

https://dspace.org/

Report of the South African Digitisation Initiative(SADI) Workshop Digitisation & Digital Libraries: Standards, Best Practices, Policies and Technical Requirements

Pickover, Michele; Mohale, Gabriele

2013-08-26
http://hdl.handle.net/10539/13074

Downloaded from a DSpace repository. https://dspace.org/

REPORT OF THE SOUTH AFRICAN DIGITISATION INITIATIVE (SADI) WORKSHOP

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

27 FEBRUARY - 1 MARCH 2013

Compiled by

Michele Pickover & Gabriele Mohale (Wits Library)

TABLE OF CONTENTS

PURPOSE OF THE WORKSHOP	2
INSTITUTIONAL PARTICIPANTS	2
WORKSHOP OVERVIEW	3
KEYNOTE PRESENTATIONS	3
ISSUES RAISED IN THE LOCAL PRESENTATIONS	5
POTENTIAL AREAS OF ALIGNMENT AND COLLABORATION	6
CONSOLIDATION OF ISSUES RAISED IN GROUP DISCUSSIONS	6
SUMMARY CONCLUSION	7
SADI WORKING GROUP MEETING	7
APPENDIX 1	9

PURPOSE OF THE WORKSHOP

Archives, libraries and museums are increasingly involved in securing the long-term access to social, economic, cultural and intellectual knowledge heritage in digital form and continued access to digital information, now and in the future, is predicated on the implementation of well-constructed strategies, policies, standards and best practice. With this in mind, representatives from institutions in Gauteng: the University of the Witwatersrand, University of Pretoria, the National Research Foundation and the Goethe Institute Johannesburg, agreed to organise a workshop aimed at reaching representatives from key South African institutions involved in digitisation projects and in establishing digital collections, to focus on the exchange of knowledge and experiences and foster co-operation in creating collections and providing online access to South African cultural and heritage resources in digital form.

More specifically, the aim of the workshop was to:

- Engage identified and relevant stakeholders so as to get an in-depth understanding of the current state of digital libraries and archives in South Africa, raise awareness about digitisation in the cultural heritage and education sectors and identify opportunities and challenges.
- Discuss ways of overcoming challenges commonly encountered when establishing and sustaining digital collections. These include the development of policies, strategy frameworks and standards to ensure the sustainability of significant institutional and national digital collections.
- Examine ways of optimising the use of existing digitisation capacity, skills development, pooling digitisation efforts and building on centres of competency.
- Examine ways of building a national portal to South African digital cultural heritage content.

INSTITUTIONAL PARTICIPANTS

The SADI workshop was attended by 48 participants from 20 South African and 3 German institutions and organisations. These included:

Cape Peninsula University of Technology
Desmond Tutu Peace Centre
German Research Centre for Geosciences
Goethe Institute, Johannesburg
Keleketla Library
Landesarchiv Baden-Würtemberg
National Archives
National Library of South Africa
National Research Foundation
Nelson Mandela Centre of Memory
North West University
Rhodes University
Stellenbosch University
The Ether Initiative
Tshwane University of Technology
UNISA
University of Cape Town
University of Johannesburg
University of KwaZulu - Natal
University of Pretoria
University of the Western Cape
University of the Witwatersrand
Walter Sisulu University of Technology

WORKSHOP OVERVIEW

Keynote Addresses were by:

- Dr Christian Keitel, Landesarchiv Baden-Würtemberg, Germany
- Dr Jens Klump, Centre for Geosciences, Potsdam, Germany.

The facilitator was Ulla Wester (Library & Information Services, Goethe-Institut South Africa).

The Workshop consisted of three core sessions:

1. 'Local Experiences and Proposals for the Way Forward'. Facilitated by Alexio Motsi (Preservation, National Archives & Records Service) and Professor Hussein Suleman (Computer Science, University of Cape Town). The presenters were: Anne-Lise Fourie (South African National Biodiversity Institute, SANBI), Prof Hussein Suleman (University of Cape Town), Douwe Drijfhout (National Library of South Africa), Ria Groenewald (University of Pretoria), Felix Ubogu (University of the Witwatersrand), Mimi Seyffert (Stellenbosch University), Marié Botha (UNISA), Charl Roberts and Michele Pickover (University of the Witwatersrand), Roger Layton (ETHER Initiative) and Heinrich Spingies (National Research Foundation, NRF).

2. 'Potential areas of alignment and collaboration'. Facilitated by Dr Daisy Selematsela (Executive Director - Knowledge Management & Evaluation, National Research Foundation).

3. 'Paving the way forward'. Facilitated by Dr Daisy Selematsela (National Research Foundation) and Mr Felix Ubogu (University Librarian, University of the Witwatersrand).

