
Consortia Repositories and Content Management for African Scholarly Publishers

Jacob M. Jaygbay

JEPS Publishing Services, Ltd.

Virginia, USA

Internet: www.jepspublishing.com

email: jaygbay@jepspublishing.com

Paper prepared for presentation at the Conference on
**“Bridging the North-South Divide in Scholarly Communication in Africa:
Threats and Opportunities in the Digital Era”**
September 6- 8, 2006, Leiden, The Netherlands

Consortia Repositories and Content Management for African Scholarly Publishers

Content management is a process, not a product. It is about getting people to work together, and it takes time and sustained effort to coordinate people.

Addey et al, 2002

Introduction

For all practical purposes, the issue of consortia repository in Africa is more a matter of process than it is of information and communication technology (ICT). We are witnessing in Africa, rapidly evolving trends in the use of ICT and 'best practices' to provide scholarly publishing solutions that improve not only the speed of scientific communication, but also the quantity and quality of information exchange among social science researchers. The use of repositories is one of the emerging solutions fuelling excitement in the scholarly publishing community. These repositories can be databases, closed file systems, or a combination of both. The small size of most social science research institutions in Africa in part makes collaboration at the consortia level more appropriate than at the institutional level. By consortium, we mean simply a grouping of individual research institutions based on region, discipline or other interests with common research and publishing agenda. There are a number of consortia research institutions in Africa. For example, the Council for the Development of Social Science Research in Africa (CODESRIA) that coordinates social science research in African institutions and universities, the Nairobi-based African Economic Research Consortium (AERC), the Organization for Social science Research in eastern and Southern Africa (OSSREA) in Addis Ababa, and the Association of African Universities (AAU), are but few that come easily to mind.

As these repositories grow, the need for an effective content management system (CMS) becomes critical. In spite of the promise that vendors of these tools may suggest, there are still a lot of unknowns in the CMS market space. The space is gradually coming to terms with its own growing pains: there are inconsistent jargons, and many vendors promise a total CMS solution but only deliver a fraction of the package. But what is it about CMS that

makes its design and implementation so intractable? Why is it being marketed as an ICT solution when in large part it really is a process? Where do CMSs fit in with a consortia repository and why do we even make the connection between the two? The simple answer is the user.

To serve users effectively and to promote scholarship, African scholarly publishers must continue to take as seriously the gate-keeping responsibility that they bear in the digital environment: editorial title management, sound copyediting and review process to ensure quality of research, marketing, meaningful information design; as they do in the print environment if they are to remain relevant in the digital age. Critics may argue that the point of the repository is to avoid the middle 'person'. However, we will agree that a giant and growing repository without clearly defined business rules and intelligent information design will quickly become unusable by social science researchers and invariably fail to support scholarship that is the *raison être* of scholarly publishing. Some of the major search engines— Yahoo, Google— for example are shifting to a more content-centric and user focused search capabilities to keep their edge in the search engine market space.

What CMS?

There are various types of CMS, ranging from the transactional type used mostly for e-commerce, to the integrated type used mainly by large enterprises to manage documents and other contents, and the LMS type used mostly for tracking the life cycle of learning content. This paper is primarily focused on publication CMS that is the type used to track and manage the life cycle of publication content like books, journals, manuals, etc.

Publications CMS is about three broadly defined sets of activities: defining, creating and organizing content; specifying the transformation standards of the content; and delivering the transformed content to end-users. Each of these three stages of the CMS process poses challenges that defy the way African scholarly publishers create, transform and publish research results today, and may force social science research institutions in Africa to reflect seriously on the global implications of the local actions taken when they create and archive their research findings. A successful consortia publishing repository calls for a fundamental

shift in the traditional business model of the individual African scholarly publisher from a print model to a multi-channel, multi-format and access driven model.

The level of effort required to identify and get the appropriate stakeholders together and take inventory of what publications are available and in what formats further complicates challenges of developing a viable consortia African social science repository. At the technological level, what are the system integration requirements? Publishing requirements? How do you shift to the use of the new solution without losing your existing readership? Importantly how do we make the published content to read natural? In the context of the latter, a meaningful African consortia repository has to deal with the pain of translation (read internationalization and localization) of its contents given the language barriers posed by the use of four languages for scientific research: English, French, Arabic and Portuguese.

