SOUTH AFRICAN DIGITISATION INITIATIVE (SADI) WORKSHOP

DIGITISATION & DIGITAL LIBRARIES: STANDARDS, BEST PRACTICES, POLICIES AND TECHNICAL REQUIREMENTS

27 FEBRUARY – 1 MARCH 2013

Speakers’ Biographies, Titles and Abstracts of Presentations

Speakers

Marié Botha ........................................................................................................................................... 2
Douwe Drijfhout .................................................................................................................................. 3
Anne-Lise Fourie ................................................................................................................................... 3
Ria Groenewald .................................................................................................................................... 4
Dr. Christian Keitel .............................................................................................................................. 6
Dr. Jens Klump .................................................................................................................................... 7
Roger Layton ......................................................................................................................................... 8
Michele Pickover (together with Charl Roberts) ................................................................................ 9
Charl Roberts (together with Michele Pickover) ................................................................................ 9
Mimi Seyffert ....................................................................................................................................... 10
Heinrich Spingies ................................................................................................................................ 11
Prof. Dr. Hussein Suleman .................................................................................................................. 11
Felix Ubogu .......................................................................................................................................... 12
Biography
Marié started working at the former Technikon SA in 1992 and with the merger between TSA and UNISA, moved to the Unisa Library in Pretoria. She is currently managing the division of Collection Developers of the UNISA library consisting of 18 staff members.

One of the areas that the Collection Developers specialises in, is the development of the Institutional repository. The repository is currently housing quite a number of research output documents as well as archival collections that were digitised and made available as part of the University's social mandate.

The UNISA Library has a symbiotic relationship with the UNISA Archives. The Archives have unique materials and, with its rare content and its superb support services to a wide range of researchers, are a flagship of the University.

In 2011 the Collection Developers identified the need to expand on our collections on the repository and a joint decision between the collection developers and the UNISA Archives was to find archival resources that were highly in demand, for possible digitisation. The official process commenced in 2011 and was completed in December 2012.

Title
Who Keeps the Joker: Collection Development’s role in the digitisation of Archival Resources: UNISA’s experience.

Abstract
At the University of South Africa, we recognised the need that the Library and Archives respond to the user needs as a single entity. At UNISA, we make every effort to service the different demands and needs of the scholarly community, as well as the general public through predominantly open access resources.

The library acquires resources already developed in an e-environment, whilst the archives need to develop digital resources of valuable and rare content. With this in mind, we knew that we needed to work together to develop a channel which is easy to access to provide the needed information at the exact right time and in a digestible format.

This places a huge demand on the skills base and competencies and abilities that are required for this very dedicated and expensive resource development. Access to unique archival resources is a very important aspect in serving the cultural fabric of the community at large.

With this presentation we will share what we consider to be important aspects to address in the working relationship between the library and the archives in reaching the user at the point and exact time of needs. We will touch on basic collaboration that was followed to digitise archival content and will briefly share some aspects that we had to consider in the outsourcing of digitisation of archival resources.
Douwe Drijfhout

Biography
In his capacity as Executive Librarian: Preservation Services he is responsible for preservation and conservation activities at the National Library of South Africa. He was actively involved in the DISA (Digital Innovation South Africa) project.

Title
Reprographic and Digital Services at the National Library of South Africa: an overview

Abstract
An overview will be given of reprographic and digital services at the NLSA. Digitisation activities included the DISA Southern African Freedom Struggles 1950-1994 project, the UNESCO Memory of the World project and the World Digital Library. Research covered digitisation in Africa as well as a review of South African policy making with reference to the preservation of digital collections. Scanning services include ad hoc reprography and digital photography on request to library users. Policy development and the NLSA’s digitisation vision is based on existing legislation and the draft National Policy for the Digitization of Heritage Resources as prepared by the Department of Arts and Culture. Future plans include the digitisation of historical newspapers representing the early years of the Black Press in South Africa. Proposed systems infrastructure (storage and preservation) will be explained.

Anne-Lise Fourie

Biography
Attended University of Stellenbosch, 1977-1980 and received a B. Bibl. in Library Science. Worked as a librarian for 32 years, starting as a librarian at the Council for Scientific and Industrial Research (CSIR) and for the last 17 years in the Mary Gunn Library, South African National Biodiversity Institute (SANBI) at the Pretoria National Botanic Garden.

Currently holds the post of Assistant Director: SANBI Libraries, responsible for: Harry Molteno Library in Kirstenbosch, Cape Town and Mary Gunn Library, Pretoria, 3 herbaria libraries, 9 Botanical garden libraries and 3 Environmental Education Centres’ libraries.

