Digitising old Yoruba newspapers at Kenneth Dike Library

Oluwaseun Obasola[1] and Rukayat Atinuke Usman[2]

Abstract

The Kenneth Dike Library and the Nigeria National Archives are especially rich in ancient collections, particularly those unique to southwestern Nigeria, home to many people of the Yoruba extraction. These facilities house print and non-print materials such as personal notes and written collections of prominent persons, old manuscripts, ancient and modern maps, journals, and old Yoruba newspapers. Many of these print materials, especially the newspapers, are deteriorating. In a bid to prolong shelf-life, access to these old materials is limited. As newspapers serve as gateways to the past, this restricted access can impact the research experience of users.

The paper begins by presenting the project framework, which was designed before the project began. It goes on to detail the nuances involved in the several stages of the digitisation process and considers the aftermath of digitising the papers in terms of ownership, storage, backup, and access. This project revealed two things: first, though digitisation solves the problem of access and preservation, it is still necessary to preserve the original materials to prevent loss due to technical issues. Second, funding, and international partnership work hand in hand with digitisation, as it is a capital-intensive activity. Last, the paper contributes to the ongoing debates on the cultural, and socio-political discourses entwined with the technical processes of digitisation. The highlighted project was sponsored by the European Research Council (ERC) in collaboration with local partners. The website, https://yorubaprints.wordpress.com/yoruba-erc-project/ raises awareness for the project.

Keywords

Digitisation, Yoruba, Old newspapers, Kenneth Dike Library

Introduction

Libraries worldwide have been increasingly focused on digitising their collections. Digitisation solves the problem of preserving older materials (Liu, 2012), and allows libraries to make their collections available twenty-four hours a day, seven days a week (Hampson, 2001). This helps libraries to meet a fundamental demand of the twenty-first century.

The evolution from industrialisation to post-industrialisation has directed the world's interest to a knowledge-driven economy. Thus, there is a corresponding increase in the need to supply and access knowledge, both new and old. As for newly created written knowledge, writers and publishers concentrate on publishing e-books or articles instead of printing them in hard copies. Climate change has made it even more imperative to seek creative ways of publishing without paper. Driving the paperless vision conserves trees, and contributes to saving the planet. These are increasingly known as "born digital" resources.

The constant creation of knowledge, however, does not obliterate the relevance of old print materials. Their relevance is especially conspicuous in the humanities, where "old" knowledge is the foundation on which newer research can stand, providing anchors for future works. The all-important nature of relics and old material necessitates their preservation. It is for this reason that libraries, as repositories of knowledge, turned to digitisation, engaging technology to digitise cultural heritage. Some of these old materials include ancient manuscripts, pamphlets, personal notes and letters, newspapers, and so on.

However, what does the term "digitisation" mean? Though this article is primarily focused on documenting the technical and ethical processes and implications of the Kenneth Dike Library (KDL) experience in digitising old Yoruba newspapers, it is necessary to define how we use "digitisation" in this context. In most articles engaging digitisation, words like – photography, preservation, conservation, accessibility, online presence, digital humanities, and heritage – frequently appear, and are sometimes used interchangeably.

King (2005) points out that digitisation can be traced back to the twentieth century when photography debuted for research purposes. This implies that digitisation is not merely taking pictures of "endangered" materials for storage. Corrado and Sandy (2017) added to the list of what digitisation is not, in the first chapter of their book. Succinctly, digitisation is not merely about having backups, as backups are more about continuity and continuity can be guaranteed via various methods. Digitisation is not merely about increasing access, as access can be ensured by creating digital or even physical surrogates. Neither is digitisation an afterthought resulting in a rush for preservation departments to preserve their holdings. Rather, it is a carefully conceived process mediated by an intersection of experts in various departments of technical and policy teams. Though Grycz (2006) acknowledges that the term "digitisation" suffices in the realm of digital imaging because this kind of imaging involves capturing the pictures of images, he also asserts that in the context of the activities of archives and libraries, digitisation is a "process". It is a *process* or series of *activities* that involves taking images of materials, recording this data, refining these images as needed, storing them in databases, and multiplying or reproducing them to guarantee access and longevity. Digitisation can thus be placed at the intersection of these terms and activities.

The Kenneth Dike Library (KDL), named after the first indigenous Principal and former Vice-Chancellor of the University of Ibadan, Professor Kenneth Dike, was founded in 1948. Up until 1970, the Kenneth Dike Library was the major national depository for Nigerian publications. From June 1970, it has shared the privilege of a central depository library with the National Library of Nigeria which became the senior partner and the centre for the compilation of the national bibliography. The library contains approximately 700,000 volumes of books and receives over 6,000 separate journals and other serials. The library collection also consists of the Africana Collection, Arabic books, and manuscript, government publications, maps, publications ordinance, theses and staff publications. Through MacArthur Foundation's funding (2002-2007), Kenneth Dike library received great boosts in all aspects of its functions and services to the immediate community and beyond. The Kenneth Dike Library has a digitisation chamber provided by the MacArthur Foundation as a part of the Systems Unit in 2010. The KDL has more than a decade of experience in digitisation. It digitises both old and new publications. The library has large deposits of national and regional collections because of its former role as a national library before evolving into an academic library.

These numerous collections, with different risk exposure levels, create a strain on the digitisation chamber, which has limited human and capital resources. Many of these collections, which are stored on different floors and segments of the library, urgently need to be digitised so they are not lost forever. This article focuses on old Yoruba newspapers held in Kenneth Dike Library, the main library at the University of Ibadan, and at the Nigeria National Archives, Ibadan.

Beyond advocating the urgent need to digitise delicate materials, this article contributes to the advancement of the nascent field of digitisation in Africa. The Yoruba-prints project revealed that digitisation projects are largely isolated from each other, and also rarely engages in knowledge-sharing, consequently creating a paucity of relevant and contextual knowledge that can guide or provide illumination to other digitisation projects. Adversely, this has produced a vicious cycle of trial and error in digitisation projects, thereby slowing down the digital humanities drive in Nigeria.