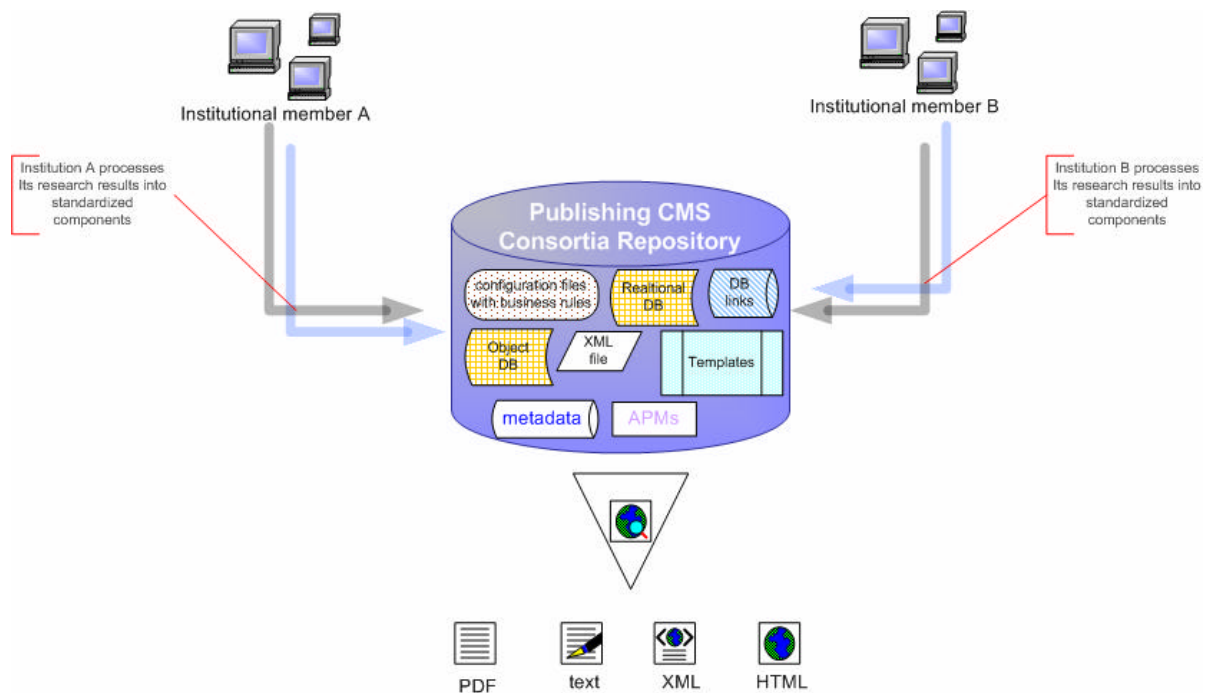


Figure 1: High Level System View of a Publishing CMS Repository

Contextualizing CMS Solution for African Scholarly Publishers

As Internet access is being slowly introduced to university students, academic institutions, and social science researchers; the publication programs of African research consortia are

feeling the need to respond to this shift in traditional research methods while still remaining relevant to their primary constituents. However, African scholarly publishers need to resist the temptation of viewing Africa as a monolithic user group. There are not only language barriers, but also stark differences in users' access to the Internet through the available ICT infrastructure to support Internet connectivity. Moreover, African publishers should not assume a digital culture in these earlier stages of the Internet on the continent. Are not cultural changes always trailing technological advances? The emergence of the Internet in Africa is no exception. To narrow the gap between technological advances and cultural changes that encourage the use of any sustainable scholarly publishing ICT solution, the targeted users have got to be central to our requirement analyses. The primary question therefore should be "What" do African researchers want from the consortia repository, as opposed to "How" do we provide a repository to African researchers. The challenge for African scholarly publishers is figuring out a way to transition to this broader collaborative publishing scheme without losing current market share but building new bridges. This transition can be viewed from two interrelated but distinct perspectives.