The workshop was concluded with a post-workshop monitoring and evaluation meeting of the nominated SADI Core group.

KEYNOTE PRESENTATIONS

1. Dr Christian Keitel is responsible for strategy on digital preservation within the Landesarchiv Baden-Württemberg. His presentation, *Trusted Digital Archives: Experiences from the Landesarchiv Baden-Württemberg, Nestor and DIN* is available on:
   http://wiredspace.wits.ac.za/bitstream/handle/10539/12501/Trusted%20Digital%20Archives01.pptx?sequence=3
   Dr Keitel leads the Nestor working group on digital preservation. The Network of Expertise in Long-Term Storage of Digital Resources (Nestor) has undertaken efforts to design a catalogue of criteria for trusted digital repositories for long-term preservation. To download the Nestor Criteria Catalogue for Trusted Digital Repositories see: http://files.d-nb.de/nestor/materialien/nestor_mat_08-eng.pdf.
   Dr Keitel emphasised the need for authenticity of digital objects, i.e. the object must contain what it claims to contain, and that a trustworthy repository allows the user to trust in the authenticity of the digital objects. Dr Keitel highlighted the DIN (German Bureau of Standards) Standard 31644 on trustworthy digital archives, which was published in 2012 and has 34 criteria which could be used by archives, both for self-

evaluation (what is missing to be a trustworthy archive) and for certification (what could be said about the trustworthiness of my archive?). The standard considers the setting of a digital archive as a multilevel process, beginning with planning, followed by specification, implementation and considerations about further improvements. The 34 criteria are grounded in four principles: documentation, transparency, adequacy and measurability.

2. Dr Jens Klum is the e-Science Project Manager at the Helmholtz Centre Potsdam German Research Centre for Geosciences**.** His presentation, *From Ions to Bits – Managing Data in a National Research Centre***,** is available on: http://wiredspace.wits.ac.za/bitstream/handle/10539/12502/Klump%20SADI%20Workshop%2020130227.pdf?sequence=1
Dr Klump highlighted the challenges of research and scientific data management and he emphasised the need for proper metadata and workflows so that we are able to deal with, and access over the long-term, the ever-increasing amount of raw research data which is being created. If this is not done we will face the problems of: inaccessible data, data that cannot be re-used, duplication of efforts and results that cannot be verified. To do this he suggested that: dealing with data has to be easy and attractive to researchers, data management tools need to be integrated into the workflow of researchers and that data management tools need to be flexible, scalable and customisable.

Both keynote presentations responded to questions from the floor on: metadata handling and ownership, quality loss in the process of migration, the management of research data in university libraries, and the evolution of libraries as data repositories and the responsibilities that come with this function.

LOCAL PRESENTATIONS

The session 'Local Experiences and Proposals for the Way Forward' included the following presenters – see also the links to their presentations on the Wits Institutional Repository:

- Botha, M., *Who keeps the Joker? Collection Development's role in Digitisation of Archival Resources: UNISA's experience.* http://hdl.handle.net/10539/12497
- Drijfhout, D., *Reprographic and Digital Services at the National Library of South Africa: an overview.* http://hdl.handle.net/10539/12498
- Fourie, A., *SANBI and BHL-Africa: preserving biodiversity literature* http://hdl.handle.net/10539/12499
- Groenewald, R., *Digitisation with the aim to long term preservation*
- Layton, R., *Towards the Digital Institution.* http://hdl.handle.net/10539/12503
- Roberts, C. & Pickover, M., *The Wits Digitisation Experience: Lessons Learned and Suggestions for National Collaboration*, http://hdl.handle.net/10539/12505
- Seyffert, M., *Here today… here tomorrow: sustainable digital collections – the Stellenbosch experience.* http://hdl.handle.net/10539/12506
- Spingies, H., *Overview of the Carnegie / NRF Digitisation Project.* http://hdl.handle.net/10539/12507
- Suleman, H. Prof., *Simple Cultural Heritage Preservation.* (http://hdl.handle.net/10539/12499
- Ubogu, F., *Policy Framework for the Digital Library.* http://hdl.handle.net/10539/12509

ISSUES RAISED IN THE LOCAL PRESENTATIONS

The presentations from local South African institutions can be grouped into the following five areas:

1. Processes and Standards (including technical requirements)
    - Seeking simplicity and minimal technical standards as solutions for digitisation projects in South Africa, i.e. use of XML encoding for harvesting.
    - Selection criteria and preparation of material before digitisation can take place.
    - Detailed description of material selected for digitisation before digitisation takes place.
    - Detailed finding aids to archival collection a prelude to successful digitisation.
    - Use of best practice and international standards and guidelines for digitisation projects.
    - Outsourcing *vs*. in-house digitisation?
    - Quality assurance for outsourcing through sound contracts with service providers and service level agreements.
    - Consideration of differences between description methodologies in libraries, archives and museums.