This article provides detailed and hands-on approach to digitising delicate print materials. Further, it contributes to the global discourse in digital humanities by interrogating the philosophical debates underlying the politics of memory, heritagisation, and digitisation such as questions on what is digitised, how, why, and for whom. We discuss this in the section titled 'The socio-cultural and political questions in digitisation'. Finally, the article contributes to the future of digitisation in Nigeria, so forthcoming projects can learn from the successes and challenges of this project, thus furthering the frontiers of digitisation, specifically for projects dealing with fragile print materials and generally for any project involving the use of photographs.

The need to digitise and early efforts

Libraries with collections of old newspapers are increasingly faced with the dilemma of meeting user needs by providing access to these fast-disappearing materials, and also preserving the old materials for future use (King, 2005). These obligations are opposed to each other, as providing unrestrained access to already low-quality or fragile papers, would lead to the rapid deterioration of the materials, which would then shorten their lifespan, so they would not be able to reach future users. Mieczkowska and Pryor (2002), while giving a brief history of newspaper production in Britain, explained the reason for printing newspapers on (lower quality) wood-pulp paper. The shortage of linen and old cotton in the 1850s - the canvas formerly used to write/print newspapers, forced production to shift to using wood pulp for paper. However, wood-pulp paper had a crucial limitation. It contained a chemical property – lignin, which turned the paper's colour from white to yellow upon interaction with light. This yellowing resulted in reducing the paper shelf life. Though the production of acid-free paper through laser technology has removed the problem of yellowing papers, papers used for producing newspapers are still very brittle and susceptible, with the earliest papers facing the most critical threat of irredeemable destruction (Corrado & Sandy, 2017).

The period in which Britain started printing on paper was not too distant from when the printing press found its way to Nigeria via British missionaries. Thus, the type of paper used in printing newspapers in Nigeria was not much different from what was obtainable in Britain.

In ancient Islamic centres of learning, direct hand copying of manuscripts was the method scholars adopted to produce multiple copies, improve access, and preserve knowledge. The digital turn was heralded when more recent scholars, particularly in the Global North (Europe and America), used microfilms to preserve early newspapers. King (2014) asserts that microfilms were quite viable in attempting to preserve the content of newspapers because if the process of creating quality microfilms (i.e., using the photographed negatives of the newspapers) is properly followed, reeled

out, and stored afterwards, newspapers can then be accessed without damaging the microfilms. They also have other advantages, such as the ability to last more than two hundred years, minimal technical know-how, and reduced storage space consumption. However, microfilm as a medium has challenges too, such as the complicated procedure of loading the reels into reading machines, it is hard on the eyes to view for prolonged periods, and it is impossible to search texts except by manually reading every single page and also requires electricity.

In recent times in Nigeria, the need to preserve historical records and heritage materials has moved the staff of galleries, libraries, archives, and museums (GLAM) to explore several options in partnerships and methods of digitising forms and volumes of heritage materials. However, the sparse availability of literature documenting these efforts has left these digitisation efforts nearly unseen (Pickover, 2009). Thus, this article was conceived in response to the recognition of the importance of documenting and publishing the nascent digitisation projects in Nigeria, and also to Hauswedell *et al's* (2020) call for countries and institutions (especially in the Global South) to document the processes and dynamics of digitisation within their contexts. Further, this article contributes to the ongoing debates on the ethical, moral, cultural, and socio-political discourses entwined with the technical processes of digitisation.

The Socio-cultural and political questions in digitisation

Many authors and scholars in digitisation have called for a reflection on deeper level philosophical and political questions on what is digitised, how, why, and for whom they are selected to be digitised on the one hand, and where and how one can access the digitised materials on the other hand (Bianchi, 2006; Pickover, 2009; Ugah, 2009; Hauswedell et al., 2020; Zaagsma, 2022). This is important because the underlying discourse proves a relationship between archives (information in general), knowledge, social memory, and power (Zaagsma, 2022), which also introduces digitisation into the politics of heritagisation. This article engages these discourses by contextualising heritage and archives digitisation in two contexts. First, the Nigerian-Yoruba or African methods of historicising, and second, the global North-South power dynamics. Finally, similar to Zaagsma (2022), we achieve this by transposing the what, how, and why questions of digitisation in the traditional archiving processes to the digital domain; processes such as selection, collection development, cataloguing and classification, circulation and access[3].

In a bid to address these critical questions and advance the debates, we emphasise that this is not an attempt to justify any form of politicisation in digitisation efforts, nor to provide solutions to the complex issues, but to explain our reasons for and limitations in making these choices. The following sub-headings address these questions.

Selection – The criteria for selecting the materials to be digitised entails a re-selection out of pre-selected and already catalogued old newspapers in the KDL and National Archives. Therefore, determining what is picked and what is left out, according to Hauswedell et al. (2020), is inherently politicised. Why were old Yoruba newspapers selected for digitisation? Why not select the oldest newspapers from each Nigerian region, given the multiplicity of tribes in the country? Where are the earliest copies of published newspapers and why are they not part of this digitisation efforts?

With over fifty thousand volumes of 19th and 20th century newspapers, KDL, like many other organisations, does not have the required technical, staff, and financial strength to digitise everything. Selecting old Yoruba newspapers as the focus of this project was based first, on a hierarchy of needs and the requirement of a project funded by the European Research Council. Early newspapers are

among the most vulnerable print materials held in libraries/archives; thus, they are prioritised for preservation above other forms of print. Yoruba newspapers top this list because they represent the earliest history of local/vernacular newspaper production in Nigeria and Africa (Omu, 1967). Additionally, KDL and the National Archives in Ibadan are particularly vested with Yoruba newspapers, under the collection allocation strategy in Nigeria. Subjecting newspapers already in a precarious state to transport from one region of the country to another would further endanger them. Neither would it be cost-effective to transfer already over-burdened experts from Ibadan to other regions to digitise other regional newspapers.

