

---

**An Audit of Digitisation Initiatives, Ongoing and Planned,  
in South Africa**

**Report to the National Research Foundation**

---

**Addressed to:**

Dr Daisy Selematsela  
Executive Director: Knowledge Management & Evaluation  
National Research Foundation (NRF)

**Prepared By:**

Roy Page-Shipp  
082 447 6289  
[roy@pageshipp.co.za](mailto:roy@pageshipp.co.za)  
31 August 2009.

## Structure of the Report

The Sections of the Report comprise an Executive Summary and 5 Main Groups:-

	Executive Summary
Sections 1, 2 & 3	Introduction, the Brief and Modus Operandi for the Audit
Section 4	An Overview of the Audit Findings
Sections 5-12	Detailed Audit Findings
Sections 13 - 17	Recommendations
Sections 18	References
Section 19 – 25	Appendices A-H

## Glossary of Acronyms

2D	2-Dimensional	NIKSO	National Indigenous Knowledge Systems Office
3D	3-Dimensional		
AACR	Anglo American Cataloguing Rules	NRF	National Research Foundation
ASSAf	Academy of Sciences of South Africa	OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
DACST	Dept of Arts, Culture, Science & Technology	OBIS	Ocean Biogeographic Information Systems
DST	Dept of Science and Technology	OCR	Optical Character Recognition
DAF	Data Audit Framework	OPAC	Online Public Access Catalogue
DANS	Data Archive and Networking Services	OSS	Open Source Software
DTD	Document Type Definition	pdf	portable text format
EAD	Encoded Archival Description	QTVR	Quick Time Virtual Reality
HEI	Higher Education Institution	RARI	Rock Art Research Institute
HPCC	High Performance Computing Centre	RDA	Resource Description and Access
HSRC	Human Sciences Research Council	RIM	Robben Island Museum
ILAM	International Library of African Music	SABTG	South African Basic Text Glossary
ISAD	International Standard for Archival Description	SANCOR	SA Network for Coastal & Oceanic Research
IT	Information Technology		
JISC	Joint Information Services Committee	SANReN	South African Research Network
JSTOR	Journal Storage	SANSA	South African Network of Skills Abroad
KZN	KwaZulu-Natal		
LIASA	Library & Information Association of South Africa	SARADA	South African Rock Art Digital Archive
LOCKSS	Lots Of Copies Keeps Stuff Safe	TEI	Text Encoding Initiative
METS	Metadata Encoding and Transmission Standard	tiff	tagged image file format
		UNISA	University of South Africa
Mp3	MPEG-1 Audio Layer 3.	VRA	Visual Resources Association
NDPSC	National Digitisation and Preservation Support Centre	wav	waveform audio format
		WWW	World Wide Web
		XML	eXtensible Mark-up Language

# Contents

<b>Executive Summary .....</b>	<b>7</b>
<i>Findings and Recommendations .....</i>	<i>9</i>
1: <i>Identification of Priority Themes for Digitisation Projects.....</i>	<i>9</i>
<i>Recommendations .....</i>	<i>10</i>
2. <i>A National Heritage Repository.....</i>	<i>10</i>
<i>Recommendation.....</i>	<i>11</i>
3: <i>Training and Skills Development .....</i>	<i>11</i>
<i>Recommendation.....</i>	<i>11</i>
4: <i>Infrastructure Needs.....</i>	<i>12</i>
<i>Recommendations .....</i>	<i>12</i>
5: <i>A National Portal to Heritage Collections .....</i>	<i>13</i>
<i>Recommendations .....</i>	<i>13</i>
6. <i>A National Digitisation and Preservation Support Centre .....</i>	<i>13</i>
<i>Recommendation.....</i>	<i>14</i>
7: <i>Permissions to Digitise: Best Practice Guidelines for South African Practitioners.....</i>	<i>14</i>
<i>Recommendation.....</i>	<i>15</i>
<b>1 Introduction.....</b>	<b>16</b>
<b>2 The Brief .....</b>	<b>16</b>
<b>3 Modus Operandi for the Audit .....</b>	<b>17</b>
<b>4 Audit Findings – A Perspective on the Digitisation Status Quo .....</b>	<b>18</b>
<b>5 Audit Findings –The Database .....</b>	<b>19</b>
5.1 <i>Identification of Likely Respondents.....</i>	<i>19</i>
5.2 <i>Collection Coverage Achieved .....</i>	<i>20</i>
5.3 <i>Completeness of data submitted .....</i>	<i>20</i>
5.4 <i>The Database – Conclusions and Recommendations .....</i>	<i>22</i>
<b>6 Audit Findings – Collections and their Status.....</b>	<b>22</b>
6.1 <i>Who is digitising, where and why?.....</i>	<i>23</i>
6.2 <i>Scale and scope of digitisation initiatives. ....</i>	<i>24</i>
6.3 <i>Size Distribution of Collections.....</i>	<i>25</i>
6.4 <i>Collections and their Status - Conclusions and Recommendations.....</i>	<i>25</i>
<b>7 Audit Findings - Focus and Prioritisation .....</b>	<b>25</b>
7.1 <i>Prioritisation of Digitisation Projects.....</i>	<i>25</i>

