software installation)	3.16	1.29	Agreed
9	Publishing through the repository	3.11	1.35	Agreed
10	Quality assurance and control	3.25	1.28	Agreed
11	Public relations and services related to marketing new digital collections	3.54	1.05	Strongly Agreed
12	Lack of established digitization plan, policies and procedures.	3.04	1.32	Agreed
13	Lack of written institutional policy on digitization	3.02	1.02	Agreed
14	Lack of suitable expertise on digitization of scholarly publications	3.10	0.95	GE
15	Lack of institutional support	3.18	0.87	GE

Corresponding author: Moruf Hawwau A.

✉ hawwau.moruf@yahoo.com

Department of Library and Information Science, Federal University Dutsin-Ma, Katsina State.

© 2023. Faculty of Technology Education. ATBU Bauchi. All rights reserved



Pooled Mean and Std. Dev.	3.23	0.25	GE
----------------------------------	------	------	----

Table 3 gives the summary of the mean and item analysis of response on challenges encountered in the digitization of scholarly research publications. The result shows that the mean range is 3.02 - 3.68. The standard deviation range is 0.47-1.35. The standard deviation indicates that the mean responses are not dispersed from each other. The result also shows that all the mean responses are above 3.0, indicating that the respondents all agreed that the 15 items listed constitute the challenges encountered in the digitization of scholarly research publications. This includes challenges in selecting collections to be digitized, Absence of

digitization workflow blue-print, Recording or taking inventory of selected materials, Quality assurance and control, Lack of established digitization plan, policies and procedures as well as Lack of suitable expertise on digitization of scholarly publications. The Pooled mean is 3.23, indicating that the respondents agreed that the items were the challenges encountered in the digitization of scholarly research publications.

Research Question 4: What are the strategies that could be used to enhance effective document preparation of research output in digitization??

Table 4: Summary of mean of respondents on strategies for enhancing effective digitization workflow management of research output in Kashim Ibrahim Library, Ahmadu Bello University, Zaria

S/N	To what extent do you agree or disagree with the following statements as strategies to improve digitization workflow management	Mean	Std.Dev	Remarks
1	Formulation of digitization plans, policies and procedures	3.10	0.95	Agreed
2	Integration of digitization skills in digital library education courses through library school curriculum	3.18	0.87	Agreed
3	Collaboration between University Library and Management information system unit.	3.07	1.08	Strongly Agreed
4	Digitization workflow should be handled purely by the library	3.61	0.54	Strongly Agreed
5	Digitization workflow should be handled by digitization experts.	3.04	1.24	Agreed
6	Strict adherence to implementing IFLA digitization guidelines on workflow and technical requirements.	3.29	1.12	Agreed
7	Digitization should be done by following a workflow system	3.72	0.68	Strongly Agreed
8	Acquisition of appropriate technology to facilitate digitization workflow	3.72	0.45	Strongly Agreed
9	Training and retraining of library professionals on digitization	3.14	1.30	Agreed
10	University stakeholders support for library	3.28	1.19	Agreed
	Pooled Mean and Std. Dev.	3.31	0.30	Agreed

Table 4 gives the summary of the mean and item analysis of response on strategies for enhancing effective digitization management of research outputs. The result shows that the mean range is 3.04 - 3.72. The standard deviation range is 0.45-1.30. The standard deviation indicates that the mean responses are not dispersed from each

other. The result also shows that all the mean responses are above 3.0, indicating that the respondents all agreed that the identified 10 items are strategies for enhancing effective digitization workflow management of research outputs. This includes formulation of digitization plans, policies and procedures, integration of digitization skills in

Corresponding author: Moruf Hawwau A.

✉ hawwau.moruf@yahoo.com

Department of Library and Information Science, Federal University Dutsin-Ma, Katsina State.