In a world where what is available online is increasingly becoming "the only accessible" data, we are aware of how the availability of one set of data builds memory politics, particularly for coming generations (Britz & Lor, 2004). However, we hope that digitisation of other critical materials in Nigerian history will follow very quickly, and even more importantly, that those digitisation processes are documented and published. This would build a framework for a holistic analysis of digitisation efforts in GLAM institutions in Nigeria.

Classification and metadata – Metamorphosing catalogued print materials into digitised materials necessitates a re-classification of the newly digitised materials to ensure that they are findable and accessible online (Pickover, 2009). How can the newly digital materials be organised and described without reproducing the existing power asymmetry in knowledge politics, while maintaining a practical approach to findability? For searchability, we adapted the Dublin Core metadata schema for the digitised newspaper. Although the metadata has its own limitations (translational issues), it was quite easy to adapt since a previous project in the library had already successfully used the schema. We adapted nine elements out of the fifteen elements of the Dublin Core metadata schema: title, date of publication, publisher, keywords, description/abstract, naming conventions, type of document, and volume and issue number.

Translational challenges – The schemas of available metadata are not quite cultural-diversity-sensitive (Ma, 2020). Despite the knotty questions on translational challenges, and the issues we considered on the ethics of using summary or "umbrella" English words[4] as keywords; we planned to adopt the Dublin Core metadata schema to define the keywords in Yoruba and English because the intent of the digitisation project is to widen accessibility to all library users, some of whom may not be speakers of the Yoruba language. We figured out the 'best' English words that capture the meanings of some Yoruba words/phrases that Google Translate could not adequately capture. Again, since the metadata is purely descriptive, the keyword search can serve as a general navigation route for non-Yoruba speakers/readers, while an advanced search would find specific Yoruba words. Though this does not solve the problem, it offers an interim approach to coping with the wider problem of African knowledge production in an academic clime dominated by non-African research methodologies.

History of newspapers in Nigeria (Yoruba land)

The modern printed newspaper has a nearly 400-year existence (Ugah, 2009). Through this period, newspapers have been used for recording and storing information, happenings of people, events, and places. Apart from merely recording texts, they also contain photographs showing the everyday life of the people they serve. The ability of newspapers to provide investigative journalism, and educate readers, is partly because they have been owned by people with strong ideological drives. Newspapers have also served as agents of social change and revolution (Curras, 1987). Local newspapers are a

primary source of history and community data. Newspapers occupy a pivotal position in transmitting information on several levels (Onwubiko, 2005). They transmit information from the past to the present and provide regular and up-to-date information (Ugah, 2009). They can show the gaps and strides of a society, and they can play entertainment roles. Mostly, newspapers are a rich source of information about politics, ideas and thoughts, commercial, economic, legal, and just about everything about the era they were published. Their relevance to issues attracts a wide body of researchers.

Newspaper production in Nigeria began in the late nineteenth century in south-western Nigeria, among the Yoruba. Thus, the earliest history of newspaper publication in Nigeria parallels the history of Yoruba newspaper production. The term 'Yoruba' describes a language of the Niger-Congo family, and the people who speak it. They predominantly occupy the current southwestern part of Nigeria, a part of Benin Republic and Togo (Ogundiji, 2003; Barber, 2014). Within Nigeria, the Yoruba are famous for ancient art works such as the Ife bronze figures, and a history replete with pockets of external influences; from Arabic traders and scholars, and later on, Western/Western-formed missionaries. The Arabic influence led to the foremost attempt to put Yoruba in written text. Thus, the earliest form of written Yoruba was formed by adapting an Arabic script called Ajami before the 1850s (Olumuyiwa, 2013). However, Ajami neither reigned for long nor spread widely. Extensive Yoruba texts began with the early stages of colonisation, in the mid-nineteenth century, with missionaries of the Church Missionary Society playing pivotal roles such as writing the tonal language in Latin text (Ogundiji 2002; Olumuyiwa, 2013). The periods between the entrance of the printing press, the development of Latinised Yoruba and inception of Yoruba newspapers coincided roughly around the mid-nineteenth century.

The entrance of the newspaper in Nigeria began with the establishment of *Iwe Irohin fun awon Egba ati Yoruba,* (popularly known as *Iwe Irohin*), a Yoruba newspaper based in Egbaland, Abeokuta, South-western Nigeria (Omu, 1967). It was founded by a member of the Anglican laity, Rev. Henry Townsend in 1859 and was published every fortnight (Omu, 1967). Its content was limited to church activities and social events. It neither reported crimes nor entertainment. The advent of Christian missionary activities, across south-west Nigeria, Abeokuta, in particular, steered the content and language towards people represented in these communities. This created a number of limitations. First, it was limited to the Yoruba elite, as formal education was not yet commonplace in the 1850s. Second, its concentration on the people of Egba, and then the Yoruba race, generally, meant the side-lining of other regions. However, these problems were justifiable because Nigeria as we now know it, was not established. Its regions were autonomous.

As the church was primarily established for religious purposes, it held no obligation to print judicial, crime, or entertainment features in its paper. Despite this, it featured strong opinions on anti-slave trade and promoted civil rights and education (Maringues, 2001). Omoloso and Abdulrauf-Salau (2014) assert that the newspaper, though not the first in Africa, is known to be the first “vernacular” newspaper on the continent. It was a widespread notion in colonial Nigeria and indeed, Africa, that local or indigenous languages were vernacular, inferring that they were inferior to the English language and other colonial languages. Reverend Townsend, initially did not appear bothered about the second-fiddle reference given to the newspaper he founded because “his motive of establishing a local language newspaper was in keeping with some of the main functions of the press, which are to inform and educate” (Omoloso and Abdulrauf-Salau, 2014) and to “beget the habit of seeking information by reading” (Alabi, 2003). In 1860, just one year after publication started, an English

translation was attached to the newspaper, making it a weekly published bilingual paper. Sadly, less than a decade later, the publication came to an abrupt end. This unfortunate development, as opined by Salawu (2004) paved the way for many other newspapers - indigenous, bilingual, and English. For example, *Iwe Iroyin Eko* which started in 1888, published monthly by Andrew Thomas was the immediate successor of *Iwe Irohin* (Omoloso & Abdulrauf-Salau, 2014). Newspapers written solely in the English language were already being published before *Iwe Irohin* folded. In 1863, *The Anglo African*, edited by Robert Campbell, was the second published newspaper.