7.2	<i>National Priorities as a Guide to Themes for Prioritisation.</i>	26
7.3	<i>Stakeholder workshop suggestion for themes.</i>	27
7.4	<i>A Tentative List of Digitisation Themes.</i>	28
7.5	<i>Justification for Digitising.</i>	29
7.6	<i>Focus and Prioritisation – Conclusions and Recommendations.</i>	30
<b>8</b>	<b>Audit Findings – Ownership, Copyright and Licensing.</b>	<b>30</b>
8.1	<i>Bad Experiences.</i>	32
8.2	<i>A National Policy on Digitisation (especially for Heritage Material).</i>	34
8.3	<i>The Digital Divide.</i>	35
8.4	<i>Possession, Ownership and Rights.</i>	35
8.5	<i>Realism, Royalties and Repatriation.</i>	36
8.6	<i>Bridging a gap with a moratorium.</i>	37
8.7	<i>Ownership, Copyright and Licensing – Conclusions and Recommendations.</i>	37
<b>9</b>	<b>Audit Findings – Staff.</b>	<b>38</b>
9.1	<i>Spectrum of skills needed.</i>	39
9.2	<i>Staff -Conclusions and Recommendations.</i>	41
<b>10</b>	<b>Audit Findings – Systems and Standards.</b>	<b>41</b>
10.1	<i>Discovery of the Digitised Images on the WWW Metadata.</i>	41
10.2	<i>Reusability of images, to minimise repeat work.</i>	42
10.3	<i>Interoperability.</i>	42
10.4	<i>Overall.</i>	42
10.5	<i>Systems and Standards – Conclusions and Recommendations.</i>	43
<b>11</b>	<b>Audit Findings – Technology and Physical Infrastructure.</b>	<b>43</b>
11.1	<i>A National Portal.</i>	44
11.2	<i>Sharing of Specialised and Sophisticated Equipment.</i>	44
11.3	<i>Back-up.</i>	45
11.4	<i>Software Selection and Innovation Management.</i>	45
11.5	<i>Technology and Physical Infrastructure – Conclusions and Recommendations.</i>	45
<b>12</b>	<b>Audit Findings – Funding and Sustainability.</b>	<b>46</b>
12.1	<i>Funding.</i>	46
12.2	<i>Sustainability.</i>	47
12.3	<i>Funding and Sustainability: Conclusions and Recommendations.</i>	47
<b>13</b>	<b>Recommendation: A National Portal to Heritage and Research Collections (including Development of a Collection Database).</b>	<b>48</b>

13.1	<i>Criteria for a National Portal</i> .....	48
13.2	<i>The Maintenance and Expansion of the Audit Database</i> .....	48
13.3	<i>Resources for Establishing the Portal</i> .....	49
<b>14</b>	<b>Recommendation: A National Digitisation and Preservation Support Centre (NDPSC)</b> .....	<b>49</b>
14.1	<i>Context emerging from the study</i> .....	49
14.2	<i>Preferred Option – A Virtual Centre of Excellence in Digitisation and Preservation</i> .....	51
<b>15</b>	<b>Recommendation: Funding for Sustainability</b> .....	<b>53</b>
<b>16</b>	<b>Recommendation: Permissions to Digitise: Best Practice Guidelines for South African Practitioners</b> .....	<b>53</b>
16.1	<i>Context</i> .....	53
16.2	<i>General Principles and Definitions</i> .....	54
16.3	<i>Decision Steps</i> .....	57
<b>17</b>	<b>Recommendation: Focus and Prioritisation</b> .....	<b>59</b>
<b>18</b>	<b>References</b> .....	<b>60</b>
<b>19</b>	<b>Appendix A: The Context for the Audit</b> .....	<b>62</b>
19.1	<i>eResearch and Digitisation</i> .....	62
19.2	<i>Some established digitisation initiatives</i> .....	63
19.3	<i>Challenges to the Digitisation Community</i> .....	63
19.4	<i>A Collaborative Digitisation and Preservation Initiative</i> .....	64
19.5	<i>Actions that gave rise to the Audit</i> .....	64
<b>20</b>	<b>Appendix B: Digitisation and Preservation Overview</b> .....	<b>66</b>
20.1	<i>The Players</i> .....	66
20.2	<i>The Principles</i> .....	67
20.3	<i>Rights to digitise</i> .....	67
20.4	<i>Means to Digitise</i> .....	69
20.5	<i>Preservation Issues</i> .....	72
20.6	<i>Marketing and Commercialisation</i> .....	73
20.7	<i>Skills and People</i> .....	73
<b>21</b>	<b>Appendix C: Submissions by Organisation</b> .....	<b>76</b>
<b>22</b>	<b>Appendix D: Respondents to the Questionnaire and their e-mail Addresses</b>	<b>78</b>
<b>23</b>	<b>Appendix E: MSExcel Version of Questionnaire</b> .....	<b>82</b>

<b>24</b>	<b>Appendix F: Staff Levels by Organisation .....</b>	<b>90</b>
<b>25</b>	<b>Appendix H: List of Contacts .....</b>	<b>92</b>

---

## Executive Summary

---

This audit is the first phase of an extensive exploration of the national development needs for effective and efficient Digitisation and Preservation of valuable Heritage and Research collections in the South African system. Sponsored by the Carnegie Corporation of New York, and managed by the National Research Foundation, the exploration was initiated following a Digitisation Stakeholders' meeting on 24 January 2008, in Johannesburg. It is driven by the impression that Digitisation and Preservation needs significantly exceed the available resources and that a more detailed perspective will enable more focused application of capacity building and project funding that will upgrade Digitisation and Preservation practice.

