Aluka: Developing digital scholarly collections from and about Africa

Siro Masinde
JSTOR, c/o National Museums of Kenya
siro.masinde@jstor.org

and

Rahim S. Rajan
Aluka’s Collection Development Manager
rahim.rajan@jstor.org
Abstract

Aluka (http://www.aluka.org/) is an initiative of JSTOR (www.jstor.org), a not-for-profit organization that is committed to preserving and expanding access to scholarly materials as broadly as possible, and to enabling new forms of scholarship and teaching. We aim to help researchers, teachers, and students improve the efficiency and effectiveness of their work and to develop new insights. Our efforts have laid a strong foundation – as of September 2008, there were more than 365,000 digital objects available in Aluka, and in JSTOR, the comprehensive archives of over 1000 leading academic journals across the humanities, social sciences, and sciences. In addition to being an access node for important and unique research and primary source material from or about Africa, Aluka also offers a way for individual scholars, libraries, societies, and cultural or research organizations to contribute digital collections to a shared online library. By contributing their collections to a shared platform, content owners have a means of enabling access to their materials to a global educational audience without having to build and support their own web delivery platform. Furthermore, the Aluka platform provides innovative tools for image processing and quality control, as well as for research, teaching, collaboration, and knowledge exchange. This paper highlights why and how Aluka has been adopted as a platform by various individuals and institutions in Africa to disseminate and publish their research and scholarly materials. We discuss the advantages of having a third party, not-for-profit initiative assist in making available online African research and scholarship. Lastly, this paper discusses some of the crucial lessons Aluka has learned thus far.
Introduction

Archives, libraries and primary source materials from and about Africa are few, highly dispersed and difficult to access. Africa has a rich natural and cultural heritage as well as important bodies of traditional knowledge and information systems that have been passed down from generation to generation. However, much of this has not been documented and there is inherent danger of distortion or complete loss. The colonial legacy resulted in the dispersion of a vast array of historical documents and knowledge from Africa. It is no coincidence that archives and museums in Europe and the United States contain much of this patrimony. For many years, people from different parts of Africa as well as other continents have learnt about Africa through accounts and collections emanating from outside Africa. Such accounts are not always reliable or accurate; it is not surprising that this has occurred though as there are significant challenges involved in seeking to access primary source data and information located or archived in Africa.

We know that the development of ICT’s have increased access to online repositories of scholarly literature in Africa. Nevertheless, access to the underlying data and source materials remains a constraint. Primary source materials are the original records about events, experiments, and research and may constitute raw data such as manuscripts, newspapers, photographs, sound recordings, letters, statistical data, artifacts and voucher specimens (cf. Lincove et al 2008). Access to high quality primary source materials is important for high quality research and scholarship. Primary source materials have traditionally been available in print, microfilm and as artifacts but digital technology offers the chance to convert these into digital formats that can be easily searched and aggregated and disseminated worldwide. The creation of digital institutional repositories is advocated as an important avenue to make raw data and other primary source materials accessible (Tennet 2002, Anbu 2006). One of the key challenges however with institutional repositories is that while they prove to be effective means of maintaining institutional memory and preservation, interoperability between different solutions and discoverability continue to pose challenges. Scholars and researchers are not always satisfied by the discoverability options available in institutional repository solutions or the limitations imposed on institutional repositories by various copyright regimes (Palmer et al 2008). There are currently a number of websites that offer primary source materials but a major challenge that researchers and students face in using these resources is the ability to evaluate their provenance, age, and authenticity.

These challenges are one of the main reasons for Africa’s low visibility in the realm of international scholarly publications. To address these challenges, a new digital library called Aluka (www.aluka.org) was founded by Ithaka and The Andrew W. Mellon Foundation in 2004. In creating Aluka, there was a recognition that as internet technologies advanced and access to scholarly knowledge online also advanced, there continued to be a need to include digital information and knowledge from parts of the world that were still grappling with the digital divide and connectivity constraints. Since August 2008, Aluka is an initiative of JSTOR (www.jstor.org), a not-for-profit organization dedicated to helping the scholarly community discover and use a wide range of intellectual content in a trusted digital archive and promote teaching in cost-effective ways. The principal focus of JSTOR has been to aggregate back issues
of scholarly journals in digital format in a broad variety of academic disciplines. JSTOR currently archives over one thousand leading academic journals across the humanities, social sciences, and sciences. More than five thousand institutions around the world access the JSTOR archives. Access to Aluka and JSTOR is free to libraries and universities in Africa.

Aluka digital library and its content

‘Aluka’ is a Zulu derived word that means to weave; it reflects Aluka’s philosophy of connecting scholars and researchers with each other and with knowledge and resources from or about Africa that are physically dispersed. The website was launched in early 2007 and currently has over 365,000 digital objects. The materials in Aluka complement the journal literature aggregated in JSTOR by uniting primary source materials with a focus on Africa.