The first is the perspective of the African scholarly publisher that operates in a rather small and segmented market environment, with all the familiar resources constraints. For the scholarly publisher CMS is in fact a major shift in its business model that may have long-term effects of their survival. Most cannot afford the upfront initial investment with any degree of certainty of recovering the return on their investment. As a collective venture, the challenges of organizing the collaborative structures and defining the set of standards to ensure an effective collection and organization of content, transforming that content for presentation, and publishing it can be more daunting than is initially envisaged. This is partly due to the fact that a CMS solution is viewed primarily as an IT exercise. The truth is that in spite of the efficiency that African scholarship can gain from ICT, content management is arguably more of a process issue than it is a software and hardware issue. Based on the author's own experience, the conservative nature of many African scholarly publishing institutions does not make the decision to make such a major investment any

easier. This conservatism is the result of real issues associated with funding and the pressures from cost cutting measures faced by these institutions.

The second perspective is that of the ICT buff who regards CMS mostly as a marketing or a knowledge management problem. This view is primarily based on the fact that current database models are designed to deal with data as opposed to content. To effectively tackle this inadequacy of database support for content, a viable CMS solution for African scholarly publishers should be able to achieve at least two broadly defined objectives:

- Recognize and handle the multitude of ambiguities posed by 'content' that are not present when dealing with 'data'. Some of the common ambiguities associated with content world include: taxonomy (*Meat vs. Beef*), synonyms (*Mimic vs. Imitate*), source language (*Man vs. Homme*), etc. The latter is particularly salient given the number of languages used to conduct social science research in Africa.
- Provide users the ability to conduct queries that combine both structured and unstructured fields in the CMS database. Such queries are not supported by independent text and XML indexing or by the more traditional B-tree indexing. This may require building a full-fledged search engine into the Database Management System (DBMS) of the CMS.

This fundamental perceptual difference leads ICT professionals to define the development of CMS solutions as primarily a matter of process as only after defined processes and rules are put around data, we give meaning and context thus bringing light to the idea of 'content' in CMS. As with institutional repository CMS, consortia repository CMS has to determine collection requirements, figure out what they want to publish and in which formats the outputs ought to be, and define a process to manage the entire project. For the consortia CMS to succeed, these three sets of requirement discussed above need to be collaboratively determined at the institutional level as well. In the case of Africa, scholarly publishers have to contend not only with only different types of contents published in the multiple languages being used for social science research but also a small and segmented market for these publications.

Analyzing these three sets of requirements and managing the tensions among competing requirements, and consolidating common requirements ultimately constitute the critical factors in determining what component belongs to which institution, what component goes to what market, and what component belongs to which page. It will also help to figure out which institutions are building what components and for which market, and what metadata should be applied to provide meaningful context to data, keeping the contents easily manageable and that it can be read naturally in a integrated context.