The brief for the audit listed the following specific elements:

- identify whom and which institutions are involved in digitisation projects;
- identify potential projects for digitisation in South Africa;
- establish on the basis of frequency, priority areas/content for potential digitisation projects, and provide a preliminary list for discussion with stakeholders;
- provide a list of current/ongoing projects, to facilitate reduction of unnecessary duplication;
- reaffirm as appropriate the merits of a digitisation centre, its activities and potential governance structure;
- collect data and information for the establishment of a portal for digitisation projects that will ensure accessibility;
- identify levels of expertise/skills availability in digitisation, across institutions;
- establish digitisation needs, i.e. skills, expertise, knowledge, infrastructure requirements; and
- generate a national database (to be updated regularly) and hardcopy report of the audit for distribution to and access by key stakeholders. This would describe the
  - content of the collections
  - types of material
  - coverage of subject areas
  - motivation for digitising each collection
  - status of the levels of expertise in the holding organisation
  - availability of software, hardware and systems
  - connectivity at the institutions, and
  - policies and procedures presently followed in relation to digital data.

'Digitisation' or the digital capture of text or other descriptive detail of items in valuable collections is driven by two forces:

- The need to have digital descriptions of the items that can be widely shared via the World Wide Web, (**Access**) and
- The creation of digital replicas of the items to insure against deterioration of the originals. (**Preservation**).

In this it differs from one of the major elements of modern eResearch<sup>1</sup>, viz Curation of Research Data, because these latter data tend to be of recent origin and are ‘born digital’. Once the digitisation process has reached the point where well organised digital collections have come into being, the data management and preservation techniques in Data Curation will apply to them as well. Prior to that a great deal of work must be done and it is such work that is the focus of this study.

It was clear from the outset that the audit involved two major elements, viz

- **Data Collection** – to establish the database, although it was appreciated that, in the short time available, the result would not be complete and provision would need to be made to continue to build up the database incrementally over an extended period.
- **Identification of Key Dynamics** in the system(s) and the extraction of meaningful recommendations from them.

Despite the shortcomings of the database, which are expanded on below, the combination of the data collection process with many follow-up discussions with role-players has yielded significant insight that will hopefully assist decision-making. A draft report was submitted to a Stakeholders’ Workshop on 18 March and the feedback that was received has been incorporated in this report.

The prescribed methodology for assembly of the database was by e-mail requests for completion by respondents of a Web questionnaire, which were followed-up with the submitters to sort out gaps and ambiguities. The research period, 21 October 2008 to 9 February 2009, was 15 weeks, of which 3-4 would be relatively sterile because of the year-end break in most institutions, especially universities. Many contributions came in during January and **40** institutions had responded by the deadline (for reporting purposes) of 27 February, describing **185** collections of varying sizes. The estimated number of digitisable items visible in or through the survey exceeds 2.7M. A number of promises of submissions – at least some of which are believed to be reliable - had not been kept by the deadline and leads to additional sources are continually being uncovered. The database had increased somewhat by the end of March 2009, when the Stakeholders’ Workshop draft of this report was submitted and still more have been received since, but the additional items are not reflected in this analysis. They have, however, been<sup>2</sup> taken up in the open access web-based database which was one of the follow-up products to the study.

The web questionnaire approach had certain limitations. Time pressures prevented a thorough prototyping and this exacted a price. Not all the questions were equally well understood and some respondents clearly could not or would not specify, for example, the

---

<sup>1</sup> eResearch is the use of information and communications technology to enable new, better research by taking advantage of global connectivity to share research processes and products, whether computational, reported text or databases.

<sup>2</sup> Available at <http://stardata.nrf.ac.za/html/database.html>



exact numbers of staff or amounts of funding available. Only some of the gaps have been filled in by follow-up discussions. Organisations owning very large numbers of collections abbreviated or aggregated their responses. There was also some duplication of collections where reports included all the aggregated collections on a particular topic (e.g. Rock Art), some of which had also been reported individually. The submission format is being reviewed in the light of experience and a simplified version will be used in an ongoing data collection process.

The potential for extracting detailed information from the database without significantly more work is thus limited for the above reasons, but a more important aspect relates to the very dispersed nature of many of the collections. Impressive examples of digitised archives, such as the Bleek – Lloyd archive, have high research value because of their relative ‘completeness’, the records of the Bleek – Lloyd team, for example, having been assembled from several university archives. The danger of ‘cherry-picking’ that focuses on only one subset, are significant; greater transparency of the various catalogues will be essential if the most complete archival sets are to be assembled. This assembly task, often a matter of complete cataloguing, is an essential precursor to digitisation in many cases and represents a potentially major funding need. It was also an unexpected finding of the audit. The perception that there are many ‘scanner-ready’ collections in South Africa is misleading; many of the collections will require significant work before scanning can commence.

The following findings and recommendations emerged in response to the brief. They are based on information and insight gathered from the two major aspects of the Audit:

- creation and population of the Database and
- interviews with major players in the system, focused on the various aspects of the brief,

The order of presentation is similar to the sequence in which the topics are listed in the brief, with some adjustments to support a logical presentation of ideas.