While the nineteenth century heralded the production of newspapers in Yoruba land and Nigeria as a whole, the twentieth century represented the glory days of Yoruba news production. In 1922, Adeoye Deniga founded *Eko Akete*, a newspaper that lasted for more than a decade and was far more diversified in content compared to earlier publications. This project engaged part of the earliest printed *Eko Akete* from 1922 into the 1930s. *Eleti Ofe* followed in 1923, *Iwe Iroyin Ososе* and *Eko Igbein* in 1925, and *Akede Eko*in in 1927.

These daily and/or weekly papers were not intended to serve as a permanent storage of information (Stoker, 1999). Thus, many are now in a deplorable condition and are being eaten up by worms and insects. Librarians and researchers have urgently searched out means to preserve them.

Pre-digitisation framework at Kenneth Dike Library

According to Holley (2003), strategic planning before undertaking any form of digitisation project is important because digitisation affects some critical aspects of the library, such as current infrastructure - ICT equipment and software, organisation structure, staffing, and service delivery. Before the project commenced on the 7th of September 2021, a framework to determine the number and variety of collections held by the library was drafted, with projected work procedures and timelines designed. The framework was agreed upon by the library, sponsors, and other partners for the project. Procurement of new equipment and employment of ad-hoc staff were also gradually put in place from about three months before the digitisation process began.

1. Number and variety of newspaper collections

The Kenneth Dike Library and the Nigeria National Archives, Ibadan, are both historically significant as holders of old and rare collections in the country. The University of Ibadan is the premier university in Nigeria and among the earliest in West Africa with the capability and capacity to hold many significant collections. The city of Ibadan is also relevant as a publishing hub, and historically, as a metropolis of the Yoruba, in proximity to other big towns famous for publishing early Yoruba newspapers, such as Lagos and Abeokuta.

The library has more than fifty thousand volumes of old newspapers, and about five thousand are old Yoruba newspapers. Out of this number, those most vulnerable to extinction were first selected for digitisation. An opportunity cost model was used to maximise the limited available resources as digitisation is labour and capital-intensive. Samples were picked from the 1920s (the year of the Yoruba newspaper publication boom) into the 1970s when indigenous newspaper publication was already on a precarious and steep decline. Samples include:

- *Eleti Ofe*- 1950 to 1965
- *Gbohun-gbohun*- 1971
- *Imole Owuro*- 1974 to 1978
- *Irawo Obokun*- 1955

- *Irohin Yoruba* – 1951 to 1973
- *Irohin Imole*- 1952 to 1966
- *Irohin Onigbagbo*- 1963/64

These volumes were packed in labeled boxes in the reference section of the Kenneth Dike Library. Some boxes did not contain the complete publications for the marked year because some publishers were inconsistent, and other volumes were simply unavailable.

During this stage of determining the numbers, type, and state of the collections available, it did not occur to the researchers to check for duplicates. It was later discovered upon the commencement of the project that some issues were repeated thrice.

Collections from the archives, on the other hand, did not have any cases of duplication. They were bound; backdated and had different publications from the collection at Kenneth Dike Library. Samples drawn from there include:

- *Eleti Ofe*- 1924 to 1927
- The Yoruba News – 1924 to 1928
- Eko Akete – 1922 to 1925
- Akede Eko – 1931 to 1937

In total, this stage took about 2 weeks.

2. Projected work equipment, procedure, and timeline

The next stage addressed the technical and systems-related concerns. This included the type of software and equipment that would be compatible with the project and what would produce the best outcome. Since the Digitisation Chamber of KDL had been engaging in digitising recent print materials before this project, the choice of software intuitively had been resolved. However, the decision on the exact type of equipment to be used was determined by the nature of the collections. Scanning using a sheet-feeder scanner, for example, was impracticable because of the size and brittle nature of the old newspapers. Also, loosening the bound collections would subject the materials to deterioration. While the use of a flatbed scanner may have been helpful, the massive size of some of the old newspapers made this impossible. *Gbohun-gbohun* 1971 for instance, measures 53cm x 35.5cm when closed, which was impossible to capture even using a large flatbed scanner which measures 45.5cm x 30cm without cutting out some parts of the large newspapers.

The content of the metadata that would be generated to describe the collection was also considered at this stage. The Dublin core metadata schema used by the University of Ibadan for digital objects was adapted for the newspaper digitisation project.

3. Approval Processes

The approval obtained for the project was in a two-part policy brief – internal and external. The internal aspects took keen consideration of the organisational aspect of the project. The collections were to be retrieved from the reference section of the Kenneth Dike Library, meaning that an internal memo needed to be written to coordinate activities between the reference and digitisation sections. Similarly, coordination was required to get approval to work with the collection from the archives. As there was no digitisation policy in place at the University of Ibadan, the project team had to work with

the digitisation manual/guidelines developed by the System Unit Team at the Kenneth Dike Library. The KDL digitisation manual was followed throughout the duration of the digitisation project.

