From DATAD to ETDs : the Way Forward

Felix N. Ubogu

University of the Witwatersrand, Johannesburg; Private Bag X1, Wits 2050. South Africa
ubogu.f@library.wits.ac.za
Association of African Universities, P.O.Box AN5744, Accra, Ghana.
www.library.wits.ac.za
mary@aau.org
Keywords: ETD Africa, DATAD, Electronic Theses and Dissertations

Abstract

The number of electronic theses and dissertations (ETDs) in Africa is increasing albeit slowly. The majority of initiatives in this arena are in South Africa but there are efforts to get institutions in other countries involved. The participation of eleven institutions in the Database of African Theses and Dissertations (DATAD) initiative has prepared them to embark on more extensive ETD projects. These institutions are now better prepared to share their information resources through improved management and provision of wider access to their collections of theses and dissertations. The institutions have established management teams comprising university decision makers, officers responsible for the administration of postgraduate studies, information technologists and librarians.

The paper discusses the DATAD initiative including its objectives, achievements, lessons learnt and the next steps. It also gives thought to ways that UNESCO could contribute to current efforts and initiatives directed towards increased involvement by African institutions in ETD projects.

Introduction

Africa has begun to participate in the global efforts to make theses and dissertations (TDs) information available, and there are increasing initiatives to that effect. As mentioned in the Unesco *Guide for Electronic Theses and Dissertations*,

> access to ETDs may be the only recourse open to those in developing countries who cannot afford to make purchases from Proquest, who cannot wait for expensive shipping of copies through interlibrary loan, who cannot attend the myriad conferences that demand the considerable expenses related to travel, or who cannot pay for expensive journals (that only may have short summaries of thesis or dissertation results) (Fox, 2001).

Presently, most of the TDs produced in Africa are not easily accessible to either students or researchers. The vast bulk of research conducted in Africa remains as grey literature and is thus effectively lost (Sturges and Neill, 1998).

Only South Africa and Egypt have computerised bibliographic information networks on which bibliographic information on theses and dissertations produced in each country can be found. In South Africa, a *Union Catalogue of Theses and Dissertations* (UCTD) is available online at http://www.sabinet.co.za/. The database contains bibliographic records of theses and dissertations at masters and doctorate levels submitted to universities in South Africa since 1918. Honorary doctorates are also included. The database is limited in its information content, lacking essential elements such as abstracts and specific subject headings. Unfortunately, on 20 February 2002, the Ferdinand Postma Library, Potchefstroom University for Christian Higher Education, informed academic libraries that the compilation of the UCTD for which it had been responsible, was to be discontinued.

In Egypt, all theses and dissertations are available on the Ain Shams University Network (ASUNET). Theses and dissertations of Egyptians who graduate from any other international university are also recorded. The database entries contain bibliographical data as well as the abstract of each thesis. The University Information Network (UIN), whose headquarters are at Ain Shams University in Cairo, serves as a national clearinghouse for theses and dissertations completed in Egypt. UIN carries out a set of activities very similar to those of UMI.

The Database of African Theses and Dissertations (DATAD)

The lack of information on African research output and the inaccessibility of theses and dissertations completed in the region has been a problem for many years. The absence of publications on these works in the international literature, has denied African scholars the recognition they deserve. The Project for Information Access and Connectivity (PIAC) convened a meeting in Nairobi in January 1998 and deliberated on these issues. An outcome of this meeting was a plan for a feasibility study for a project to index, abstract, and distribute theses and dissertations completed in African universities (Materu-Behitsa and Levey, 1998). The meeting recognised the critical need to get research on Africa, by Africans, into the global knowledge pool. The Database of African Theses and Dissertations (DATAD) initiative was born out of these discussions and the Association of African Universities (AAU) has since then led the DATAD process.

The DATAD objectives are to:
- Enhance the capacity of African universities to collect, manage, and access their theses collections;
- Provide visibility and improve accessibility to the work of African scholars both within and outside of the continent;
- Provide support to AAU member universities, and cooperation among them;
- Facilitate the development of relevant copyright procedures and regulations, which would promote the protection of the intellectual property rights of African university scholars and researchers; and
- Create capacity in African universities to use electronic media to collect, manage, and disseminate theses completed at their institutions.