## **Findings and Recommendations**

### **1: Identification of Priority Themes for Digitisation Projects**

*The audit reveals a daunting quantity of material, both explicitly and implied, that is deemed by its custodians to be worthy of digitisation. Resource constraints will necessitate choices, both within institutions, and within funding bodies. There is also a need to avoid overlap. Against this one must consider the autonomy of the organisations holding the collections.*

*The motivation for digitising, viz Access and/or Preservation should be the starting point. Collections that are threatened with imminent decay should clearly have high priority, especially when they fall within a category that has been prioritised. An inclusive prioritisation process needs also to be designed.*

## **Recommendations**

- i. *That a representative prioritisation Task Team, under the Chairmanship of a respected figure in the system, be appointed to identify priority categories.*
- ii. *That the Task Team should use, as a starting point, the list of priority topics identified in the study, including in the Stakeholders' Workshop, viz*
  - a. *Department of Science and Technology National Priorities*
    - *“Farmer to Pharma” strengthening the bio-economy value chain*
    - *Space science and technology*
    - *Energy security*
    - *Global change science with a focus on climate change*
    - *Human and social dynamics*
  - b. *Department of Arts and Culture Strategic Plan*
    - *Heritage Sites (Note that Rock Art is already being comprehensively dealt with)*
    - *Indigenous Languages*
  - c. *From the Stakeholders' Workshop*
    - *Project specific, e.g. Rivonia trial records*
    - *Records at risk: e.g. National or United Party records*
    - *Broad categories: Natural history and biodiversity.*
    - *Art and architecture*
    - *Economic development and trade*
    - *Cultural and political history*

*(NB The Freedom Struggle and Indigenous Music are already being dealt with by DISA and ILAM respectively)*

*The Database can also be consulted for possible further categories.*

- iii. *Transparency in prioritisation is obviously desirable and a method for prioritising collections within categories or institutions is given in 7.4*

## **2. A National Heritage Repository**

*The scale and complexity of the digitisation projects vary enormously: from 'routine' ETD repository creation at the 'simple' end of the spectrum, to digitisation of tens of thousands of documents and artefacts of diverse shapes and sizes, sourced from half a dozen owners, each with a unique view of value and permissions, at the 'complex' end. Not all the institutions with collections to be digitised will be able to create and sustain online repositories: of the 185 collections listed in the audit, only 43, representing 29 organisations, provided web addresses for their collections. This suggests that a National Repository would have a base to start from and considerable opportunity to host many more of the collections. This latter feature would also enable custodians in institutions that are difficult to access via the internet, to position a mirror site on the server that carries the National Portal. (See Findings and Recommendations #5) The Repository would also provide a home for*

*collections digitised by independent researchers. In due course it could be linked, via the Portal to all other repositories in South or Southern Africa and thus form part of a virtual online collection of heritage items.*

*This repository should be hosted at an established, well-governed institution, with sufficient autonomy to ensure sustainability.*

### **Recommendation**

- iv. *That a National Heritage Repository be established at a selected institution to hold collections for which no institutional repository is available.*

## **3: Training and Skills Development**

*The principal obstacle to smooth and rapid development of capacity to digitise the backlog of collections and establish a sustainable system for preserving and marketing the digital collections is the lack of informed and experienced practitioners of all types.*

*Pockets of expertise exist in various institutions, but most are sub-critical, with precious expertise bound up in only one, or at most two, individuals; often the ones whose enterprising spirit has established the digitisation programme. These persons are critical to success in their current environments.*

*The need arises in essentially two categories:*

- *existing staff in institutions who will be running digitisation projects in the short term to catch up with back-logs. They will require guidance and training that will assist them to upgrade their skills. Training programmes aimed at these candidates are likely to attract large numbers of trainees in the first few years, but demand will tail off as the skills spread and the backlog of digitisation projects is worked off, and this should be taken into account in planning.*
- *'Digital Librarians' who can take charge of the planning and management of projects and the development of repositories. The needs of existing incumbents of these roles can be met by workshops and networking.*

*Provision should also be made for tertiary Departments of Librarianship to include the planning and management of digitisation in degree courses.*

### **Recommendation**

*That three training interventions be undertaken:*

- v. *Short, start-up courses for librarians, archivists and others who wish to become familiar with the basics of digitisation.*
- vi. *Professional development support for persons who will be taking charge of significant digitisation exercises to enable them to perform as Digital Librarians. This will be best achieved by workshops and the creation of best*

*practice guidelines (possibly emerging from the workshops). There is more than enough talent and experience in the digitisation community to furnish the necessary expertise.*

- vii. *University level Professional Education for young aspirants who wish to make a career in Digital Librarianship or Archiving.*

#### **4: Infrastructure Needs**

*Basic digitisation equipment for flat specimens  $\leq$  A4 in size, is widely available, but specialised or high production equipment is rare and then tends to be fully occupied. Many teams have large scanners (A3 or larger).*

*Sharing of sophisticated equipment is taking place on a limited scale, but is unlikely to be a good long-term solution to resource optimisation.*

*Where a large quantity of material has been assembled, from within or across institutions, consideration should be given to driving the production of digitised images as a factory-type operation, with attention to appropriate staff skills, production targeting and quality control. These do not often feature in the day-to-day running of libraries and museums and specialist advice may be required.*

*There is interest in 3D techniques, but only one large scale facility, which is presently used for site-scanning although it could, in principle, be used for smaller objects. Facilities for object copying in the automotive industry are also potentially available for smaller items, such as fossils.*