2. Access and Copyright
    - Equal partnerships.
    - Copyright and free access issues in overseas funded digitisation of South African heritage.
    - Access to digitised collections through library catalogues and portals.
    - The challenge of provision of open access to digital resources and the protection of copyright.
    - Compliance with South African legislation when providing access to digitised material, such as the Protection of Personal Information Bill and Copyright.

3. Skills and Training (including human resources)
    - Challenges and problems of investing in the education & training of archival staff.
    - Developing digitisation into a core skill in memory institutions.
    - Need to develop digitisation and management of digitisation project skills in memory institutions.

4. Digitisation policy (including. decision taking at institutional level)
    - DSpace as a suitable and sustainable platform for some digitised material, but not all.
    - Visible policies and standards for digitisation at institutional level, including IT storage and security measures and standards.
    - Conservation and preservation of original items as an essential component of institutional digitisation policy.
    - Digitisation as a way to sustain a heritage department through outreach to remote users, as well as a component of open distance learning.
    - Processes for the outsourcing of digitisation.
    - Storage systems are governance decisions: The Cloud, server, file formats, etc.
    - Role of the future National Digitisation Policy as a guiding document.

5. Collaboration and Lobbying
    - Resource sharing for digitisation.
    - Fragmentation of the heritage material landscape in South Africa, and resulting need for an updated digitisation audit.
    - Lobbying for digitisation, particularly funding needs and skills training.
    - Legal deposit for digital publications.
    - The framework developed by the NRF incorporating the establishment of a national portal for publications, a digital initiatives register, to be a heritage repository and to advice on training modules.

POTENTIAL AREAS OF ALIGNMENT AND COLLABORATION

The National Research Foundation put forward the following six areas of possible alignment and collaboration:

1. Legal – including collective licensing; legal deposit of digital material; extended collective licensing for cross-border access.
2. Organisational – including good practice; regions that have not been well presented in the digital field; collaboration; shift from projects to programs to core activity or organisations nationally.
3. Standards – including delineating interoperability standards; platform-agnostic digital preservation requirements, move from a repository-centric worldview; standardisation of requisite skill sets, codes of practice; engagement of users in determining if these standards are relevant.
4. Technical – including the development of evaluation protocols and benchmarking; approach interoperability rigorously, understanding of the value of inter-system communication.
5. Economic – including an awareness about sustainable digital preservation; establishment of a digital preservation resource centre; development of case studies for costing; define selection criteria to build a common framework for content selection decision (use triage chart); to study and promote community sourced solutions and explore opportunities for public-private partnerships; define core services; support research and development.
6. Education – including the development of an international certificate program in digital preservation; accreditation of curriculum, providers and metrics, benchmarking courses and cooperation between international providers of education and training; addressing the supply and demand for qualified trainers and to develop a pool of qualified trainers; engage with employers and professional bodies, to endorse skills or identify gaps that need to be addressed; improve cooperation in defining skills-sets and develop ways to education and training options.

CONSOLIDATION OF ISSUES RAISED IN GROUP DISCUSSIONS

1. Need for cross-sectorial (LAMs) collaboration and coordination
    - What structural form should this collaborative initiative take?
    - Name of this initiative? (Forum, Consortium, Association, NGO)
    - Membership
    - A core group be “elected” or mandated to set up this initiative

2. Proposed Priorities and Activities of Proposed Initiative
    - Guidelines and standards (Digitisation process and procedures, Copyright, Tools, Metadata, Content selection, Digitisation strategy, Policies)
    - Training
    - Portal – Two-pronged: Resource and Information sharing (Share content, Resource discoverability, Open Access)
    - Fundraising for Activities of Initiative and sustainability of Resource
    - Lobbying and Advocacy (At national level: HESA, Parliament, Networking: regionally and internationally, Curriculum changes to accommodate digitisation requirements and skills)
3. Open Issues For Further Discussion
    - How should the differences in methodologies etc. be bridged between libraries, archives and museums
    - Lack of inventories, lack of arrangement and description of collections and items under threat and at risk
    - Lack of Institutional resources (human resources and equipment)
    - In relation to content available through the Portal: (What content? Scope of content? What is meant by Memory? Will what will be included in the portal resource influence our ability to fundraise and if so how?
    - Should data management be part of this initiative's ambit?
    - Role of NRF?

SUMMARY CONCLUSION

Workshop participants concluded the workshop with the agreement that SADI should continue its work as a national initiative under the adopted name 'South African Digitisation Initiative (SADI).