DATAD is an interesting case and strategic starting point for national and even regional ETD initiatives in Africa. The DATAD program by the AAU was conceived as a regional initiative to address the problem of production, management, access and dissemination of research information from universities and institutions of higher learning in Africa. A three-year pilot project in February 2000 and involved eleven (11) institutions, namely: University of Zimbabwe, Addis Ababa University, Cheikh Anta Diop University, CODESRIA, University of Dar es Salaam, Eduardo Mondlane University, Ain Shams University, Yaounde1 University, University of Ghana, Makerere and Kenyatta Universities.

The pilot has made significant impact on institutions. Noteworthy areas of impact on participating institutions include:
- **Institutional policies and practices**: The establishment of management teams, involving institutions' management and decision makers, librarians and IT staff, has facilitated discussions about institutional policies and practices. Issues addressed include copyright, anomalies in thesis submission formats across faculties and departments, public access policies and practices for theses and dissertations, and the desirability of electronic submissions. In some cases, individual institutions have started using the database to identify/ verify new research areas and research profiles for university staff.
- **Capacity building at the participating institutions:** Provision of hardware and software, multifaceted training for staff, record formats and data entry guides.
- **Networking and Team Spirit:** Through workshops, Advisory Committee meetings and a DATAD LIST-SERV participants in the project are able to share experiences and pool resources to collectively overcome obstacles and challenges.
- **Capacity building at the AAU** for providing intellectual input and leadership for the process and the development of working manuals, coordination, management, training and quality control, design and maintenance of the DATAD database and Home Page and providing online access.

- **Tangible outputs:** In addition to the eleven active DATAD sites, the DATAD online database (DATAD OnLine *http://www.aau.org/datad/database/*) is available for online searching over the Internet and will later be available on CD-ROM. The DATAD Methodology Manual (*http://www.aau.org/datad/reports*) is a tool to guide new members joining the initiative.

Even before the main database was launched in April 2003, institutions have reported various roles it is already playing:
- as a research results dissemination tool it gives a summary of the findings.
- as a networking tool it gives the names of the authors, the supervisors, and their contact information.
- as a quality control tool it serves to bring to light the quality of research undertaken in African Universities
- it serves to identify research gaps and reduce duplication
- as a motivation tool, it puts to light the works of African scholars and lecturers

Conceived with the long objective to *Improving Management and access to African scholarly work,* the DATAD program SYSTEMatically introduces ETD to its participants, prepares for the eventual development of a digital library of theses and dissertations. The following programs are planned to evolve out of the DATAD initiative:
1. The Database of African Theses and Dissertations (DATAD)
2. Digital Library for Theses and Dissertations (DLTD)
3. Guide for Copyright and Intellectual Property Rights

While DATAD focuses on retrospective digitisation, the objective of the Digital Library for Theses and Dissertations (DLTD) is to initiate electronic submission of theses and dissertations.

Reports received from the participating institutions reflect a general awareness among many universities in Africa of the need to digitise TD information. Some universities are now apparently requesting students to submit digital copies of their TDs. This is a welcome development that will lead to the addition of TDs produced in Africa to the global knowledge pool.

DATAD Infrastructure and Management

DATAD advocates a combination of both centralized and decentralized architecture at management and technical levels. This was found most acceptable by institutions and technically feasible given the relatively poor and unreliable ICT infrastructure in most parts of Africa.

The **centralized** aspects and activities include:
- Standardization of Record formats
- Quality control monitoring
- Distribution / dissemination of the
- Minimum specifications for hardware and software
- Data integrity for the central server

- Project documentation
- Management and maintenance of the central server

  Decentralized aspects and activities are:
- Data entry and editing
- Quality control
- Storage, archiving and preservation
- Access authorization to full text/document delivery
- Copyright and IPR control for full text

DATAD uses a bibliographic management software (Procite) as the bibliographic management tool. It uses common operating SYSTEM, (windows 98, 2000, NT), a simple web posting software for the LAN (Reference Web Poster), common word processor (Word Perfect, or MS Word), optical character recognition software (OmniPage Pro11) to scan text for speedy data entry.