*There is enough skill and experience in the system to set up a South African portal, which will enable minor players to make their material more readily available and to give a South African 'brand' to our heritage and research material. This is dealt with in more detail under 5.*

*Software choices abound, and some standardisation can have benefits. More important is the availability of good objective advice, although the history of software in most application areas inclines towards diversity rather than standardisation. Some degree of coordination or provision of expert advice could be provided by a user groups convened by National Digitisation and Preservation Support Centre. (See 6)*

#### **Recommendations**

- viii. *That virtually every library, museum or other institution with heritage items, should be provided with a basic digitisation setup, which need not be expensive.*
- ix. *That sharing of sophisticated equipment, e.g. for large items or 3D imaging, is unlikely to be successful, especially because user skills are usually critical*

*for success Support of requests for funding of such equipment should therefore take account of*

- a. The scale of work involved – small numbers of items should preferably be sub-contracted to well-equipped teams with a productivity focus.*
- b. Availability of surplus capacity on existing infrastructure*
- x. That user groups should be set up to provide good advice on software choices for new entrants.*

## **5: A National Portal to Heritage Collections**

*Of particular importance is the availability of portals. The Aluka portal, which is focused on African collections, is well-known and the University of Pretoria team also has experience of portal design and use. In the South African context a National Portal would make a unique contribution; in addition to providing, and managing, the conventional ‘one-stop’ access to a multiplicity of linked South African collections, with federated search of all of them, a South African portal would be a ‘nationally owned channel of access’ to this South African material, and mitigate some of the reservations felt by librarians and curators about access to ‘our’ material being via foreign-based portals. The portal would obviously provide managed access to the National Repository.*

*The Portal could also provide user-friendly access to information and support with respect to current Best Practice in digitisation and preservation. It would be logical for the Heritage Projects database that was compiled as part of the audit and will continue to expand, to be accessed via the Portal.*

*This is a rapidly developing field and there are lessons to be learnt from the leaders. Any shared servers or portals need to be carefully located for optimal performance and access, and especially with regard to updating and upgrading performance.*

### **Recommendations**

- xi. That an institution should be identified to host a National Portal and an advisory group established to direct development.*
- xii. That the advice of global leaders in the field should be sought.*

## **6. A National Digitisation and Preservation Support Centre**

*It is clear from the above that there are a number of developmental and support issues relating to Digitisation and Preservation of Heritage material that could well be dealt with on a common or coordinated basis in the South African heritage community. If these were to be dealt with ‘centrally’, it would be important to ensure an enlightened style that promotes support, rather than control, effective use of distributed skills and expertise, and a high degree of*

*transparency and good communication to promote progress and collaboration. Many commentators expressed concern that an energy-sapping bureaucratic 'dead hand' could be laid on the present initiatives. This is particularly acute where funding is concerned and there is strong opposition to any central control of the flow of funds from donor and other sources to institution-based projects.*

*Section 14 includes a possible staffing structure for the Centre, which focuses strongly on the use of skills in the system, rather than a large Centre staff.*

*The following Recommendations are made in this context.*

### **Recommendation**

- xiii. That a National Digitisation and Preservation Support Centre (NSDPC) should be funded, preferably close to or within one of the emerging centres of expertise, that would, inter alia,
  - a. Take responsibility for keeping in touch with global developments and ensuring that these are made known to and, where appropriate, useful to local role-players*
  - b. Support training at all levels in the system.*
  - c. Maintain an advisory network that would ensure effective sharing of Best Practice and expertise in the digitisation community, this to include organising workshops and seminars.*
  - d. Provide a forum for identifying skills and services that can be out-sourced**
- xiv. That the National Portal and/or the National Repository could be hosted in the Centre.*

### **7: Permissions to Digitise: Best Practice Guidelines for South African Practitioners**

*There is widespread nervousness over raiding of Intellectual Property, usually grossly exaggerated, especially where the real nature of copyright protection is not properly understood. 'Looting by cultural imperialists' is a much more substantial threat and widely publicised bad experiences have led to a situation where the professional consensus on this matter is one of fear. However, much if not all of the risk of such looting can be eliminated if collection custodians observe appropriate procedures in participating in digitisation partnerships, and a draft set of such procedures is presented in the report. (Section 16)*

*Concerns over the 'management of the historical narrative' are more intractable, at least in part because they are difficult to define in manageable terms. The task of producing a 'National Digitisation Policy' has been devolved to the National Department of Arts and Culture, but there is no sign of the early emergence of a product acceptable to all the stakeholders. In the meantime, progress is likely to suffer unless a few high profile champions can be persuaded to take a lead.*

### **Recommendation**

- xv. *That, pending the formulation of a National Policy on Digitisation, professional organisations should formulate their own best practice consensus for the guidance of members.*

---

## 1 Introduction

---

In January 2008, the Carnegie Corporation of New York through the office of the Programme Officer: Libraries and Information, Dr Rookaya Bawa, convened a meeting in South Africa with South African digitisation stakeholders. The meeting was partly prompted by the fact that the Corporation had recently received a number of proposals from different organizations and institutions in South Africa for funding to digitise South African resources as well as South African material at the United States Library of Congress.

The meeting recommended, *inter alia*,

- that Carnegie Corporation should support a project, led by the NRF, to establish a national digitization task team which would lead to the formation of a national digitisation body or institute.
- One of the urgently required tasks was an audit of existing projects that might take the form of a database.