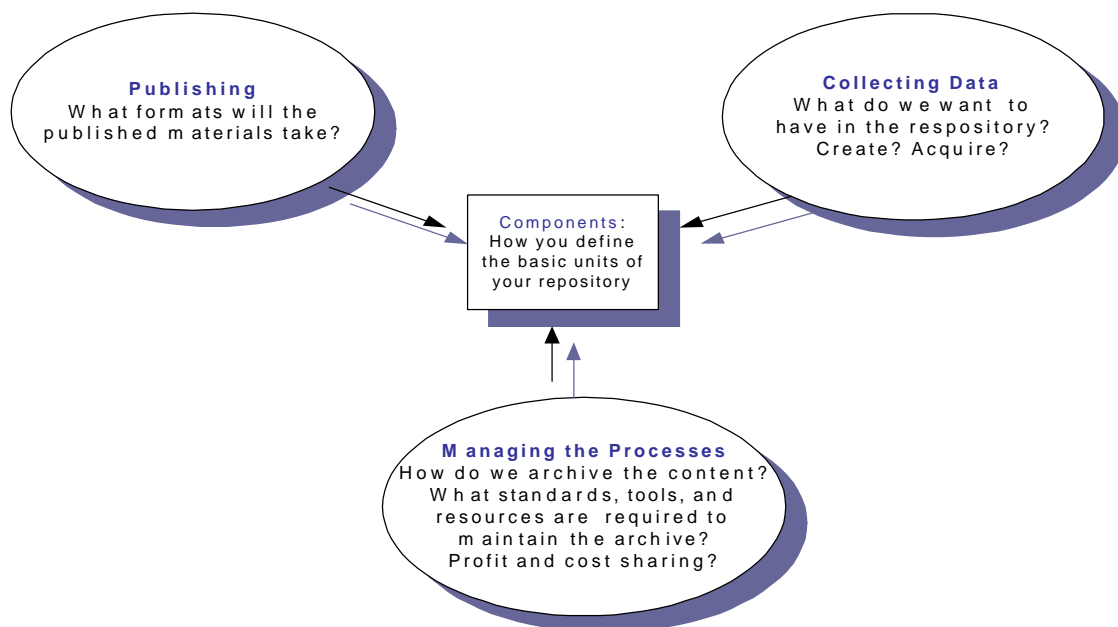


Figure 2: Defining the Content Type for your CMS

Current Trends in Scholarly Publishing

There is growing interest by institutional scholarly publishers in developing electronic archives. These archives are usually in the form of Open Access journals or institutional repositories. The latter is less known among researchers than the former. The trend is partly explained by the rise in the number of open access scholarly journals. A recent survey shows the number of open access scholarly journals grew by 18 percent worldwide from January 2004 to July 2005 (Rowlands and Nicholas 2006). In addition, there are a number of partnership-type projects such as the Journal Storage Project (JSTOP), funded by the Mellon Foundation; the Electronic Publishing Trust for

Development, the International Publishing Network for the Availability of Scientific Publications (INASP), and the National Inquiry Services Center (NSC) sponsored Africa Journal Online (AJOL). There are some positive points about this growing interest. It will help to reduce the rising cost of printed scholarly materials, and improve access to research findings. Both factors tend to advance scholarly publishing because they support the traditional goal of scholarly research institutions, namely conducting research, using the research findings to teach and facilitating learning. Obviously, scholarly publishing is less about teaching than it is about research and publications.

In February 2005, Oaster indexed over five million items in over 405 institutional repositories around the world. How much of those items are from African repositories is unclear. The representation of African published scholarly publication remains marginal. Just a quick query of <http://scholar.google.com> can give you some clues. However, we are witnessing a trend that only seems to be gathering momentum. It is a trend that is bound to affect the way social science research is conducted on the continent. The question is no longer, if some form of content management repository system is in order, but what is the best way for African scholars to adopt this solution. Can we reasonably assume that if we build this capability, African scholars and the rest of the academic community will come on board? How do we gauge success?

Is it Time for an African Consortia Repository?

It is not clear if one can easily answer this question affirmatively with an acceptable degree of certitude. However, it is evident that scholarly publishing is changing and this change is being driven partly outside of the three traditional influences: the university library, scholars, and scholarly publishers. ICT is not a new ally to this tripartite but it becoming a driving force that is changing the traditional publishing landscape faster than the traditional allies are accustomed to. To get involved and move forward, African research institutions, university libraries, and scholarly publishers will have to form new alliances with private and corporate partners to get actively involved in any discussions about a consortia repository and do so on their own terms and not remain a powerless consumers of knowledge

produced elsewhere and by interests that may not be so favorable to their research, teaching and scholarly publishing needs.

How much longer can African scholarship be left on the fringes of global knowledge exchange? The argument that the 'prestige' of African scholarly journals (read publications) makes them less favorable for listing in 'recognized' social science indexes is vacuously true because African scholarship remain largely underrepresented in most indexing services. It is the mere absence sometimes not the quality of the African publications that feeds the cynicism. There is growing evidence that improving the online presence of African scholarship through open access journals or a viable African consortia repository can increase the citation of African researchers and enhance the credibility of their scholarship. The CLASCO experience in Latin America reported by Babani and Smart (2006) is a glaring example.