An SQL server manages the integrated database. SQL was chosen because it is open source and has the potential to evolve to UNICODE.
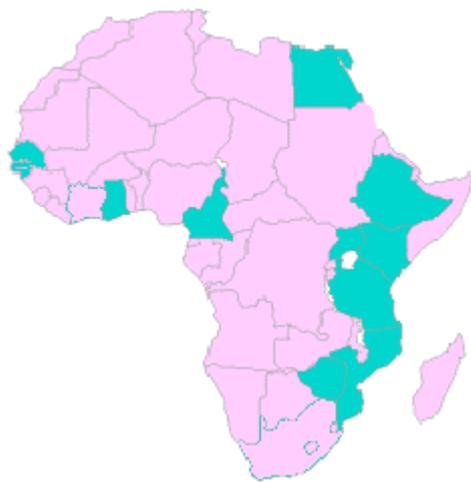

Figure 1: DATAD: Participating Countries (Countries colored Light Blue)

Electronic Theses and Dissertation Initiatives in Africa

The number of electronic theses and dissertations (ETDs) in Africa is increasing albeit slowly. The majority of initiatives in this arena started in South Africa. In the past three years, there have been initiatives from other countries. The participation of eleven institutions in the DATAD initiative has prepared them to embark on more extensive ETD projects. These institutions are now better prepared to share their information resources through improved management and provision of access to their collections of theses and dissertations. The institutions have established management teams comprising university decision-makers, officers responsible for the administration of postgraduate studies, information technologists and librarians. For example, Makerere University has held several workshops on the subject of Electronic Resources Library Management. The MakLIBIS project, within the wider library automation project, will undertake digitisation of library materials and establishment of full text databases. The University of Zimbabwe, through DATAD is lobbying for the introduction of a policy that will make it a requirement to submit theses both in print and electronic media. The Council for the Development of Social Science Research in Africa (CODESRIA), through its Small Grants Program for Theses Writing has accumulated close to a thousand theses. CODESRIA is now planning to digitize its collection and make it available on the web and CD-ROM.

In South Africa, development in the ETD arena started in 1996 when Rhodes University (RU) undertook an investigation into the holding of electronic copies of theses and dissertations. The Senate of the University in April 1997 approved mixed submission of both print and digital copies of student theses and dissertations. With the support of Virginia Polytechnic Institute and State University (Virginia Tech), which donated free software to Rhodes, the University mounted its first digital thesis on the World Wide Web in 1998, and became the first institution in Africa to do so. Since then, the number of theses and dissertations available at RU's electronic archive has steadily increased.

Of the 37 universities and technikons (polytechnics) in the country, Rhodes University *http://www.ru.ac.za/library/ electronic_resources* and the University of Pretoria *http:// upetd.up.ac.za/ETD-db/* are the only two institutions that make the abstracts, and in some cases full text, of their TDs available on the Internet.

The slow pace of development in the ETD arena in South Africa may be attributed largely to lack of expertise, and budget constraints to buy necessary equipment (Ubogu, 2002).

There seems to be a general awareness among many universities in Africa of the need to digitise TD information; and some universities are now apparently requesting students to submit digital copies of their TDs. This is a welcome development that will lead to the addition of TDs produced in Africa to the global knowledge pool. Many institutions have adopted the technologies used by the Virginia Polytechnic Institute and State University (Virginia Tech) to implement their ETD database. The software is available free of charge to members of the Networked Digital Library of Theses and Dissertations (NDLTD). The ETD database is a series of web pages and perl scripts that interact with a MySQL database. These scripts provide a standard interface for web users and researchers, ETD authors, graduate school personnel, and library personnel to enter and manage the files and metadata related to a collection electronic theses and dissertations *http://scholar.lib.vt.edu/ETD-db/*.

The recommended hardware is a UNIX-based web server platform with adequate memory and enough disk

space. The software is mostly freely available without charge to education institutions, and include MySQL, perl, CGI.pm, DBI and DBD:Mysql modules for perl, web server software *http://scholar.lib.vt.edu/ETD-db/developer/ require.html*.

The ETD Africa initiative will utilize the NDLTD software at institutional and national nodes. The full text ETDs will be indicated and referred to in DATAD On-Line.