This report describes the audit and the resulting database, and draws on this and a series of stakeholder interviews to make recommendations for further action.

---

## 2 The Brief

---

The brief for the investigation that gave rise to this report provided the following specific elements:

- identify whom and which institutions are involved in digitisation projects;
- identify potential projects for digitisation in South Africa;
- establish on the basis of frequency, priority areas/content for potential digitisation projects, and provide a preliminary list for discussion with stakeholders;
- provide a list of current/ongoing projects, to facilitate reduction of unnecessary duplication;
- reaffirm as appropriate the merits of a digitisation centre, its activities and potential governance structure;
- collect data and information for the establishment of a portal for digitisation projects that will ensure accessibility;
- identify levels of expertise/skills availability in digitisation, across institutions;
- establish digitisation needs, i.e. skills, expertise, knowledge, infrastructure requirements; and
- generate a national database (to be updated regularly) and hardcopy report of the audit for distribution to and access by key stakeholders. This would describe the
  - content of the collections
  - types of material
  - coverage of subject areas



- motivation for digitising each collection
- status of the levels of expertise in the holding organisation
- availability of software, hardware and systems
- connectivity at the institutions, and
- policies and procedures presently followed in relation to digital data.

---

### **3 Modus Operandi for the Audit**

---

The core process used was Internet-based, driven by e-mails to identified stakeholders, who were invited to complete an online questionnaire. The lists of addressees and respondents are given in Appendix D, and an MS Excel version of the questionnaire is given in Appendix E. (The latter was provided to respondents who expressed a preference for this format above the web-based questionnaire.) The submissions were recorded in a database forming part of the NRF's Nexus system, which has since been made available online at <http://stardata.nrf.ac.za/html/database.html>

The core set of addresses was made up from lists of key staff members at all the Higher Education Institutions (HEIs) and Science Councils, Public Sector Museums and the Public Libraries in the major centres, as well as a selection of NGOs from the Prodder listing ([www.prodder.org.za](http://www.prodder.org.za)). National Government Departments that were believed to have relevant databases were also included as well as obvious candidates such as the National Library and National Archives. This was followed up by personal contact with stakeholders known to the contractor, who typically identified further leads for follow up. Many interviews were conducted with key role players and an international Digital Heritage Repatriation Workshop at Rhodes University, which occurred fortuitously within the research period, was attended.

These 'audit' activities were augmented by desk-top research and focused site visits aimed at familiarisation with the technologies in use for digitisation in South Africa. From these inputs the first draft report was compiled and submitted to an NRF Review Panel, whose inputs were used to formulate a second draft that was used as a basis for discussion at the Stakeholders' Workshop on 18 March. Input from this workshop and feedback from Carnegie Corporation has, in turn, been used to produce this Report.

The time schedule followed is given in Table 3.1

As the questionnaire respondents were either self-motivated or 'encouraged' by the contractor, the responses were not expected to be, and indeed cannot be viewed as, statistically representative. They have accordingly not been quantitatively analysed to any great extent. However, the combination of responses received and interviews and other interactions with participants enabled a qualitative view of the South African digitisation scene to be obtained, and this is believed to be sufficient for some general conclusions to be drawn.

TABLE 3.1 Time Schedule for Audit	
Activity	Date or Time Period
Preparation in anticipation of Contract	September/October 2008
Signature of Memo of Agreement with Service Provider	21 October 2008.
Survey/ Research period	21 Oct 2008 – 9 Feb 2009
Submission of First Draft Report	9 February 2009
Feedback from Review Panel	13 February 2009
Submission of Second Draft Report	2 March 2009
Stakeholders' Workshop	18 March 2009
Submission of Post Workshop Report	3 April 2009
Report including Carnegie Feedback	31 August 2009

---

#### 4 Audit Findings – A Perspective on the Digitisation Status Quo

---

It was clear from the outset that the audit involved *two major elements*, viz

- **Data collection** – to establish the database, although it was appreciated that, in the short time available, the result would not be complete. In the longer term, the database would be incrementally developed.
- **Identification of key dynamics** in the system(s) and the extraction of meaningful recommendations from them that would promote effective and efficient expansion of the work currently under way.

As indicated in Section 3, above, the audit did not produce a statistically representative sample and had other shortcomings, which are expanded on below. However, the combination of data collection with many follow-up discussions with role-players has yielded significant insight that will hopefully assist decision-making. This perspective has the following chief characteristics:

There is *major concern (and significant misunderstanding) re 'copyright'* which is stalling progress in many sectors. This is extensively discussed in Section 8 - Audit Findings on Ownership, Copyright and Licensing.

There is *much activity, a great deal by part-timers*. The greater proportion of the collections have been or are being digitised by staff whose main functions lie elsewhere, typically in libraries. *Organisational operating budgets* provide the funds and *rudimentary equipment and infrastructure* serve many of these efforts. Many digitised collections are already *available on Open Access websites*.

*Some institutions*, having started small and part-time, are now *expanding* their teams and upgrading facilities to make greater impact.

*Some large-scale undertakings* have been tackled, e.g. the Freedom Struggles collection by DISA, the multiplicity of Rock Art collections consolidated by the Rock Art Research Institute at Wits and the ILAM collection of ethno-music recordings at Rhodes. All employ specialised teams trained up for the purpose and have developed

considerable skill and experience in their fields. These projects usually use donor funds, specially procured for the purpose.