Main job purpose is to manage SANBI Libraries, provide services for all information needs of SANBI staff and stakeholders and to meet demands for comprehensive, easily accessed information on the biodiversity of sub-Saharan Africa. Completed a USB-EDU Junior Management Diploma (2 year), presented by University of Stellenbosch in 2008. Organised and presented the first Biodiversity Heritage Library (BHL) Africa Organisational & Planning meeting at Kirstenbosch in Cape Town, South Africa, 14-15 June 2012. Currently planning the official launch of BHL Africa in April 2013.
Title
SANBI and BHL-Africa: preserving biodiversity literature

Abstract
The Biodiversity Heritage Library is a consortium of natural history and botanical libraries that digitise and make accessible the biodiversity literature held in collections and also literature available open access. Published literature on biological diversity has limited global distribution; much of it is available in only a few select libraries, yet literature about the biota existing in developing countries is often not available within their own borders. Access to published literature is one of the chief obstacles to research.

The BHL consortium works with the international taxonomic community to ensure that biodiversity heritage is made available to a global audience through open access. BHL members digitise public domain books and journals and obtain permission from publishers for materials still under copyright. In partnership with Internet Archive, taxonomic communities and local digitisation efforts, BHL has digitised almost 40 million pages of scientific literature. BHL provides a range of services and APIs which allow users to harvest data files, species information and reuse content for research purposes.

Since 2009, BHL has expanded and gBHL (global biodiversity heritage library) is a network of autonomous members operating programs and projects to make biodiversity literature. They are:

- BHL-Europe
- Chinese Academy of Sciences
- Atlas of Living Australia
- Brazil (through SciELO)
- Bibliotheca Alexandrina
- Coming soon is BHL Africa

BHL Africa is an inclusive network of African libraries and institutions in Western, Eastern, Central and Southern Africa. BHL Africa’s guiding principles and values are: Open Access, Collaboration and Transparency.

SANBI is legally mandated to manage biodiversity information. The Biodiversity Information Management Directorate (BIM) is responsible for the effective management of information and such information can be accessed through the Biodiversity Advisor http://biodiversityadvisor.sanbi.org. As the leading biodiversity institute in Africa, SANBI Libraries has been given the task to officially launch BHL Africa in April 2013.

Ria Groenewald

Biography
Ria Groenewald has been involved in several e-initiatives of the Department of Library Services, University of Pretoria (UPLS) since 2000. Ria is well trained in digitisation and digital preservation, and has published an international award winning paper together with Amelia Breytenbach. She is the Digitisation Coordinator of the UPLS and is responsible for driving the digitisation initiatives as well as for training and marketing thereof.
She presented papers nationally and internationally and does undertake internal and external professional consultation where required. Ria also formed part of the World Digital Library’s technical committee in 2009 and is involved in the National Research Foundation's Carnegie Digitisation project to enhance digitisation in South Africa.

Ria is working on the digitisation of valuable collections housed at the University of Pretoria, which includes projects such as the digitisation of Africana Special Books, Die Huisgenoot, the digitisation of several Veterinary Science projects and several collaboration projects with other institutions. Ria serves as a committee member of the South African Preservation and Conservation Association (SAPCON).

Ria is currently working on digital preservation and the delivery of information on mobile devices.

Title
Digitisation with the aim to long-term preservation

Abstract
South Africa has over the past 5 years seen different initiatives to raise awareness about digitisation especially in the cultural heritage and education sectors, however, a topic that has just recently started to get the necessary attention in our country is that of digital preservation and the long-term sustainability of the original scans.

Data curation forms part of the data created by researchers, gathered in offices and consulting rooms, but the original archival scan of a source is not always seen as something that needs to be preserved in its original file format because the physical paper copy is still available. The vulnerability of libraries was again emphasized by the vandalism of the Library in Mali where valuable manuscripts were stolen or burnt – a lost heritage.

Institutional repositories are mostly populated with material converted from the original to another format (i.e. PDF) or derivate from the original scan to a compressed and thereby smaller version of the original (i.e. TIFF to JPEG). Such formats must not be seen as the archival format but merely the format that is best suitable for web display at the time and period when submitted to the digital repository, and it will be preserved and archived by the repository software as such. We are entering a period where our users start to use their smart phones and/or tablets more often to access information. These devices, however, are coupled with data availability per month. Smaller data sizes do not necessarily mean poorer quality, as the dimensional size of a pixel is now much smaller than the original screen display. We no longer click but scroll, we use HTML5 and CSS3 to create formats such as EPUB, these technologies compel us to go back to the original drawing board.

