DESIGN AND DEVELOPMENT OF DIGITAL MANUSCRIPT REPOSITORY

BY

ZAKARI MOHAMMED PH.D
DEPARTMENT OF LIBRARY AND INFORMATION SCIENCE,
AHMADU BELLO UNIVERSITY, ZARIA

A Paper Presented at the International Conference on The Problems And Prospects of Public/Private Arabic Manuscript Repositories in Nigeria held at Arewa House, Kaduna on 15th July, 2010.

ABSTRACT

The paper examined the place of manuscripts as primary sources of information service delivery and socio-political, economic and knowledge acquisition in various subject areas of academic disciplines. It highlighted the need for their digitization, establishment of Digital Manuscript Repository (DMR), taking note of its structure, content, services, systems requirements and sustenance. It concludes by emphasizing the need to digitize and house manuscripts in networked digital repositories to prevent them from extinction and unduely inaccessible and also ensure their local, national and global access and use.

INTRODUCTION

Manuscript can simply be referred to as unpublished document containing vital information on specific theme in a given subject area of

interest or specialisation. As primary source of unedited information, manuscripts are useful for conveying and retaining specific pieces of information useful to students, scholars, researchers, administrators, etc in their academic, research, administrative, socio-economic and political activities. Essentially, the extent of the usefulness of a manuscript largely depends upon the theme, subject, content, coverage, and the audience it addresses. Implicitly therefore, there is no limitation as to the type of subject coverage or theme of a typical manuscript content since the various authors, writers or composers only wrote in their areas of interests or specialization to address specific issues of interest and expertise.

The emergence of manuscript writing and manuscript collection especially among Islamic scholars, poets, administrators, students, etc pre-dated the advent of conventional printing presses. It serves as veritable source of scholarly communication and dissemination of vital information among the stakeholders as well as for storing the recorded knowledge acquired from variety of sources. Manuscript writing and collection has since time immemorial been conceived as part and parcel of academic activity and scholarly advancement, especially in the Islamic world as by-product of the teachings and practices of the Islamic religion. Hence, the prevalent of Arabic Manuscripts in Islamic societies, education, institutions,

mosques, palaces and among scholars, students, missionaries and traders. Evidently, the manuscript collections constitute such outfits as private libraries, school libraries, mosque libraries, etc containing variety of manuscript formats and contents covering various subject areas of interest and expertise.

NEED FOR DIGITISATION OF ARABIC/AJAMI MANUSCRIPTS

There is no doubt in the fact that manuscripts, especially the Arabic/Ajami manuscripts are in the threshold of getting extinct and unduely inaccessible continually due to threats of theft, mutilation due to excessive use, inheritance due to death, over possession and protection by family owners, and human and natural disasters leading to their under utilization.

The major objectives of digitising Arabic/Ajami manuscripts and indeed any other type of manuscripts of any sort is to secure, preserve and conserve them in different media and formats for easy access and use regardless of time factor. Manuscripts can be stored in such media as tapes, discs, optical discs, microforms, films, etc. However, with the advances in Information and Communication Technology (ICT), manuscripts can be digitized and stored in computer files, dataware house, CD-ROMs and

online computer networks such as the Internet System and services, LAN and WAN, portals, online databases and world wide web, etc.

Manuscripts are digitized –

- to prevent the effect of overuse of the documents, and incidence of changes in the content;
- to elongate the life span of the text and ensure its continued usability;
- to ease validation and retrieval of the original text;
- to ease mobility of the text thereby facilitating remote access;
- to facilitate multi-access by users especially when its resident in network environment;
- to facilitate duplication, formatting, dissemination and acquisition of the document;
- to retain the quality of the document;
- to facilitate widening of the manuscript holdings among the stakeholders;
- to prevent loss and destruction of the physical document;

- to ensure the safety and return of the original manuscript to the owner, etc.

DIGITAL MANUSCRIPT REPOSITORY (DMR)

The conceptual share sizes of Arabic Manuscript collections in different sites and with different individuals. scholars, families, communities and institutions vis-à-vis the likelihood of losing them continually especially due to incidences of disasters, theft, mutilation, and inheritance as a result of death points to the urgent need to not only digitize them but continually to also establish Digital Manuscript Repositories (DMR) to house them electronically, especially in network environment for easy access and use irrespective of location and time. The DMR should serve as manuscript digital archive for the preservation and dissemination of Arabic Manuscripts and indeed any other relevant manuscripts for existing and future access and use. The DMR is in variance with the paper-based archive that is prone to undue challenges of ownership, multi-access and use; storage; security; preservation; retention and backup. Rather, it is a digital system where digital assets and contents are electronically housed and managed for existing and future users to easily access their contents in simple or multiple sites/stations especially in network regime. Essentially

therefore, the DMR could be said to be a set of services for housing, managing and making available digital manuscripts regardless of their source, owner, contents and language of communication for access and use.