4. Procedure for digitisation

After taking a long-range view of the process, we developed an action plan to achieve the digitisation of the newspapers. It is important to note that the framework drafted was constantly updated, as the plan and the reality of executing the plan turned out to be different. For example, the framework considered using the software, Macromedia Fireworks for editing the captured images of the newspapers. However, after a test run, it was discovered that the newspaper titles edited using this software were not accessible after converting them to TIFF. Simply put, they could not be viewed. This necessitated reviewing the software and switching to using Lightroom for editing the images. This does not mean that the Macromedia Fireworks software is not suitable for digitisation. Other digitisation works such as digitisation of book materials and other texts, carried out in the digitisation chamber of the Kenneth Dike Library are edited and cleaned using this software. The processes below sequentially describe how the old Yoruba collections were digitised, while simultaneously noting the lessons learned on the way.

i. Set up – This project began by setting up the purchased hardware components. Except for the copy stands, cameras, laptops, and Uninterrupted Power Supply (UPS), other equipment used in the project was locally sourced. The capturing/placement boards and laptops were bought in the Netherlands and shipped to Nigeria. They include:

   - Uninterrupted Power Supply (UPS)
   - Standing LED lights
   - Copy stands
   - A server
   - Laptops
   - Cameras
   - Solar power installations
   - Router

The project did not start with the procurement of all the listed equipment. For example, the server, solar power installation, and laptop arrived several weeks after the commencement of the project. The project commenced with the initial procurement of the following materials: Two Canon 4000D digital cameras, two placement boards, and two laptops. Later, a router for internet connections, solar installations for constant electricity, and a server were procured. A pack of nose masks and a bottle of hand sanitizer were procured for the staff to observe COVID-19 protocols. It was important to protect the equipment from the dust on the old volumes of papers. A space within the digitisation chamber was cleared to make room for the new equipment. The capturing boards were laid on a regular worktable. Each board had a stand attached to the top of its mid-section, where a digital camera was fixed. The camera was then connected to the laptop/server through a wire connection. Before the cameras were fixed, the settings such as the light scale, size, etc. were set on the same parameters for both cameras as a quality control measure, to ensure similar output. Factors like different lighting in the areas where the two capturing boards were placed, weather Conditions such as rainy seasons, when it got dark due to clouds, and most importantly, the different hues of yellow, brown, and grey of different newspapers meant adjusting the camera settings until we got a desired pictorial result. Also, acquiring devices from outside the country, i.e., the laptops, was not a promising idea, as just

two weeks into the project, they malfunctioned. The laptops tended to overheat while in use, perhaps because of the climate differences between tropical Nigeria and the cold.

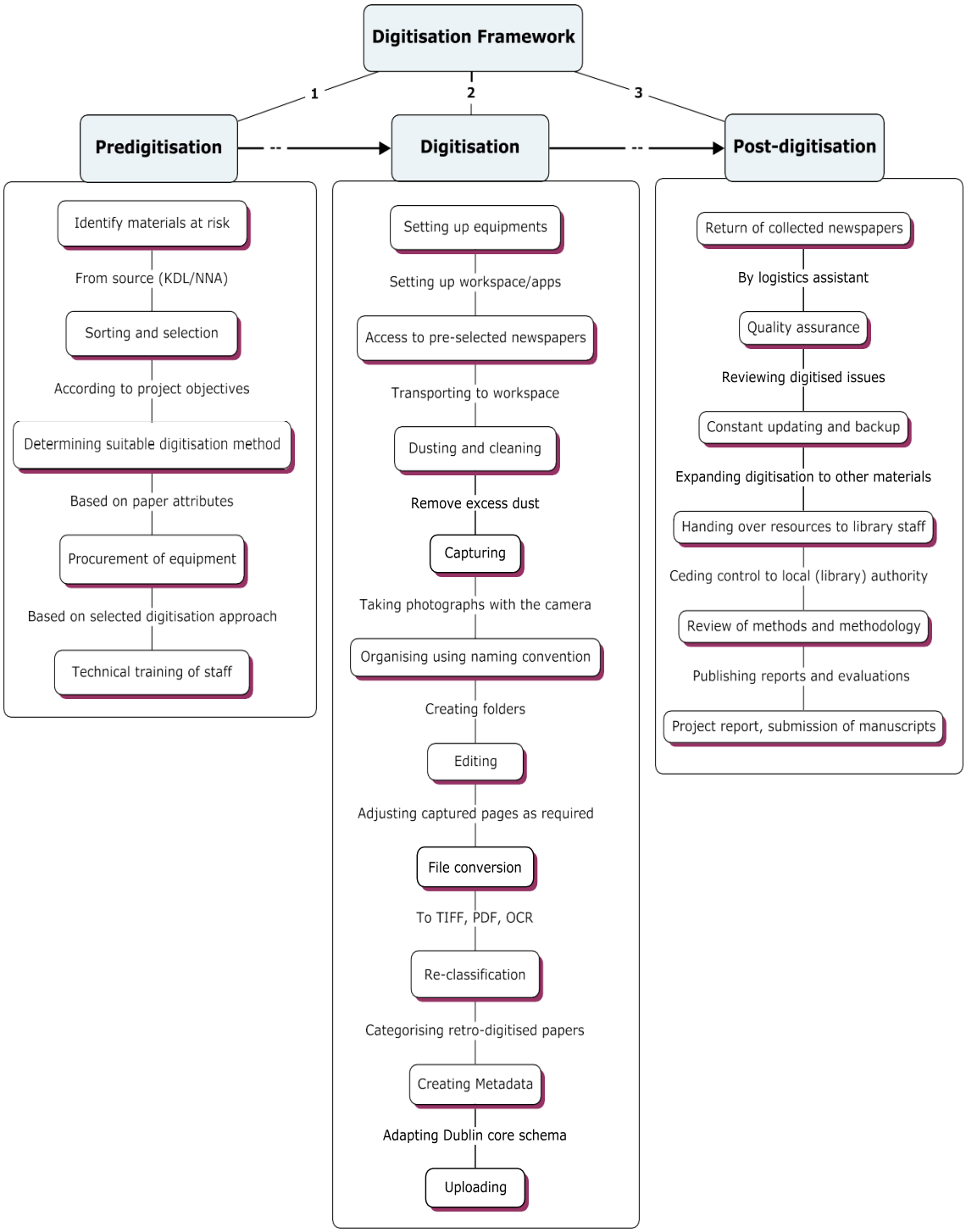

Figure 1: *Schematics showing the framework adopted for the digitisation of Yoruba Newspapers at KDL*