Internal policies regarding digital collection building should clearly address the need to preserve and sustain the original digitised formats for long-term access. The responsibility of maintaining the ongoing operation and enhancement of digital collection building should be defined as far as
possible when starting a new digitisation project to prevent data from becoming obsolete because of newer advancing technologies.

Dr. Christian Keitel

Biography
Dr. Christian Keitel received his PhD in medieval history in Tübingen and a diploma in archival science at the archive school in Marburg. Since 2000, he is responsible for strategic issues on digital preservation within the Landesarchiv Baden-Württemberg (regional state archive). Between 2006 and 2009, he led a working group which developed a system for the long-term preservation of digital objects (DIMAG) and the open source software IngestList.

His fields of research cover problems of long-term preservation of digital and digitised material (e.g. significant properties or archival description) as well as questions on authenticity and certification (co-leader of the Nestor working group on certification and the DIN working group on trustworthy digital archives).

Dr. Keitel teaches several courses on digital preservation at three German universities.

Title
Trusted Digital Archives. Experiences from the Landesarchiv Baden-Württemberg, nestor and DIN

Abstract
The presentation will address various solutions to users’ needs regarding authentic digital objects. It will describe the approach of a specific archive, the German standards on trustworthy digital archives and the procedure of the Nestor certification.

The most important factor for digital archives is their designated user. The user decides if he can trust the objects of the digital archive and with that he decides about the credibility of the entire archive. Hence, questions related to authenticity and trustworthiness are fundamental to the success or failure of the archive.

We could distinguish two different user groups in this regard. Some users question the authenticity of individual archival objects (group 1), while most users look at the trustworthiness of the archive as a whole (group 2). The digital archive should serve both groups.

The Landesarchiv Baden-Württemberg is a regional state archive in Germany. We have designed and implemented a digital magazine (DIMAG) in 2006 and the free software tool IngestList in 2008 (http://sourceforge.net/projects/ingestlist/). Both tools have some built-in functions to address the needs of group 1.

nestor is the German competence network for digital preservation. It supports the cooperation between the main actors in the heritage sector (archives, libraries and museums). A nestor working group compiled a criteria catalogue on trustworthy digital archives between 2004 and 2008 (English version on http://nbn-resolving.de/urn:nbn:de:0008-2010030806). On this basis, a
follow-up working group published the German standard DIN 31644 Trustworthy Digital Archives in 2012. Both this standard and the original criteria catalogue address the needs of both user groups.

DIN 31644 stands between the Dutch Data Seal of Approval and the ISO 16363 Audit and Certification of Trustworthy Digital Repositories in terms of complexity and requirements. Each of these standardization initiatives has developed its own certification procedure. These three certification procedures compose the European Framework for Audit and Certification of Digital Repositories (http://www.trusteddigitalrepository.eu). Finally, the presentation will outline the central elements of the nestor certification procedure versus DIN 31644.

Dr. Jens Klump

Biography
Dr. Jens Klump earned a degree in geology and in oceanography from the University of Cape Town (UCT) and received his PhD in marine geology from the University of Bremen, Germany.

Dr. Klump is a senior research scientist at the German Research Center for GeoSciences (GFZ) in Potsdam, Germany. His field of research is data driven science. Research topics in this field are numerical simulation of natural gas reservoirs, virtual research environments, remotely operated instruments, high performance and cloud computing, long-term preservation of research data, and the development of system solutions for geoscience projects.

Title
From Ions to Bits – Managing Data in a National Research Centre

Abstract:
Managing data in active research projects is a challenging task. The innovative nature of research requires a flexible data infrastructure that is able to adapt to ad-hoc changes. How can this be reconciled with the necessity to streamline infrastructure services in order to keep cost at a sustainable level? What must data management infrastructure look like to integrate well into the everyday work of a researcher?

In the past the focus of attention has been on large volume research data. However, most research data is small and complex, already highly enriched with contextual information. Managing this “long tail” of research data is labour-intensive and requires new strategies and technological solutions to allow sustainable operation.

Eventually, the results of a project are published in the literature and should be accompanied by data publications. The data, now being part of the record of science, has to be citeable and has to be curated for a long period of time. Data publication and long-term preservation call for new services and for cooperation between infrastructure providers (computing centre) and memory institutions (library).
This talk will investigate the challenges and solutions for managing research data, taking research at GFZ as an example.

Roger Layton

**Biography**

Roger Layton is the CEO of Roger Layton Associates, and also the founder of The ETHER Initiative, which has been created for the pursuit of an eternal heritage using digital technologies. He was the project leader of the National Policy on Digitisation for the Department of Arts and Culture in 2009-2011 and is the primary author of the Digital Heritage Body of Knowledge (DHBOK).