It was suggested that a Working Group should be formed, consisting of the SADI preparatory group and three additional volunteers. The workshop mandated the Core group to get the initiative started.

The workshop participants thanked the organisers not only for a useful workshop and for kick starting this national digitisation initiative.

SADI WORKING GROUP MEETING

The SADI Working Group consists of the following members:

Bailey, Razaan (Desmond Tutu Peace Centre)
Groenewald, Ria (University of Pretoria)
Khoza, Thuli (NRF)
Mohale, Gabriele (Wits University)
Motsi, Alexio (National Archives)
Pickover, Michele (Wits University)
Selematsela, Dr Daisy (NRF)
Suleman, Prof Hussein (University of Cape Town)
Ubogu, Felix (Wits University)
Wester, Ulla (Goethe Institute)

It was agreed that:

1. SADI should be an open forum, consisting of institutions and individuals who are prepared to assist and offer expertise.
2. The type of entity/structure SADI should be needs to be explored.
3. SADI should provide leadership in terms of digitisation guidelines, standards and processes. To this end a national resource portal should be initiated. The Report of the proceedings of the workshop will also be made available on a new SADI platform https://sites.google.com/site/sadi2013wg/. Ria Groenewald is the contact person. Working documents, policies, links to resources etc. can now be put up on the website, as well as contact details for working group members.
4. SADI should create a national resource of digital heritage materials through a searchable portal. The portal should also be an information resource for digitisation processes, best practices, expertise, general discussion etc.
5. SADI initiatives should speak to the draft National Policy on the Digitisation of Heritage Resources and could eventually feed into the structures suggested by the draft National Policy on the Digitisation of Heritage Resources. SADI guidelines etc. could also provide the necessary authority for standards and processes.
6. SADI needs to coordinate various functions such as training and lobbying. Members of the Working Group should assume responsibilities for specific immediate tasks.
7. In preparation of the SADI Resource Portal and for harvesting of metadata purposes, an assessment needs to be done of the different schemas, descriptive methodologies etc. used in various participating institutions.
8. The structural form of SADI will have to be explored and SADI's Mission, Vision and Goals need to be determined. For the vision and goals put forward by the initiating Gauteng grouping see **Appendix 1**.
9. Ideas for a SADI logo need to be investigated.
10. Collaborations and contacts with other, including international, institutions with digital resources has to be investigated to establish a network of participating entities.
11. The sustainability of SADI needs to be investigated with the view that it needs to be linked to the establishment of guidelines and structures as set out in the Draft National Policy on the Digitisation of Heritage Resources. A formal inaugural meeting of SADI needs to take place where the recommendations and outcomes will be presented.

APPENDIX 1

AIMS AND OBJECTIVES OF SADI

Vision:
To establish a National Digitisation Initiative that will engage with all the core elements of digitisation, including long-term digital preservation.

Goals:
1. Establish a South African National Digitisation Initiative and engage identified and relevant stakeholders to get an in-depth understanding of the current state of digitisation initiatives.
2. Include in SADI all custodians of heritage, research and cultural information, including, but is not limited to: libraries, archives, museums and private collections.
3. Make digitised material available on an open access basis on the Web.
4. Ensure interoperability for continued access.
5. Secure access to information with the ultimate goal of supporting research, innovation and development in general.
6. Coordinate all digital initiatives currently running in South Africa.
7. Create awareness towards long-term digital preservation.
8. Develop policies, strategy frameworks and standards to ensure interoperability and sustainability of the digital material.
9. Optimise the use of existing digitisation capacity, pool digitisation efforts, identify training needs, and develop expertise and skills.
10. Navigate legal issues such as copyright, privacy and intellectual property
11. Lobby policymakers, negotiate funding and apply for grants, by providing a strong national platform.
12. Engage identified and relevant stakeholders of digitisation initiatives in South Africa.
13. Examine ways of building a national portal to South African digital cultural heritage content.

The South African Digitisation Initiative would assist in:
- Providing a framework for online access to South Africa's heritage.
- Establishing an on-going process for communications to educate about digitisation.
- Identifying opportunities to collaborate co-operate and share information with each other.
- Building collections of digital cultural heritage resources nationally.
- Avoiding duplication of effort.
- Providing access to expertise.
- Resource sharing.
- Profile-raising for individual organisations.
- Improving access to unique and commonly held collections.
- Providing better access to collections and institutions.
- Sharing information and developing expertise in the regions and nationally.
- Negotiating funding and apply for grants, by providing a strong national platform.
- Providing an opportunity to develop common standards for interoperability
- Linking people and linking data.
- Navigating legal issues such as copyright, privacy and intellectual property
- Identifying training needs.