Ding (2000) in Usman (2007) elaborated on the works of Line (1996), Getz (1997), and McKintey (1997) on the process of digitization and advantages to the higher institutions of learning and pointed out that

(1) digital materials can be stored, transmitted and retrieved easily;
(2) access to electronic information is cheaper than its print counterparts when all the files are stored in an electronic warehouse with compatible facilities and equipment; and
(3) online computer catalogue databases, electronic journals and Internet based resources are expensive initially, but worthy of efforts because of much enhanced accessibility.

THE ESSENCE OF DMR

The establishment of DMR is based on the simple philosophy of providing open circulation and access to the manuscripts held by individuals, families, institutions, etc for use. The mission of DMR is to provide reliable and long-term access to digitized manuscripts for use by the targeted stakeholders. The functions of DMR include:

- Preservation of digital Arabic Manuscripts and other relevant manuscripts.

- Provision of assurance to authors, writers, users and owners of the retention and availability of original physical manuscript in electronic format.
- Provision of efficiency in access to digitized manuscripts especially in multi-sites environments.
- Provision of access points to digitized scholarly manuscripts for the use of scholars, students, researchers and other users.
- Provision of opportunities for collaborative activities for shared access and use of digitized Arabic and other relevant manuscripts; and
- Provision of access levels to the digital collections through online catalogue (OPAC) and search – and retrieve technologies, etc

STRUCTURE AND CONTENT OF DMR

For the DMR to achieve its objectives and function more effectively and efficiently, it should be established within the framework of Digital National Manuscript Repository (NDMR) with multi-sites located in the 6 geo-political locations in Nigeria. They should be interconnected via an Internet Website for easy access and use of their collections regardless of location, access time and in a real-time regime.