© 2023. Faculty of Technology Education. ATBU Bauchi. All rights reserved



digital library education courses through library school curriculum, digitization workflow should be handled purely by the library, digitization should be done by following a workflow system as well as training and retraining of library professionals on digitization. The Pooled mean is 3.31, indicating that the respondents agreed that the items were the strategies for enhancing effective digitization workflow management of research outputs.

DISCUSSION OF FINDINGS

The findings of this study showed varying means with most (Items 1-5 and 7-8) have mean responses above 3.0, indicating that the respondents agreed that these items have a positive influence on visibility of research outputs. This includes identifying documents with special handling or treatment before digitization, aligning documents accurately on the digitization equipment, automatic identification and extraction of data visualizations using algorithms, organizing documents based on type, size and format as well as checking document suitability. The Pooled mean is 2.77, indicating that the respondents agreed to a moderate extent that document preparation techniques has a positive influence on visibility of scholarly research publications. The null hypothesis test indicates that there is a significant influence of document preparation on visibility of research output in Ahmadu Bello University, Zaria.

The method of preparing digital documents is vital for its accessibility and visibility. Where the documents are prepared in formats that restrict access or inability to copy, or designed with too much colours, it might hinder visibility and access. This finding is in agreement with Ong et al. (2023) who studied the digitization process of natural history collections to understand the workflow in digitization process during a workshop. Findings of the study revealed the importance of document preparation in pre-production, production and post-production phases in digitization. It also proved that inadequate document preparation in post-production phase influences visibility of such document. This finding is also in line with Han and Wan (2018) who researched on preparing a text

document using PDF/A which involve real-life experience of converting TIFFs/JPEG2000s to PDF/A. The study found that using PDF/A gives institutions advantages of the newly preferred master file format for digitization of text documents over TIFF/JPEG2000. Results of findings demonstrated the ease, the reasonable runtime, and the availability of open source software to perform such conversions. From both the theoretical analysis and empirical evidences, the study showed that PDF/A has advantages over the traditional preferred file format TIFF for digitization of text documents for enhancing its visibility and usability. The study of Matthias (2019) corroborated that documents with low precision and high recall lead to difficulty of identification, extraction and visualization as this affects retrievability and analysis of such documents, thereby affecting visibility in the long run.

With regards to challenges encountered in the digitization of research outputs, the findings also showed that all the mean responses are above 3.0, indicating that the respondents all agreed that the identified 15 items are the challenges encountered. This includes challenges in selecting collections to be digitized, Absence of digitization workflow blue-print, Recording or taking inventory of selected materials, Quality assurance and control, Lack of established digitization plan, policies and procedures as well as Lack of suitable expertise on digitization of scholarly publications. The Pooled mean is 3.23, indicating that the respondents agreed that the items were the challenges encountered in the digitization of scholarly research publications. The identified challenges are similar to Rafiq et al. (2018) who examined the barriers to digitization in university libraries of Pakistan.

The finding revealed the barriers which were hindering the digitization activities in university libraries of Pakistan to include lack of technical-know-how of digitization and its requirements among library professionals. This finding is further corroborated by Nasidi and Jika (2023) whose findings indicated that there were digitization efforts and functional digital repositories in the three selected universities

Corresponding author: Moruf Hawwau A.

✉ hawwau.moruf@yahoo.com

Department of Library and Information Science, Federal University Dutsin-Ma, Katsina State.

© 2023. Faculty of Technology Education. ATBU Bauchi. All rights reserved



studied but the digitized resources are not well structured to specifically link resources in the departments, faculties and other units of the universities with the central repositories. It was also found one of the challenges of digitization is that digital repositories were not always accessible, possibly due to inadequate ICT infrastructure and necessary staff skills for their effective maintenance and visibility. This finding is also supported by Kim and Maltceva, (2022) who identified the main constraints on digitization within libraries, archives, and museums to include; lack of written policy on digitization, inadequate ICT infrastructures and manpower, inadequate funding and government support, low level of document usage due to little or no visibility as well as lack of adequate user education on utilization of digitized resources and services.