Much of the literature on developing a publications CMS tend to agree that the barriers to developing a viable consortia CMS are more political and cultural than they are technological (Mackie 2004; McLean 2004). In Africa, what particular factors will have the most influence on the implementation of a viable publications CMS solution for Africa? The volume of social science research publications out of Africa today is unprecedented. Some journals are now being released both online and in print and this makes accessibility easier. However, there is still a lot that can be done in terms of improving the accessibility to African published social science journals. We see new forms of web-based citation tools but most still lack a meaningful representation of the volume of social science research work being done on the continent. That has to change as African scholars themselves increasingly reference these repositories. For example, scholaruniverse.com is a database of social science researchers all over the globe. However, most of the researchers covered come from South Africa. They have shown interest in incorporating other existing databases of scholars from other universities in Africa. A consortia repository can make such collaboration not only easier to develop but also mutually beneficial to both Africa and scholaruniverse.com or other relevant business partnerships.

On the Internet, as is with printed scholarly publications, African social science research institutions have to deal squarely with the issue of legitimacy that characterizes the publication and distribution of scholarly journals globally. While we continue to foster new collaborative mechanisms to promote the presence of African social science research and published publications, legitimacy of African-published journals is better secure if African scholars are part of the content creation and archiving process of these repositories. The www.scholaruniverse.com, a web-based citation indexing service, is a classic example.

Define Clear Business and Development Goals

Having access to the larger market does not in itself translate into profits for small African scholarly publishing institutions. In fact each African scholarly publisher, as members of the consortia repository, needs to move beyond its comfort zone to understand what is it that they are selling, when and to whom they are selling their publications. Otherwise, marketing resources not properly targeted can fail to yield meaningful return on investment (Kumar et al 2006). The publishers that benefit most are those that realize that their role as facilitators for the advancement of scholarship in the larger scholarly community is key to their success in the global market environment. The success of a publishing CMS solution has to be based on precise business goals that are communicated to each participating institutional member. The goals should inform not only the processes of acquisition, transformation and publishing of contents, but also measure impact on the advancement of scholarship. The latter is the developmental angle of our role as scholarly publishers. Are we reaching our targeted markets? How do we gauge the impact of the consortia repository on the presence of African scholarship globally?

There are serious implications for the way African scholarly publishers do business. In many cases, this will require thinking global even in the context of local publishing activities. Parsing your contents so that it fits and reads naturally when combined with other contents is not an easy feat but is crucial for sharable electronic repositories. It may require publishers to create better metadata about their contents if only to ensure better accountability at the consortia level.

Additionally, one of the advantages of a publishing repository is reuse. The consortia repository is capable of providing services that are not offered by any of its individual

members. The more thorough we can define the requirement for the market, and the basic components of the repository, the better we can cater for the global market. This is not to downplay the politics of scholarly research and the exchange of scientific knowledge which the author has discussed elsewhere (Jaygbay 1998).

Workflows and Multi-language Support

Beyond the traditional research institutions, and universities, Africa today can boost of numerous associations of research institutions that provide space for collaborative social science research and outlets for coordinating and publishing research results. These institutions are of varying sizes and research capacities. A serious look has to be given to institutional membership and the issue of copyright at the individual, institutional and consortia levels. The role of each of these individual institutions to the success of consortia repository is paramount. To ensure that each member institution is functioning by the standards of pre-defined processes, CMS solutions will need to devise workflows for the CMS that allow them to apply the rules that are collectively defined during the requirement phase. These rules are primarily geared toward enforcing standards through the granting of privileges and assignment of tasks to each member of the consortia to ensure the team functions within the boundaries of the pre-defined processes. The fact is that many scholars are members of various research institutions, and these institutions are themselves associated in many ways. Managing who does what and during what phase and timing in the process is of paramount import. Additionally, consortia CMS solutions in Africa have to contend with the issue of translation that further complicates the definition of content and publication requirement.

One of the undisputable benefits of an African consortia CMS solution is to take advantage of the economy of scale. The not-so-obvious advantage of a successful CMS solution is cost effectively providing localized publications to a global market space where African scholarly published literature is usually absent. I must add sometimes by design. With CMS solutions, we are seeing a number of vendors also offering translation solutions. The integration of translation solutions can help African scholarly publishers to break the cycle of the ever-encroaching boundaries posed by the small-segmented markets with language barriers. Effective CMS solutions tend to control the impact of rising translation cost. This will require African publishing institutions to maintain their contents in their

primary language and other languages as well. The challenge is doing so within budget and time constraints.