Figure 2: *A camera, copy stand, two light stands and a computer system (server) with UPS used for the digitisation of Yoruba newspapers at the Kenneth Dike Library, University of Ibadan (2021).*

Software - Several kinds of software were bought and installed on the computers used for the execution of the project. They include:

- Lightroom - Used for editing and converting images captured with the cameras into cleaner versions.
- Adobe Acrobat 7 - Converts images in TIFF format (Tagged Image File Format) as saved in Lightroom, to PDF (Portable Document Format).
- AbbyyFineReader - Files are converted into searchable files. It changes the newspaper form from images into text (i.e., OCR - Optical Character Recognition).
- Microsoft Office Applications - Word and Excel were employed. Excel was used for the collection of the metadata gathered from the newspapers while Word was used to type documents.
- Anti-virus - "Bitdefender" was bought to protect the computers and files from viruses.

ii. Acquisition and dusting - The newspapers or print materials were first acquired from the National Archives and/or the reference section of the University of Ibadan library. They were then cleaned

with soft dry towels to remove dust, carefully spread out to stretch wrinkles from squeezed papers, and placed on the capturing boards.

iii. Sorting and capturing - The newspapers were stored in boxes in the reference section. Each box was examined to check for duplicate copies and to confirm the year(s) of publication. The sorting of duplicates is another way in which the planning stage differs from that of the execution. While compiling boxes for this project, the labels on the boxes were considered to be accurate and properly sorted. But on commencement of the project and more thorough examination, it was discovered that some newspaper issues had many duplicates, sometimes, even packed into different boxes. This slowed down workflow for a while, as sorting duplicates was not designed in the initial framework drafted before the project began. After the duplicates were removed, the remaining issues were sequentially arranged so each staff could clearly mark the issues they discovered. This made later collation of all the captured images easier. One after the other, the newspapers were placed on the capturing boards. The camera was then set at a height that adequately covered the face of the print material and a picture was captured using the Canon 4000 D camera. The image was saved in a document format known as RAW or CR.2. To avoid a mix-up among the newspaper issues, separate folders were created for new newspaper brands and their different issues.

iv. Renaming – Cameras automatically title the file of the images taken during the capturing process. However, these titles do not reflect the standardised naming convention designed for each volume. Each captured image/file had to be renamed to reflect the naming convention. The naming convention involved using the essential details of a newspaper in a shortened format. For example: "EO_JUN_17-23_1961_XIII_694" represents – *Eleti Ofe* (newspaper title), June 17 to 23 (date), 1961 (year), XII (volume) and 694 (issue). Codes like this were designed for each type and volume of newspapers digitised.

v. Editing – It involved cropping, cleaning, and sometimes rotating the pages of the captured newspapers until they appeared in the standard form (portrait and straight view). Lightroom software by Adobe was utilised for this stage. Each page was cropped, cleaned, lightened, and adjusted as required, and then compiled by issues in different folders. A crucial factor that needed harmonising during editing was the setting of pre-sets. A pre-set is the adjustment of all elements, such as texture, shadow, brightness, and clarity. These options in the Lightroom software were saved and titled according to each newspaper brand. A pre-set is usually designed on the server and exported to each laptop using the router, so they are accessible on multiple laptops for use by other staff. This is to control quality and ensure a uniform appearance for all the files. However, this approach was plagued with challenges. The yellow shades of the newspapers, even if they were of the same title, were different. Some appeared yellowish-brown, while some tilted towards grey and orange. In fact, various hues were observed across different papers, sometimes contained within a particular volume. This meant that the pre-set had to be manually adjusted, issue by issue, and page by page, to produce a good result. The advantage of this manual process is that after a while, one becomes familiar with the elements of the software and the newspaper. Knowing what to adjust, and to what degree, becomes relatively easy. It must be noted that it was easier to edit on the server than on other laptops used for the project. This is because the server hosts all of the original captured newspapers. The other laptops needed to be connected to the server via the router, or directly copied using a hard drive to the memory of the laptops, and then had to be re-transferred to the server after editing. Hard drives were used for copying instead of flash drives because the latter are highly susceptible to corruption. After editing, Lightroom saved the files in TIFF.

vi. Converting to searchable format - The TIFF format in which Lightroom saves the files cannot be viewed by some image or document viewers nor can they be seen when readers search texts. To enable these features, the pictures were first converted into PDF using Adobe Acrobat 7.0 and then converted from an image to searchable texts using Optical Character Recognition (OCR), also known as searchable PDF, with AbbyyFineReader. Each of these formats was carefully collated into different

folders and labeled appropriately. Early versions of the files, i.e., CR.2/RAW and TIFF were not discarded. They were kept as a check in case of any dysfunction that might arise while transforming file formats and as an additional backup. After conversion to OCR, the team vetted the searchability of the words. Searches using the English keyboard did not yield particularly distinct search results. We attribute this to the Latinised rendition of the Yoruba alphabet by the software.

vii. Collation of metadata - Metadata is an important part of information storage, sorting, retrieval, and use. The thought of having to retrieve a piece that has no title, or any form of description is mentally challenging. This illustrates the necessity of generating metadata for collections. According to Corrado and Sandy (2017), metadata is structured information about a resource. Debates exist among digital librarians most especially, on the classification of metadata, but Corrado and Sandy (2017) stick with four types: descriptive, technical, administrative, and structural metadata. This project was concerned with descriptive metadata, the basic information that makes it easy to retrieve a document. Upon completing the digitisation of a box or bounded copy, the information/metadata from the newspapers was extracted and organised on Excel spreadsheets. The details include the title of the newspaper, date of publication, publisher, keywords, description/abstract, naming conventions generated for each paper, type of document, and volume/issue number. Each Excel spreadsheet had columns created for each piece of information. Separate Excel files were created for each newspaper brand, while separate sheets within a file were opened for different newspaper years (year of publication), to make for a neat and clear organisation. To save time, metadata could be created by a staff thereby creating specialization in the workflow. In the alternative, each staff engaged in the digitisation process could enter metadata into a pre-designed template and convert the same to PDF. The latter method was used in this exercise.

viii. Uploading digital objects into the institutional repository (IR) - This involves the compilation of the final format of each file (i.e., OCR) and collecting their metadata to be inputted into the online institutional repository of the Kenneth Dike Library, University of Ibadan. The online IR of the university is hosted by DSpace and administered by the staff of the systems unit of the library. It is helpful to have a stable Internet connection, to facilitate prompt uploading of files, for access at any time, from anywhere in the world.