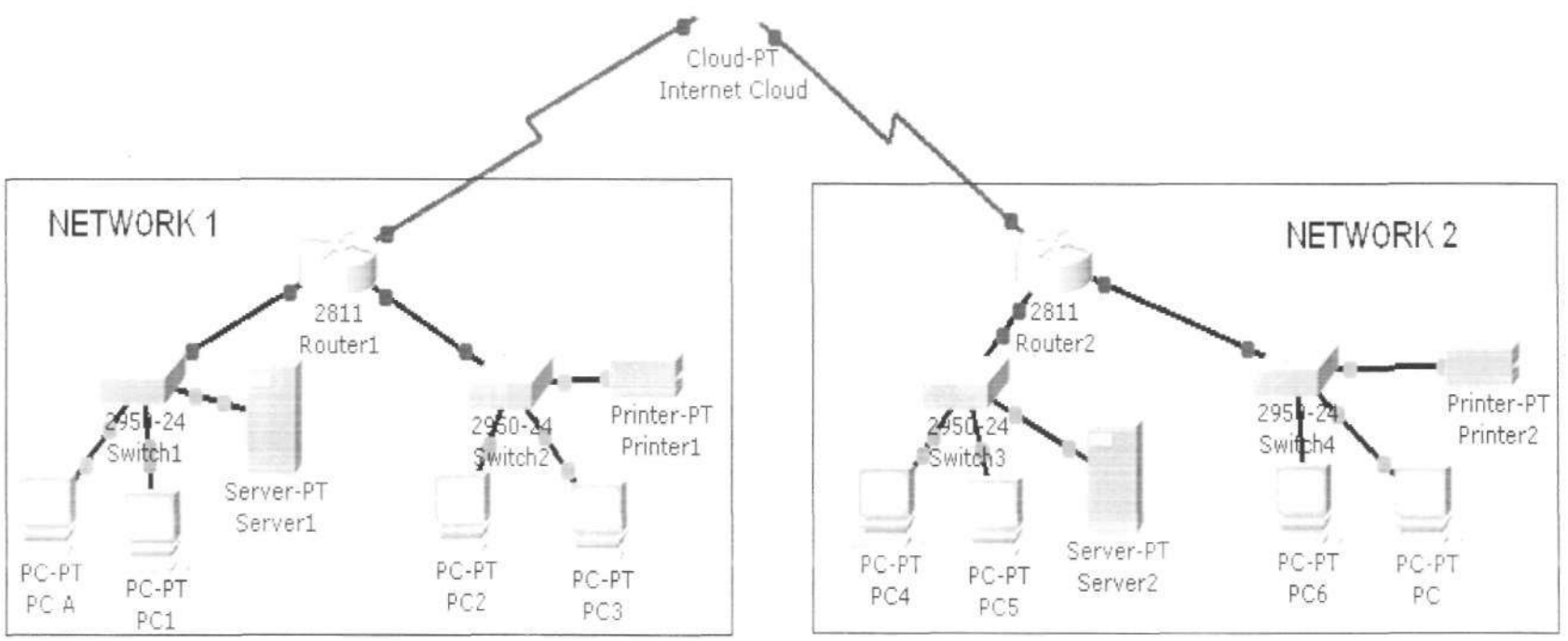


Fig.1 .Distributed Network of DMR System Structure

The content of the DMR should include all foreseeable subject areas in which relevant manuscripts irrespective of language of communication were found and digitized. It should have no restriction to the source from where manuscripts are acquired, the language of communication and also the date of composition/writing. Besides, the development of the DMR collections should be cumulative and perpetual to ensure continued acquisition, organisation, storage and the general management of the manuscript resources in perpetuity. Hence the need for a strategic policy on manuscript collection, processing, the users; access; use; funding and system development and management.

DMR SYSTEMS AND SERVICES

The DMR should make the best use of ICTs to facilitate its functions and provision of relevant services to meet the yearnings of the target audience. Essentially, it should be hosted on automated infrastructure using the Internet facilities to provide real-time services and connectivity. The DMR system should be open and inter-operable to allow for reasonable interaction with other repositories within and outside the country with no access barriers of any kind so as to ensure global access, use and also prevent systems obsolescence.

The DMR may provide such services as:-

- Provision of online services such as e-manuscripts and e-monographs delivery;
- Provision of 'print-on-demand' manuscript publications;
- Burning of manuscripts on CD-ROMs;
- Provision of desiderator on manuscript holdings in a given subject area, etc;
- Development of bibliographies and indexes to specialized manuscript collections.
- Creation of metadata for digitized manuscript resources;
- Provision of free or free-based access to e-manuscript service.
- Provision of electronic delivery service where the user could be served regardless of location; and
- Acquisition and digitization of manuscript resources which would otherwise have been lost, have been inadequately processed or were known only to a few individuals.

The type of services provided by the DMR will however be dictated by the type of computer hardware available, the software platform installed in the system and the type of scanning machine available, and the capacity and skill of the staff.

DMR SYSTEMS REQUIREMENTS

The type of services provided by the DMR and the extent of its use will be contingent upon the type of computer hardware and software

platform installed in the systems to ensure quality service and ease of use by the end users. Issues to be considered along in the choice of software platform include the nature of metadata, manuscript submission workflow, content types, type of server and operating system, type and nature of databases and search mechanisms. Barton, R, Mary, et al (2005) suggested what to consider, the features to look for and the strengths of a typical digital repository platform to include:

- Basic technology building blocks
- Product features to look for
- Technology product models
- Other technological aspects of running a service.
- Implementation stages
- Cost consideration
- Major Institutional Digital Repository (IDR) software providers
- Features checklists.

They added what features to look for when examining a software platform for IDR to include:

- File formats supported: text images, databases, video, audio, etc.
- Metadata standards (descriptive, technical, preservation, right).
- Interoperability: Open Archive Initiative (OAI) compliance, Z39.50, SRW, etc.
- Permanent item address or locator (e.g. Persistent URL)

- Search/browse of metadata
- Full-text search;
- Workflow, submission for content approval;
- User authentication and authorization;
- Back-end: content contributor, editor, administrator, metadata, editor
- Front-end: end-user access to content;
- Customisation: Application Programming Interface (API) for customizing the software, extending features as needed.

The metadata to be adopted for the DMR system should be clearly defined to enable the end-users to easily and efficiently search for, find and retrieve any information needed about any manuscript entry. Similarly, the DMR system should be interoperable with multiple search enquiries and discovery tools to allow for different end-users search different databases and still locate a given manuscript resource in the repository. The interoperability of the DMR will increase its visibility and provide a greater number of citations of resources in the repository.