African consortia scholarly publishers will need to take serious look at how to deal with issues of translation not only to reach a larger global market space but also as a part of their collaborative venture as the repository will have to support multiple languages to reflect prominent languages used to for social science research. The requirements for internationalization/localization should be considered seriously.

Conclusion

African scholarly publishers have produced meaningful amount of social science research data and continue to do so albeit with some difficulties. Most of the data were archived without much thought given to retrieval, transformation and publishing in a collective repository. Collecting, organizing data from multiple sources, in different formats and languages is not only a complex task but can be a daunting one. It is also a task that cannot be easily automated through the use of ICT. The rules will change as the needs of researchers do and the tools available to them for research evolve. The Author's own experience with building requirement, finding a suitable vendor, and implementing a Learning Management System (LMS) CMS for a major software company in the Washington, DC metro area indicates that most of the work required is process intensive and can be time consuming even for a single language CMS. Given the intricacies of the African context characterized by multiple languages, varied and multiple formats of existing data, the familiar socio-economic constraints, this process can be more daunting than most African scholarly publishers might anticipate.

Additionally, the issue of standards needs to be addressed upfront especially in the context of such a vast collaborative effort as an African consortia repository. The standards issue is twofold. On one hand, African scholarly publishers will have to review the standards that emanate from the ICT industry like XML, J2EE, SOA, etc. to see which ones are adaptable to the business needs of their CMS solution. Most of these standards are more stable and widely used in the software industry and can be adopted fairly easily if the

process begins with clearly defined set of system requirements. On the other hand, and more important is the standards that African scholarly publishers will have to create to ensure interoperability among the membership of the consortia repository. The issue is that there are no dominant open source standards to this date, and the two that are most widely cited, Darwin Information Typing Architecture (DITA) and DocBook, may not be adequate.

Importantly, we need to remember that an African consortia repository solution will not only grow but is bound to change as Africans social scientists change the way they conduct research and the ICT that drives the CMS solution changes. To be successful, any consortia repository solution that is developed has got to plan to accommodate both growth and the need for change. The conservative nature of academia does not immunize the African scholarly community from the impact of technological change. The way African scholars acquire and consume scientific knowledge is changing, and African scholarly publishers can change along with it to remain relevant to African scholarship or simply wither.

Bibliography

Addey Dave, Ellis James, Suh Phil, and Tiemecke David (2002), Content Management Systems, USA, Canada, Glasshaus.

Boiko, Bob *Content Management Bible* (2005), Indianapolis, IN, Wiley Publishing, Inc.

Dominique Babini and Pippa Smart (2006) 'Using digital libraries to provide online access to social science journals in Latin America' in *Learner Publishing* vol. 19, No. 2, 107-113p.

Jaygbay, Jacob (1998) 'The politics of and prospects for African scholarly journals in the information age' In *Knowledge dissemination in Africa: The role of the scholarly journal*, ed. Philip G. Altbach and Damtew Teferra, 63–74. Chestnut, Hill, MA: Bellagio.

Kumar Kumar, Rajkumar Venkatesan and Werner Reinartz (2006), "Know what to sell, When and to Whom" in *Harvard Business Review*, March. Pages 131-137.

Mackie, M (2004) "Filling institutional repositories: practical strategies from the DAEDALUS project" see <http://ariadne.ac.uk/issue39/mackie/>

McLean, N. (2004) "Repository Management and implementation". Paper presented to the E-learning Interoperability Seminar, August. See http://standards.edna.edu.au/idea/neil_malean.pdf

Rowlands Ian and Nicholas, Dave (2006) "The changing scholarly communication landscape: An international survey of senior researchers" in *Learner Publishing* vol. 19, No. 1, 31-55p.

Watkinson, Anthony (2006) 'Open Access: A publisher's View' in *Logos* vol. 17, No 1.