Challenges

The digitisation of old Yoruba newspapers at the Kenneth Dike Library had many challenges. The challenges are highlighted as follows:

1. Image recognition deficit – The pages of newspapers are usually dotted with several images and photographs which may add meaning to a written piece or stand as its own body of knowledge. This project was technically limited in having fully detailed images like the OCR done on texts.
2. OCR limitation – Significant variation in font sizes; elaborate or tiny font, and blurred texts, often make it difficult for the OCR to conduct a credible search of the affected words. This could lead to a false search. The settings did not affect the reading of the characters when converted to OCR. However, when we tried searching for keywords in Yoruba with ascent marks using the search box, it was a bit difficult to get accurate results. Searching using the installed Yoruba keyboard helped.
3. Different screen types of the computer systems used - A total of five computer systems were used, though not more than three were active at a time. None of these systems were the same. They had different specifications such as screen resolution, capacity, dpi/ppi (dot per inch/pixel per inch), and so on. Also, the laptops were business editions; they were not graphics-dedicated nor did they have Nvidia (a feature that is usually on graphics computers). This caused occasional software problems. For example, pages could suddenly re-arrange themselves in a mixed form of

portrait and landscape view, the software could hang up for a while or even suddenly close or stop responding.

4. Irregular columns – The text layout in columns of some of the newspapers is sometimes bent, bending into the page borders, even outside of printed areas. This makes cropping and page size harmonisation a little difficult. Thus, we sometimes had no choice but to leave some pages and texts bent.
5. Loss of text - Many reasons account for this: segments eaten up by bookworms, brittle and lost parts of the papers, creased papers (this also sometimes affects OCR), and the spine curvature of bound volumes may make text difficult to be captured clearly.
6. Indexing - Indexing of these old papers is laborious for several reasons. Many of the papers lack consistency in frequency of publication. For example, within the archival materials, one can find *Eleti Ofe*, a supposed weekly publication as written on the paper, dated 17th to 23rd of February 1924, and the next issue is dated 19th of February 1924. The publishers sometimes repeated issue numbers when this occurred. The publisher and editor could also change abruptly. These subjects are relevant to enable a keyword search and systematic arrangement of these materials. This makes metadata collation very tasking and important. Publisher inconsistency thus limited our ability to provide an authoritative digital data navigation keyword guide for users. These problems also made the use of indexing software impracticable.
7. Format migration issues - This is mostly caused by software problems. The page orientation and colour consistency of some newspapers could be distorted while changing them from one format to the other. The page arrangement also gets distorted if the pages exceed nine (9) pages while editing on the Dell XPS laptop. They are also varied while copying from one system to the other, especially if the different formats are archived on several different systems/computer units. Data could be lost in the format migration process.

Conclusion

Digitisation is only the first step in the efforts being made by scholars towards digital preservation. Corrado and Sandy (2017) corroborate this as they stated that digital materials are unlike "physical collections of books, manuscripts, or artifacts that can be neglected for years without significant loss or additional expense," because digital materials can be affected by bit rot, software malware and many other types of "digital accidents and disasters".

The first step towards preserving the digitised items is to store them on an open-source digital library software such as the one already employed by the Kenneth Dike Library, DSpace. It continues with digital asset management as technology is constantly evolving. File formats, for example, may become obsolete at any time. With an already set up stand-by digitisation chamber/unit at the Kenneth Dike Library, the digitised materials are unlikely to become obsolete without warning. Staff are trained and re-trained in the face of technological development and trends. These staff, in conjunction with the systems unit, keep a watchful eye on these trends.

Repository maintenance/administration, access provision, access control, and user support are provided by the systems unit team of the university library, thus, reasonably securing the future of the digitised materials.

Intellectual property rights - Data protection regarding this project is not typically concerned with the local distribution of these newspapers, either in physical or digital form. It is concerned rather, with the protection of a national and cultural heritage. In local settings, the old Yoruba newspapers were published and widely circulated particularly in the Western parts of Nigeria; their content was meant to be seen by its audience, not kept away for covert use, thus, there is little to no copyright breach with the digitisation and local circulation of the papers. However, these papers have evolved from

being mere articles carrying details of the everyday significant happenings of the past to becoming subjects of national and cultural heritage. Thus, opening access to them beyond the community they were created to serve, might be crossing the line of copyright and intellectual property rights. In putting this reality into perspective, the digitised collections are not only kept online but in the open-source institutional repository of the Kenneth Dike Library. The library has the purview of granting access and access control. We are aware of the limitations this could create, such as debarring access to users outside of KDL, thus limiting the overarching purpose of accessibility. Or the issue of excluding this heritage from the pool of digitised global resources or heritages. This is an emerging and rapidly evolving sector, especially in Africa. Hopefully, as newer and better methods emerge, holistic solutions to the complex issues in digitisation can be adequately addressed.

Copy-right law

There was no need to seek permission from authors of articles in the newspapers or the publishers; as the library is protected by the Fair Use Act under the copyright law (Second Schedule), as long as the materials digitised are used exclusively for teaching, learning, and research. Further, most of the newspapers digitised are in the public domain as the date of first publication for some was over 50 years.

Acknowledgement

The project was funded by the European Research Council Starting Grant 2020 – "Yoruba Print".

References

Alabi, S. (2003). The development of indigenous language publications. *Issues in Nigerian media history: 1900- 2000AD (*Akinfeleye R. and Okoye I., Eds.*).* Lagos, Nigeria: Malthouse Press Limited. Anthropology and Education. Philadelphia: University of Pen.