Roger and The ETHER Initiative are part of the consortium for a major EU-funded project called EUROPEANA INSIDE, which is creating a best-practice network of digitisation organisations (including content owners, aggregators and software companies) with particular reference to the EUROPEANA digital library. The ETHER Initiative is the only Africa-based organisation among the 36 partners in this project and the goal is to help to use this experience to contribute to a similar best-practice network in Africa.

Roger has presented many papers at national and international conferences including a paper on “Digital Provenance” at the OpenCulture conference in London in June 2012, and also on “The Principles of the Digital Heritage” at the International Conference on African Digital Libraries in Johannesburg in November 2011. His work includes the development of semantic-based repositories and collection management systems for museums, archives, libraries and the oral history.

Roger has been in the IT industry for 43 years and has largely focused on the digital heritage in the past 8 years.

**Title**

Towards the Digital Institution

**Abstract**

We are living at a unique point in history in which all memory institutions are being transformed from purely physical institutions towards the digital institution. This transformation could not occur in previous generations, since the technologies were not in place to support this change. This transformation should also be largely completed by the next generation in 20 years or so, with all institutions will have a digital structure which reflects and complements their physical structure. It is thus our generation’s responsibility, and perhaps ours alone in the long history of memory recording, to ensure that this transformation occurs properly, so that it benefits all future generations.
This transformation is a one-off event and it encompasses the entire institutional operation and is not confined merely to the capturing digital images and the creation of repositories. Rather, this transformation requires a total re-engineering of the institution, and requires a rapid catch-up in knowledge, capacity and technology within a sector which is largely populated with older specialists within insufficient knowledge of the digital world.

This paper presents some key elements of this transformation with which we are currently engaged to help various institutions through this sea change. Our experience has been that few institutions have started to make significant moves into this digital transformation, although everyone has accepted its necessity, and we are now at the start of the real work which is needed throughout the entire sector. This has led me to identify three major areas of work to be done.

**Michele Pickover (together with Charl Roberts)**

**Biography**
Michele Pickover is the Principal Archivist: Digital Archives and Library at the Wits Library. She is based in the Historical Papers Research Archive and has been involved in digitisation projects, both at an institutional and national level, since 1999.

**Title**
The Wits Digitisation Experience: Lessons Learned and Suggestions for National Collaboration

**Abstract**
The presentation will provide a brief background to the Wits Digitisation Centre, the services it delivers and strategies for the sustainability of the Centre. It will discuss our vision for an Integrated Resource Discovery platform for digital resources in the Library as well as digitisation policy development at Wits. Finally, it will allude to some of the technical and intellectual issues that need to be engaged with for collaboration across institutions and a possible way forward for a South African Digital Libraries and Archives Portal will be proposed.

**Charl Roberts (together with Michele Pickover)**

**Biography**
Charl Roberts is currently the Manager – Library Systems and Digital Initiatives at the Wits Library. His responsibilities include managing all ICT activities and projects in the Library. This includes projects associated with the Wits Digitisation Centre. Charl was a part of the Library team tasked to setup and support the Wits Digitisation Centre.

**Title**
The Wits Digitisation Experience: Lessons Learned and Suggestions for National Collaboration
Abstract
The presentation will provide a brief background to the Wits Digitisation Centre, the services it delivers and strategies for the sustainability of the Centre. It will discuss our vision for an Integrated Resource Discovery platform for digital resources in the Library as well as digitisation policy development at Wits. Finally, it will allude to some of the technical and intellectual issues that need to be engaged with for collaboration across institutions and a possible way forward for a South African Digital Libraries and Archives Portal will be proposed.

Mimi Seyffert

Biography
Mimi Seyffert is currently Head of Special Collections at Stellenbosch University Library and Information Service. In this capacity she heads the Africana collection, the Rare Books Collection and the Manuscripts Section and is also responsible for managing digital projects and initiatives in the section. Mimi holds a B.Bibl. degree from the University of Stellenbosch and has been a Special Collections Librarian at Stellenbosch since 1998. She is passionate about the preservation of South African heritage and history and promoting research in these areas.

Title
Here today... here tomorrow: sustainable digital collections – the Stellenbosch experience

Abstract
The aim of this paper is to share the experiences of Stellenbosch University Library and Information Service in developing a digital collections repository, with specific focus on heritage collections.

The Library’s strategic intervention of supporting, developing and contributing to high-level scholarly publication output includes the strategic goal of digitising unique collections within the Library and Information Service. This goal has recently been given new impetus with Stellenbosch University joining the Research Libraries Consortium and benefiting from a Carnegie Corporation of New York grant for the digitisation of heritage collections.