There are *many untouched collections*, because custodians lack skills, staff or the funds to pay them, and/or are awaiting resolution of ‘copyright’ issues. This is likely to be the most under-reported category of collections in the database.

*Standards are reasonably well managed*, with the pioneering teams being sensitive to domain-specific requirements.

Awareness of the *need for long-term sustainability* is visible, but *few plans* are in position to ensure this.

There is a need for *capacity building* and *skills upgrading* which requires a combination of *training, advice from specialists and peer work-shopping*.

Overviews of the context for the audit and the main issues involved in Digitisation and Preservation are provided in Appendices A&B, Sections 19 and 20 of the report.

---

## **5 Audit Findings –The Database**

---

### **5.1 Identification of Likely Respondents**

The initial challenge was to identify likely respondents to the request for information within the relevant organisations. Various databases were used: from the NRF, LIASA, and the Prodder website ([www.prodder.org.za](http://www.prodder.org.za)). Where possible at least two contact names were used per organisation for the initial e-mailing, which comprised over 280 addressees. All science councils, higher education institutions, museums, provincial libraries and the public libraries in the larger cities were included, as well as various entities in the Department of Arts and Culture (National Library, National Archives) and about 90 NGOs that seemed likely to own relevant collections. Follow up activities were guided by redirection within organisations, personal contacts and referrals from respondents. At time of writing the report, 40 organisations had responded – some with multiple submissions. (See Tables 5.1 and 5.2, below.)

Many of the addressees had not responded at all and it is not known whether this was because the contact address was incorrect, or because they have no collections to report. The expected contributors in the Universities and Science Councils have responded and most of the relevant NGOs appear to have been identified.

Appendix D (Section 20 of the Report) gives the details, which are summarised in Table 5.1.

<b>TABLE 5.1: Tally of Submissions by Organisation Type</b>		
<b>Organisation Type</b>	<b>No of Organisations</b>	<b>No of Submissions</b>
<b>TOTAL</b>	<b>40</b>	<b>185</b>
Business/Commercial	3	4
Government	6	19
Museums	3	15
NGO's	2	5
Science Councils	12	30
Universities	14	112

## **5.2 Collection Coverage Achieved**

It was immediately apparent that, with the exception of museums and libraries, who specialise in the maintenance of collections, most organisations have no centralised picture of the collections in the organisation that are suitable candidates for digitisation. In many cases it seemed to be assumed that all digitisation needs would be visible from the institutional library. This was demonstrably not true; three reported 'research' collections at universities came directly and not via the library; there are probably many more. Although many of the requests had clearly been referred to university researchers in various disciplines, for example, there remains no way of telling to what extent full coverage of the SA Heritage scene has been achieved. .

The prospect of completing large numbers of forms towards the end of the final semester of the year was clearly not especially attractive to some participants and a great deal of encouragement and negotiation was required to extract returns. The NRF IT systems underwent a major modification during the survey and unhappy experiences with it discouraged some contributors, who had to be re-energised. Some institutional archives house rather too many collections for the custodians to wish to complete questionnaires exhaustively; the UNISA Archive<sup>3</sup> has 700 catalogued collections! For such organisations, with significant numbers of collections already digitised and more to come, it was thus necessary to negotiate a compromise that would reduce the effort required to manageable proportions. In some of these cases, all digitised material has been aggregated as a single collection and a second 'collection' was created by assessing the scope and scale of the 'to be digitised' category. Extensive details was given for only the current and/or completed activity although the Technical (Part 2) and the Resources (Part 3) Sections of the questionnaire were usually completed in detail.

## **5.3 Completeness of data submitted**

Ideally, the questionnaire should have been piloted with insightful stakeholders and any ambiguities and difficulties repaired. The time available for this did not permit a thorough job and the penalty was that some sections of the questionnaire were less well

<sup>3</sup> Interview with Ms Marie Coetzee, Archives & Special Collection, Unisa Library

completed than others, although this was not solely due to questionnaire shortcomings. The following is an overview:

***Part 1: Management and Governance.***

The data provided were always sufficiently complete to identify the institution and its key role-players and to enable follow-up contact.

***Part 2: Description of the collections digitised and to be digitised.***

*Descriptions of collection items* were usually adequate at a high level, e.g. nature of material, numbers of items etc, but only relatively few respondents provided insight into the content in the ‘Additional Information’ field. Furthermore, those collections yet to be digitised have often not been sufficiently well scrutinised for scope and scale to be accurately assessed. The overall number of collections is known to have been under-reported.

*Copyright details* were not always provided and very few organisations recorded a *digitisation policy*.

*Technical aspects* were better recorded in completed or current work. An interesting range of Metadata Schemas was revealed (See Section 10. Audit Findings – Systems and Standards).

The *collection evaluation* approaches, which were requested to provide information that would inform the draft prioritisation guidelines, were usually present but no sign of intensive internal prioritisation of collections, despite significant competition for resources, was evident.

***Part 3: Resources – Skills, Technology and Funding.***

*Staff numbers and expertise* rarely gave a complete picture, probably because most organisations are depending on part-time contributions from persons who have other primary duties.

*Equipment details* provided were usually confined to make and sometimes, model of equipment, but no specification details.

The need for *Quality Assurance* is clearly recognised, but quality production standards are mostly not formalised.