Barber, K. (2004). Bibliographical Survey of Sources for Early Yoruba Language and Literature Studies, 1820-1970. *Research in African Literatures*, *35(1)*, 203-204. https://doi.org/10.1353/ral.2004.0004.

Bianchi, C. (2006). Making online monuments more accessible through interface design. *Digital Heritage—.*(pp. 445–466). Routledge

Britz, J., & Lor, P. (2004). A moral reflection on the digitization of Africa's documentary heritage. *IFLA Journal*, *30*(3), 216–223. https://doi.org/10.1177/034003520403000304

Corrado, E. ., & Sandy, H. (Eds.). (2017). *Digital Preservation for Libraries, Archives, and Museums* (2nd ed.). Rowman and Littlefield.

Curras, E. (1987). Information as a fifth vital element and its influence on the culture of the people. *Journal of Information Science*, *13*(3), 27–36.

Grycz, C. J. (2006). Digitising rare books and manuscripts. In *Digital Heritage* (pp. 33–68). Routledge

Hampson, A. (2001). Practical experiences of digitisation in the BUILDER Hybrid Library Project. *Program: electronic library and information systems*, *Vol. 35 No. 3, pp. 263-275* https://doi.org/10.1108/EUM0000000006950

Hauswedell, T., Nyhan, J., Beals, M. H., Terras, M., & Bell, E. (2020). Of global reach yet of situated contexts: an examination of the implicit and explicit selection criteria that shape digital archives of historical newspapers. *Archival Science*, *20*(2), 139–165. https://doi.org/10.1007/s10502-020-09332-1

Holley, R. (2003). Developing a digitisation framework for your organisation (feel the fear and do it anyway). *Lianza*, *October*, 1–8. https://doi.org/10.1108/02640470410570820

King, E. (2005). Digitisation of newspapers at the British Library. *The Serials Librarian*. *49*, *1–2*, 165–181. https://doi.org/10.1300/J123v49n01_07

Liu, A. (2012). The state of the digital humanities: A report and a critique. *Arts and Humanities in Higher Education*. *11*, *1–2*, 8–41. https://doi.org/10.1177/1474022211427364

Ma, R. (2020). Translational challenges in cross-cultural digitization ethics: the case of Chinese Marriage Documents, 1909–1997. *Libri*, *70*(4), 269–277. https://doi.org/10.1515/libri-2020-0088

Maringues, M. (2001). The Nigerian Press: Current state, travails and prospects. In K. Amuwo, D. C. Bach, & Y. Lebeau (Eds.), *Nigeria during the Abacha Years (1993-1998)* (1–). IFRA-Nigeria. https://doi.org/10.4000/books.ifra.640

Mieczkowska, S., & Pryor, K. (2002). Digitised newspapers at Norfolk and Norwich Millennium Library. *Collection Building*, *21*(4), 155–160. https://doi.org/10.1108/01604950210447395

Ogunbiyi, I. A. (2003). The Search for a Yoruba Orthography since the 1840s: Obstacles to the Choice of the Arabic Script. *Sudanic Africa*, *14*, 77-102. http://www.jstor.org/stable/25653397

Olúmúyìwá, T. (2013). Yoruba writing: standards and trends. *Journal of Arts and Humanities*, *2*(1), 40-51. https://doi.org/10.18533/journal.v2i1.50

Omoloso, A. I., & Abdulrauf-Salau, A. (2014, May 4). *Indigenous Language Newspapers in Nigeria from 1914–2013: A Review*. In Amalgamation National Conference of the Department of Political Science & Department of History and International Studies, Ibrahim Badamasi Babagida University, Niger State, Nigeria.

Omu, F. I. A. (1967). The "Iwe Irohin", 1859-1867. *Journal of the Historical Society of Nigeria*. *4*, *1*, 35–44. https://www.jstor.org/stable/41971199

Onwubiko, P. (2005). Using newspapers to satisfy the information needs of readers at Abia State University Library, Uturu. *African Journal of Education and Information Management. 7*, *2*, 66–88.

Pickover, M. (2009, July 1). *Contestations, ownership, access, and ideology: policy development challenges for the digitization of African heritage and liberation archives*. 1–10. First International Conference on African Digital Libraries and Archives (ICADLA-1), Addis Ababa, Ethiopia. https://wiredspace.wits.ac.za/server/api/core/bitstreams/f76d3475-e18a-4a48-9ab8-0fbb605a49d0/content

Salawu, A. (2004). The Yoruba and their language newspapers: Origin, nature, problems and prospects. *Studies of Tribes and Tribals*, 2(2), 97-104. *2*, *2*, 97–104. https://doi.org/10.1080/0972639X.2004.11886508

Stoker, D. (1999). Should newspaper preservation be a lottery*? Journal of librarianship and*

information science. *31*(3). https://doi.org/10.1177/096100069903100301

Ugah, A. (2009). Strategies for preservation and increased access to newspapers in Nigerian University Libraries. *Library Philosophy and Practice*(e-journal), 270. https://digitalcommons.unl.edu/libphilprac/270

Zaagsma, G. (2022). Digital History and the Politics of Digitization. *Digital Scholarship in the Humanities* 38(2), 1–22. https://doi.org/10.1093/llc/fqac050

Endnotes

[1] Oluwaseun Obasola is the Digitisation Librarian at the Kenneth Dike Library, University of Ibadan, Nigeria. She can be reached by email: seunobash@gmail.com

[2] Rukayat Atinuke Usman is a postgraduate student at the Institute of African Studies, University of Ibadan, Nigeria.

[3] The problem of access was implicated in many sub-headings but fully addressed towards the end of the article.

[4] For example, in the abstract of each newspaper volume, words like “Yoruba culture” were used to describe a range of events already tagged in the Yoruba language like the Egungun festivals, naming ceremonies, coronations, excerpts on indigenous hairstyles, and many others.