The findings on digitization management of research outputs shows mean responses above 3.0, indicating that the respondents all agreed that the identified 10 items are the strategies for enhancing effective digitization workflow management. This includes formulation of digitization plans, policies and procedures, integration of digitization skills in digital library education courses through library school curriculum, digitization workflow should be handled purely by the library, digitization should be done by following a workflow system as well as training and retraining of library professionals on digitization. The Pooled mean is 3.31, indicating that the respondents agreed that the items were the strategies for enhancing effective digitization workflow management of scholarly research publications. This is similar to Boros et al. (2020) who studied the manner of alleviating digitization errors in named entity recognition for historical documents.

Findings showed that the proposed model clearly improves the results on both historical datasets, and does not degrade the results for modern datasets. This study thus justifies the need for adopting named entity recognition (NER) strategies to digitized scholarly research texts. This finding is reinforced by Kavishe and Dulle (2016) which recommended integration of digital library courses into library and

information science programs in library schools with a view to inculcate technical skills in managing digitization workflow. It has been affirmed that activities in digitization workflow are alternative strategies to increase the visibility and accessibility to digitized and stored collections (Corona, 2023; Nelson et al., 2015).

CONCLUSION

Based on the findings of the study, it is concluded that digitization workflow is an on-going process in the university and it is happening at a moderate rate. The outcome of the study shows that there is a positive influence of document preparation on the visibility of research output in Ahmadu Bello University, Zaria.

RECOMMENDATIONS

Based on the findings of the study, the following recommendations are made

1. Institutions should have a tailored focus on the kind, volume and quality of information resources and services that they seek to digitize and make public.
2. The preparation of documents and conversion to digital formats can be enhanced through the purchase and application of digital conversion machines and equipment.
3. Tertiary institutions should identify and partner with indexing companies and associations, this is to help institutions get their digital files indexed in order to enhance their visibility.
5. Workflow management system which guides the digitization processes should be developed and tailored to the need of each institution library.

REFERENCES

- Boros, E., Hamdi, A., Pontes, E. L., Cabrera-Diego, L. A., Moreno, J. G., Sidere, N. and Doucet, A. (2020). Alleviating digitization errors in named entity recognition for historical documents. In *Proceedings of the 24th conference on computational natural language learning* (pp. 431-441). Association for

Corresponding author: Moruf Hawwau A.

✉ hawwau.moruf@yahoo.com

Department of Library and Information Science, Federal University Dutsin-Ma, Katsina State.