The paper will look at the various processes we undertook to realise the aims of the project: from selecting and preparing collections for digitisation, liaising with external service providers, choosing a platform to host the content, making the content accessible and ultimately adding value to the collections. I will also discuss the challenges in keeping the integrity of original heritage collections and not jeopardising their context when transforming hard copy to digital. Finally, the paper will examine the ways in which we as the Library and Information Service ourselves have made investments to make these projects sustainable and to ensure not only the maintenance, but also the growth and further development of our digital collections.
Heinrich Spingies

**Biography**

Hein has over 20 years LIS, KM, IM, DM and RM experience. He holds a Master’s degree in LIS obtained from the University of Johannesburg (then RAU). He is currently employed by the National Research Foundation where he heads up the Information Resources & Services unit. Prior to him joining the NRF in 2010, he was employed by the Independent Development Trust where he implemented a KM strategy for the organisation with the aim of using project feedback to design and execute new developmental projects. He was a management consultant with IBM Global Business Services and PwC Consulting and has worked across borders implementing a library system for the Botswana National Library.

Title

**Overview of the Carnegie / NRF Digitisation Project**

**Abstract**

In 2008, the NRF was awarded a grant by the Carnegie Corporation of New York to implement a Digitisation and Preservation Centre. The presentation will provide the delegates with an overview of the project, the accomplishments and challenges experienced to date and the way forward.

Prof. Dr. Hussein Suleman

**Biography**

Hussein Suleman is an Associate Professor in Computer Science at the University of Cape Town, where he directs the research of the Digital Libraries Laboratory. He completed his undergraduate degrees and MSc at the then University of Durban-Westville and finished a PhD at Virginia Tech in the USA in 2002, in the area of component-based digital libraries.

He actively advocates for Open Access in South Africa, and works closely with the Networked Digital Library of Theses and Dissertations (NDLTD), which promotes/supports the adoption of electronic theses and dissertations and generally digital libraries worldwide.

He currently manages the South African ETD portal as well as the international ETD Union Archive. He actively collaborates with the Centre for Curating the Archive at UCT, developing software tools for preservation, dissemination and discovery for the Bleek and Lloyd and related collections. Hussein’s main research interests are in digital libraries, information retrieval, cultural heritage preservation, Internet technology and high performance computing.
Simple Cultural Heritage Preservation

Software systems coupled with digitisation for the digital preservation of heritage have become the norm because of the opportunities for vastly better discovery and dissemination over traditional preservation techniques. However, these systems often are built on the premise that digitisation by itself will lead to preservation and little thought goes into the design of the software systems.

It is proposed that the design of these software systems should be principle-based, where the core of such principles can be extracted from an analysis of successful preservation systems. Systems built on these principles will arguably provide the best platforms for digital preservation of heritage. Among these principles, the notion of simplicity is key. Recent and ongoing work provide many operational examples of how simplicity and related ideas can be incorporated into the design of systems to maintain or increase the level of functionality and, arguably, improve the ability to preserve both content and services in the long term.

Felix Ubogu

Felix Ubogu has been University Librarian at the University of the Witwatersrand, Johannesburg, South Africa since December 2000. In 1998, as University Librarian at Rhodes University, he developed the first digital theses initiative in Africa by making Rhodes University’s theses and dissertations available on the web. He chaired the Advisory Committee of the Database of African Theses and Dissertation (DATAD) Pilot Project from July 2000 to February 2005. From 2003 to 2004 he was the Project Manager of a UNESCO pilot project to improve the management of and access to theses and dissertations prepared at universities in two UNESCO Member States in Africa, Ethiopia and South Africa. He was also instrumental in the development of the South African Electronic Theses and Dissertations project for South African university libraries and has been a key player in the South African Digitisation Initiative.

In 2009 he was one of the organisers of the First International Conference on African Digital Libraries and Archives (ICADLA-1), held in Addis Ababa, Ethiopia. He has served on the Board of the Networked Digital Library of Theses and Dissertations, the Editorial Board of African Journal of Library, Archives & Information Science, as well as for some years as a member of the Board of the South African Library for the Blind.

A Policy Framework for the Digital Library
Abstract
The paper argues that it is essential that an institution has an appropriate policy framework in place in order to ensure the long-term success of its digital library programme and that institutions do not have to reinvent the wheel but can adapt policies already in use by other institutions. Information gathered on digital libraries and repositories shows that many repositories do not have appropriate policies with regard to Content, Preservation, Metadata Re-use and Data Re-use. A comparison of the policies also shows a wide degree of variation, particularly in terms of depth and breadth of coverage. In South Africa, most institutions confirmed that they either had no policy or were in the process of drafting one.