Aluka’s digital library is being developed in collaboration with academic, cultural, and research institutions around the world. There are currently over 100 institutional and individual content contributors in over 30 countries, including 16 African countries. Aluka is a platform for the aggregation of related but disparate items, and for the sharing and dissemination of high quality multimedia content comprising text, images, 3-D objects and video among others. Aggregation of disparate but related data creates the possibility of additional quality control and checks on consistency. Aluka’s engagement with institutions outside of Africa is also contributing to the “virtual” repatriation of African materials. The presentation of a diversity of objects in the form of text, images, and audio visuals provides powerful tools for teaching and learning. Among Aluka’s innovative tools for manipulating digital data, and which provide new opportunities for investigating digital objects, include high resolution viewing and measuring tools.

The objects in Aluka are organized in three initial thematic areas, namely, African Plants, African Cultural Heritage Sites and Landscapes, and Struggles for Freedom in Southern Africa. The content is applicable to research and teaching across several disciplines.

The African Plants content comprises a scientific collection of data and information about Africa’s flora. These comprise high resolution images of type specimens, photographs, drawings, handwritten field notes by European explorers such as David Livingstone, botanical art, and taxonomic and reference data. Content contributors to the African Plants theme comprise over 60 scientific institutions from Africa, Europe and the USA. Among them are herbaria, botanical gardens, museums and universities that are repositories for African plant voucher specimens.

The African Cultural Heritage Sites and Landscapes comprises in-depth visual, spatial, and research documentation of cultural heritage sites and landscapes in Africa, a number of which are designated as UNESCO World Heritage Sites. The materials include rare slides from early documentation projects and excavations; aerial, panorama, and satellite photographs of landscapes; virtual and 3-D models of sites and structures; unique GIS data sets for each site; antiquarian maps and digital site plans. Fifteen selected sites have so far been documented among them Kilwa Kisiwani, a medieval city on an off-shore island in Tanzania; the rock-hewn churches of Lalibela, Ethiopia; the Asante Temples at Basease and Patakro in Ghana; and the celebrated mosques in the Sudanic style in Djenné and Timbuktu, Mali. The digitization of rock art images across Africa and the rare, fragile medieval manuscripts of Timbuktu are also continuously expanding the library. In this thematic area, Aluka is working with a group of researchers at the University of Cape Town.
The *Struggles for Freedom in Southern Africa* collections document the struggle against apartheid and the wars of liberation in Southern Africa. The liberation of Southern Africa from colonial rule was one of the major political developments of the 20th century. The end of Portuguese colonial rule and white settler domination as well as the dismantling of the apartheid regime had far-reaching consequences not only for the people in the region but also for the global community. The initial content focuses on six countries, namely Angola, Botswana, Mozambique, South Africa, Namibia, and Zimbabwe. The national advisory committees, consisting of leading scholars, archivists, and public intellectuals in each of the six countries, joined by a few scholars from outside the region, play a critical role in selecting materials and forging collaborations. Among the documents in this collection include nationalist publications, colonial government reports, local newspaper articles, UN documents, monographs, oral testimonies, speeches, and posters.

**Partnerships and community participation**

Aluka is designed to be a community resource that relies on the participation and contributions of many institutions and individuals across a broad range of disciplines and countries. Such a grand partnership requires mutual trust and commitment by all players at the outset. All partners have to recognize that each contribution is important in setting up a shared resource. Aluka recognizes each institutional and individual contribution by citing contributors appropriately. Having a shared resource promotes and enhances communication and interaction within the research and scholarship community. The public can also have access to resources that were hitherto the domain of subject discipline specialists before the advent of digital libraries.

Aluka has invested in capacity building. In partnership with collaborating institutions, Aluka has set up digitization laboratories at various partner institutions. Training and support for digitization is provided so that partners can carry out the digitization by themselves and share the results with JSTOR/Aluka, as well as the possibility of making these digital assets available on their own sites or via other distribution channels. It is important to note that Aluka doesn’t seek exclusive access to these digital assets. The key issues that have had to be addressed in building Aluka pertain to content development, intellectual property rights, technology and the sustainability of the digital library.