© 2023. Faculty of Technology Education. ATBU Bauchi. All rights reserved



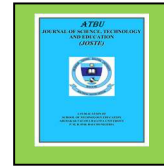
- Computational Linguistics.
<https://doi.org/10.18653/v1/P17>
- Chisita, C. T., Durodolu, O. O. and Ngoaketsi, J. (2021). Evaluating the processes and procedure of Digitalization Workflow. In Digital Libraries-Advancing Open Science. IntechOpen. 12pp.
- Corona, L. (2023). Digitization for the visibility of collections. *Collection and Curation*, 42(3): 73–80.
- Coulbourne, M., Pike, R. and Draper, B. (2022). Preparing and Digitizing Brittle 19th and 20th Century Newspapers. Mid-Atlantic Regional Archives Conference.
- Davidavičienė, V. (2018). Research methodology: An introduction. Modernizing the Academic Teaching and Research Environment: Methodologies and Cases in Business Research, 1-23.
- Dieffenbacher, S. F. (2023). Digitization vs digitalization: differences, definitions, and examples <https://digitalleadership.com/blog/digitization-vs>
- Han, Y. and Wan, X. (2018). Digitization of Text documents Using PDF/A. *Information technology and libraries* pp.52-64.
- Inuwa, S. and Abrizah, A. (2018). Embedded librarianship in research in Nigerian Universities: Practices and sources of practice knowledge. *The Journal of Academic Librarianship*, 44(6): 738-746.
- Javed, K. and Shafait, F. (2017). Real-time document localization in natural images by recursive application of a cnn. In 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR) (Vol. 1, pp. 105-110). IEEE.
- Kavishe, G. F. and Dulle, F. (2016). Preservation skills and strategies of electronic information resources in digital era: Case of University of Kwazulu-Natal libraries. *Library Philosophy and Practice (e- journal)*. 1451. <http://digitalcommons.unl.edu/libphilprac/1451>
- Kim, H. and Maltceva, N. (2022). Digitization of Libraries, Archives, and Museums in Russia. *Information Technology and Libraries*, 41(4): 23-31.
- Matthias, H., Pomp, A., Erki, K. and Meisen, T. (2019). Data-Driven Recognition and Extraction of PDF Document Elements. *Technologies*, 7, 1-19. doi:10.3390/technologies7030065 www.mdpi.com/journal/technologies
- Mohammed, A. A., Saka, K. A., Babalola, G. A. and Ahmed, A. O. (2019). Access to digital materials on research output: Experience of Federal University Libraries users in northern states of Nigeria. *Samaru Journal of Information Studies*, 19(1): 77-97.
- Nasidi, N. A. and Jika, A. Z. (2023). Digital Archiving and the Establishment of Open Access Digital Repositories in Selected Nigerian Universities. *Library Philosophy and Practice*, 1-23.
- Nelson, G. and Ellis, S. (2019). The history and impact of digitization and digital data mobilization on biodiversity research. *Philosophical Transactions of the Royal Society B*, 374(1763): 20170391.
- Ong, S. Q., Mat Jalaluddin, N. S., Yong, K. T., Ong, S. P., Lim, K. F. and Azhar, S. (2023). Digitization of natural history collections: A guideline and nationwide capacity building workshop in Malaysia. *Ecology and Evolution*, 13(6): e10212.
- Oyewole, O. A. (2020). Toward Digitization of Files and Records in MDAs (Nigeria). Institute of Information Management (IIM). EDMS-Digitization Workshop-Attendees_Copy.
- Prescott, A. and Hughes, L. M. (2018). Why do we digitize? The case for slow digitization. *Archive Journal*, 1-15.
- Rafiq, M., Ameen, K. and Jabeen, M. (2018). Barriers to digitization in university libraries of Pakistan: A developing country's perspective. *The Electronic Library*, <https://doi.org/10.1108/EL-01-2017-0012>

Corresponding author: Moruf Hawwau A.

✉ hawwau.moruf@yahoo.com

Department of Library and Information Science, Federal University Dutsin-Ma, Katsina State.

© 2023. Faculty of Technology Education. ATBU Bauchi. All rights reserved



- Ramdhani, T. W., Budi, I. and Purwandari, B. (2021). Optical Character Recognition Engines Performance Comparison in Information Extraction. *International Journal of Advanced Computer Science and Applications*, 12(8): 21-30
- Sharma, K., Shingatgeri, V. M. and Pal, S. (2021). Role of data digitization on data integrity. *Quality Assurance Implementation in Research Labs*: 221-245.
- Towle, M. (2022). Digitization Guideline, Enterprise Information Management. <https://imtpolicy.sp.alberta.ca>
- Ugwuona, C. S. & Dike, V. W. (2015). Relationship between library resources and research productivity in five Nigerian health-based research facilities. *Library Philosophy and Practice (e-journal)*. 1267. Accessed on 21st November, 2017, from: <http://digitalcommons.unl.edu/libphilprac/1267>
- Zhou, Y., Zuo, S., Yang, Z., He, J., Shi, J. and Zhang, R. (2023). A Review of document image enhancement based on document degradation problem. *Applied Sciences*, 13(13): 7855.

Corresponding author: Moruf Hawwau A.

✉ hawwau.moruf@yahoo.com

Department of Library and Information Science, Federal University Dutsin-Ma, Katsina State.

© 2023. Faculty of Technology Education. ATBU Bauchi. All rights reserved