The Way forward

Efforts should be made to drive ETD development in Africa at national levels. Iinstitutions that award higher degrees should participate in national efforts requiring theses and dissertations to be prepared and submitted electronically. The national initiatives should be predicated on the need to share software, metadata, documentation and training tools. The pooling of expertise and resources may in particular benefit those institutions where staff with highly developed IT skills are in short supply. Sharing of detailed documentation on all aspects of operating an ETD programme is a very cost effective method of collaboration. The development of generic procedures and training programmes, which can be customized for local conditions, may facilitate the participation of institutions in an ETD programme. Sharing of documentation is also likely to reinforce the use of standards, which will ensure that ETDs are readily retrieved (Ortuzar, 2001).

Resources should be mobilised locally and from development partners to initiate the ETD process at national levels.

The specific objectives of the initiative should be:
- To make African theses and dissertations accessible online to scholars and users worldwide;
- To build capacity at the institutional, national and regional levels on digital libraries;
- To promote regional cooperation on digital library SYSTEMs;
- To train scholars in the technique of electronic publishing and electronic access to documents; and
- To build on the network established by DATAD and to increase participation in this network.

The initiative should adapt the training and course materials developed in identified pioneering institutions for use in the implementation of projects. Institutions should be trained to adapt the protocol of the Networked Digital Library Of Theses And Dissertations (NDLTD). They should also establish project teams where none exists presently, and these project teams should initiate ETD projects in their institutions.

While most institutions would be willing to make the abstracts of their TDs accessible in a central database, they may not be prepared to give up control of the full text of these TDs. Institutions could retain responsibility for the control of such intellectual property. This means that the abstracts of TDs will be made accessible in the DATAD database but the full-text will remain under the control of the institutions.

In this distributed model, the participating institutions will be primarily concerned with developing content, while DATAD will provide the common service components. This will include harvesting metadata from participating institutions, and DATAD will continue to maintain a common search interface for distributed national resources, without interference in the intellectual property rights of individual institutions. The DATAD website will be enhanced for these capabilities and DATAD will host ETDs of partner institutions if required to do so. DATAD will continue with its other interventions, setting standards for best practice and quality control, and developing an awareness of the scope of institutional commitment to formats, funding, storage and migration strategies.

Governance

Each country would establish LEAD institution that will serve as a national node. The lead institution will co-ordinate ETD initiatives in that country and collaborates with AAU-DATAD as necessary. The lead institution and AAU-DATAD will lead discussions on policy, strategy, and financial issues. Wider involvement from the academic communities will come through focus groups, workshops and discussion lists. In meeting its objectives, the project should, wherever possible, make use of work that has already been done and to build upon existing expertise in digital preservation and digital collection management.

Conclusion

Africa is on the road to making the intellectual products, TDs, of its higher education institutions available in the international knowledge pool through improved management. The support of development partners will go a long way to bring about the efforts made presently.

References

Association of African Universities /Association des Universités Africaines. www.aau.org/coreprog/0104/ themes.htm#Sub-Theme 5: Improving management and access to

Fox, Edward (2001). Enhancing access to university research. *In:* The Guide for Electronic Theses and Dissertations. UNESCO. Available at: etdguide.org/

Larney, Tom (2002). Discontinuation of UCTD. Email to heads of academic libraries.

Materu-Behitsa, Mary and Levey, Lisbeth A (1998). Database of African Theses and Dissertations: Report of a Feasibility Study. December 1998, p.iv. Also available at: www.piac.org/datad/

Ortuzar, Gabriela (2001). Standards, cooperation, and collaboration. *In:* The Guide for Electronic Theses and Dissertations. UNESCO. Available at: etdguide.org/

Sturges, Paul and Neill, Richard : *The Quiet Struggle: Information and Libraries for the people of Africa*. 2nd ed. Masell, London and Washington, 1998, p.168.

Ubogu, F.N. (2002). ETDs in South Africa: Current Status. Available online at: www.access.wvu.edu/esra/ESRASearch_VSearchResult.cfm

UNESCO. Project for the development of an international framework for the creation of electronic theses and dissertations (ETDs). Available online at: portal.unesco.org/ci/ev.php?URL_ID=1580&URL_DO=DO_TOPIC&URL_SECTION=201&reload=1039193861