**Content development**

Aluka obtains digital content in general by working with partner organizations. Typically, these organizations carry out the digitization, following standards and best practices that Aluka has provided. Aluka facilitates the sourcing of funds for setting up digital labs and also provides training in digitization in order to achieve efficiency and high standards of output. The results of these efforts are then shared on the Aluka platform and made available to a larger community of scholars and researchers around the world. Participating institutions benefit by having a digital lab and developing digitization skills at the technical and management levels. The second model is where institutions and individuals contribute already digitized materials. This model takes advantage of the network by enabling individuals and institutions to contribute their own scholarly collections and the ability to make them available to larger audiences of scholars, students, and researchers.
Intellectual property rights

Intellectual property rights are addressed at the outset before any digitization or a contribution is incorporated in Aluka. This provides an opportunity for Aluka and the partner to negotiate and fully understand the terms of the partnership. The copyrights of materials in Aluka are retained by the contributors of content. The contributors provide Aluka a non-exclusive license to host and preserve materials for educational and non-profit purposes such as teaching or research endeavors.

Technology

Aluka offers a high quality shared web platform. Building a high quality website requires a great amount of skill and other resources to adequately address the hardware and software requirements. Aluka supports the necessary technical infrastructure, including servers, networking, security, and data storage. The aim is to make the platform sophisticated and scalable but user-friendly. The technology teams ensure that the organizational structure of the data meets the requirements of many institutions and individuals as well as the different kinds of media types. The web site is currently available in three languages (English, French and Portuguese). We have also sought to build into the architecture of the site multiple options in viewing and accessing the data, trying to take into account the bandwidth limitations that exist in various parts of the developing world. Hence a user can download a text version of a digitized document or view a simple jpeg image instead of a higher resolution version. These features while relatively simple to implement, provide users with choices about how they wish to interact with the digital library.

Sustainability

The sustainability of the Aluka collections is built around community participation and contribution. Institutions and individuals are encouraged to make content contributions or participate by providing feedback in the form of corrections, comments, and suggestions. African institutions are given free access to the resource whereas institutions outside Africa are encouraged to sign up and pay modest participation fees to help cover our operating and ongoing technology costs. JSTOR also seeks philanthropic support for capacity building and new digitization projects so that the digital library can grow and expand.

Role of a third party host in aggregation and long term preservation

Arms (1999) observed that ‘long term preservation requires organizations that are committed to the long-term’. These happen to be not-for-profit entities such as scholarly societies, national and university libraries and charitable foundations that are committed to the future, and whose motivation to execute projects is not necessarily for financial rewards. The technical, organization and financial requirements needed to ensure long term preservation are unlikely to appeal to profit-making organizations. For example, shareholder profit and increased revenue generation do not necessarily promote the digital preservation of scholarly material. Much of the innovation in digital preservation of scholarly materials has come from non-profit initiatives such as Portico (www.portico.org) and the CLOCKSS initiative (www.clockss.org). Not-for-profit entities are created to fulfill a particular mission and it is these types of organizations that are more likely to pay special attention to the technical and organization requirements that can guarantee long term preservation and access to the materials (Guthrie 2001). It is also important however that these not-for-profit endeavors provide tangible benefits to various members of the
scholarly communications enterprise. JSTOR provides a unique example insofar as it provides important preservation and archiving benefits to participating libraries, as well as important discoverability and search benefits to participating journals and publishers.

A third party not-for-profit can play a valuable role in aggregating electronic resources of different kinds. By taking a non-discipline centered approach, a third party organization avoids being sucked into discipline specific debates or rivalries between various institutions or groups of scholars. A neutral party can effectively enable interdisciplinary research due to the aggregation of materials from varied disciplines as well as from individuals and institutions. Aluka has materials pertaining to Anthropology, Archaeology, Botany, Geography, History, Law, Sociology, and so on. The effective search and browse functionalities enable users to discover data and information in fields beyond their primary specializations. Many institutions and individuals in Africa fail to share research because of distrust and fear about competitors taking the results and moving ahead of them. Yet, we know from recent trends in technology development that the sharing of knowledge and proprietary information is an important driver for innovation and development.

**Concluding remarks**

The Aluka example demonstrates that there is great potential to advance the online availability of research and knowledge in Africa by focusing efforts on tangible benefits for all members of the scholarly community, adopting technology platforms that are scalable and flexible, genuinely seeking to meet the demands and constraints on the ground, and participating in discussions and workshops such as today’s, where we can benefit from suggestions and input from members of the African publishing and scholarly communities. Working with a mission based not-for-profit organization presents advantages such as long term sustainability and scale that are difficult to achieve when institutions work individually. Our engagements have also shown that there are many concerns (such as intellectual property rights) that tend to hinder progress even when the technology and digitization skills are available. Decision makers at African scholarly and library institutions need to continue addressing these concerns in order to enable better access to data and information and optimize communication among scholars and researchers in their own institutions, countries, and worldwide. The benefits of digitization and the sharing of data and information far outweigh the disadvantages. It is up to institutions to establish enabling policies and for countries to facilitate digitization and access to the World Wide Web in order to ensure that African scholars and researchers can optimally exploit the opportunities made possible by recent advances in information technology.
References