Wright, Richard (2007) discussed the Metadata Harvesting Protocol developed by OAI to promote interoperability between digital repositories and between e-print archives. That they enable metadata from a number of archives to be harvested and collected together in a searchable database. He observes that the OAI protocol requires repositories to include metadata

fields employed in the unqualified Dublin Core Metadata Initiative (DCMI) standards which include:

- Title of the article/material;
- Creater-the name(s) of the author(s);
- Subject matter of the material
- Description of the content of the material;
- Publisher (Printer);
- Contributor – a person or corporate contributor to the material's content.
- Date the material was created, modified or made available;
- Type – the nature or genuine of the material's content;
- Format in which the material appears;
- Identifier – an unambiguous reference to the material within a given context – using a string or a number conforming to a formal identification system such as ISSN, ISBN, etc.
- Rights – information about the rights held in and over the material including intellectual property rights, copying rights, etc.

Customarily, it should be at the disposal of the DMR to introduce some changes in metadata structure and composition or impose some additional metadata fields such as whether the manuscript/information resource has been peer reviewed, whether full-text of the resource is available and where, an indication of the status of the resources (such as not published, published, in press). Consequently, it is suggested that the DMR

should install Greenstone Software (http://www.greenstone.org) as against other softwares available to repositories such as:

- Dspace (http://www.dspace.org),
- Eprints (http://www.eprints.org), and
- Flexible Extensible Digital Object and Repository Architecture (FEDORA) (http://www.fedora.org).

The Greenstone software is strong in its interoperability with other platform and protocol as it allows for import and export of collections from one platform to another. It is a multilingual thereby making it fit for the DMR project. Besides, it is distributed freely by the UNESCO who developed it in partnership with Human Infor NGO and issued as open source under the GNU General Public License.

The major hardware requirements for the DMR are:

- ❖ HP ML 370T Prolient Server
- ❖ Sony 13.6MP Camera
- ❖ HP Desktop 7800 with Flat Screen
- ❖ HP 4015 N Printer
- ❖ HP 4015 Toner
- ❖ Bookeye-3 Repro Color System A1 scanner; Book Scanner ScanRobot SR301 or Xerox Documate 752 scanner.

However, a decision could be reached on the number of hardware items to acquire and also on whether to acquire HP products or Dell products. The

same decision on the type and model of scanner should be reached. Alternatively, X-rox Scanner Products with similar capacity could be purchased. Provision should also be made for other related consumables.

The DMR should be manned by skilled and experienced Librarian and ICT personnel with requisite knowledge of programming, networking and systems design and maintenance.

DMR SYSTEMS SUSTAINABILITY

A study conducted by Usman (2007) on the status of digitization in 30 Nigerian Universities identified the obstacles militating against digitization projects. These include:

- ❖ Lack of funds;
- ❖ Lack of appropriate facilities;
- ❖ Lack of skilled manpower and
- ❖ Lack of constant electricity supply.

This findings point to the fact that much need to be done to ensure the success of the DMR project. Hence, the DMR should be seen as perpetual system with no predetermined end. What is therefore required is to develop some strategies for its survival and sustenance. The strategies should cover such areas as:

- ❖ Development of DMR Management structure

- ❖ Staffing and Human Resource Development
- ❖ Funding and finance
- ❖ DMR policies and policy implementation strategies.
- ❖ DMR strategic plans
- ❖ Systems quality assurance, monitoring and evaluation mechanisms;
- ❖ Systems infrastructure/facilities development and maintenance;
- ❖ Systems development, operations and services; and
- ❖ DMR systems Security.
- ❖ Major and alternative source(s) of electricity supply.

CONCLUSION

It is generally agreed and established that manuscripts will continue to have great role to play in knowledge and information acquisitions and access. As primary sources of information delivery and knowledge acquisition in diversed areas of socio-political, economic and academic disciplines, manuscripts would and perhaps are facing threats of extinction and unduely inaccessible continually due to theft, mutilation, inheritance, over possession and protection by the owners and human and natural disasters leading to their inaccess and under utilization. What is urgently needed is their digitization and storage in networked digital repositories to enhance their preservation, protection and conservation for effective access and use at local, national and global levels.

REFERENCES

BARTON, Wary R, et.al (2005). "Creating an Institutional Repository" LEADIRS workbook. (http://dspace.org/implementation/leadir.ptf).

USMAN, Ibrahim Alhaji (2007). " Digitisation of Past Question Papers, Dissertations and Theses: A Case Study of 30 Nigerian University Libraries". The International Information and Library Review. Vol.39(3-4), September/December, pp 230-234.

WRIGHT, Richard (2007)."Digital Repositories Explained, Presto Space Deliverable DB.4 Repositories" (www.prestospace.org/project/deliverables/d.3-4 pdf).