*Funding* details were rarely completed adequately, apparently because most organisations do not have a budget line item for this activity, and also because of the ad hoc allocation of resources. It is in the nature of the funding and operation of many of the organisations involved that budgets and cost-accounting are not atomised to the level of projects. Even organisations with digitisation grants are usually contributing management and accommodation ‘costs’ from normal operating budgets, and much of the valuable work that is being done has become part of ‘normal’ operations and is thus difficult to separate as a distinct item.

#### ***Part 4: Sustainability***

Very few respondents could identify a long-term plan.

#### ***Part 5: Lessons learnt***

Most respondents were reticent or silent on this topic, but some thoughtful insights were received.

### **5.4 The Database – Conclusions and Recommendations**

It is unrealistic to expect a questionnaire survey to develop a **complete** database of digitisable collections, especially where some organisations have clearly not identified their own in detail.<sup>4</sup> It was also clear that the necessary information, especially the more technical aspects, was not available in some institutions. Financial information may have been withheld for confidentiality reasons, but it is more likely that digitisation-specific costs are difficult to distinguish, except where capital equipment is concerned.

The coverage of active institutions exceeded expectations, but some important ones may still be missing. Once the existence of the database is known, it is likely that more organisations will come to the party and provision should be made for this.

There is a case to be made for continuing to build the database and make it visible, but revised protocols relating to the detail required will need to be negotiated with the stakeholders. A brief standard description along the following lines could be proposed: *Nature and number of items, free text descriptions of their value and the justification for digitising as well as contact details* for the collection champion will constitute a useful core. Some statement of the likely *benefits* to contributors, especially in relation to *project funding*, will allow custodians to market their collections.

---

## **6 Audit Findings – Collections and their Status**

---

One of the principal objectives of the audit was to compile a data base of collections digitized, in progress and needing/appropriate for digitising.

---

<sup>4</sup> If an organisation decides to undertake an internal audit, a well researched Data Audit Framework (DAF) methodology (Jones, 2008) has been developed for JISC.

Collection of this information was ruled to be complete, for the purposes of the Audit, on 27 February 2009, and this database is analysed in the accompanying Table.6.1. Note that in calculating the total numbers of items in collections the arithmetic midpoints of the ranges were used.

<b>TABLE 6.1: Collection Status and Number of Items</b>					
<b>Category</b>	<b>Digitised</b>	<b>Digitising under Way</b>	<b>Planned</b>	<b>Not Planned</b>	<b>TOTAL</b>
Number of Collections	18	55	64	48	<b>185</b>
Number of Items*	180 300	1 354 710	695 090	487 070	<b>2 717 170</b>

\* Calculated using mid-points of ranges: biased by large collections.

The institutions reporting completed or current digitising projects numbered 28 out of the total of 40, an indication of a significant spread of expertise

## **6.1 Who is digitising, where and why?**

Scrutiny of the database reveals a number of approaches to digitising and it is instructive to tease these out a little. The following categories are neither watertight nor exclusive, but they give some insight into the scope and dynamics of digitisation activity in South Africa.

### **6.1.1 Functional digitising**

Much of the digitisation activity is purely functional, aimed at improving the ease of large scale access to material already in regular and fairly widespread use. Examples include

- SABS availability of the SA National Standards to clients and users via the WWW.
- Various examples of teaching material, from microscope slides for pathology students to historical documents used by history lecturers.
- Specimen accession registers in museums that point to the availability of physical specimens.
- Preservation-oriented digitisation of material such as nitrate-based photographic negatives that are inherently unstable and mini-disc recordings with irregular formats that are becoming difficult to transcribe.

### **6.1.2 Digitising for academic visibility/reputation**

Electronic Theses and Dissertation Repositories lead the field here, although in many cases the bulk of the material is current and 'born digital'. This category is important because in many institutions it has provided the skills and infrastructure platform from which larger scale digitisation activities have developed. Other examples are

- The UCT Geomatics programme of laser-scanning of heritage structures in Africa. (not in this database)
- The Transvaal Museum's joint activity with the University of Toulouse to digitise the skull of Mrs Ples in order to reveal the brain structure.

- The Africa Online Journal collection at Sabinet.

### **6.1.3 *Assembly of (relatively) inaccessible heritage material***

Many collections of related material have been held in dispersed archives that are accessible in principle but only at significant cost of time and effort. The aggregation of these resources in digital form, accessible via the WWW, makes them readily accessible to a wider audience, including many potentially interested parties who were ignorant of their existence. Examples include

- The Bleek and Lloyd Archive of notebooks, photographs and artworks relating to the San peoples of Southern Africa.
- The Rock Art Collection from the SA Rock Art Digital Archive
- The Freedom Struggles collection being assembled via DISA
- The Hugh Tracey recordings of ethnic music in the International Library of African Music (ILAM) at Rhodes University.

### **6.1.4 *Digitisation for commercial purposes***

Items from various collections are made available from websites at a cost, or access is provided on subscription to a website, but in most cases these are 'break-even' or subsidised ventures with doubtful commercial viability. A possible exception is the African Image Pipeline created for Gijima KZN by Africa Media Online to promote sales of the work of local photographers to a wider market. (Not in this database)

## **6.2 *Scale and scope of digitisation initiatives.***

The spectrum of projects in terms of size and scope is large. Many projects are small (<100 items) and can be dealt with easily by the part-time activities of library staff, provided that they have the minimum of equipment and training. In many libraries this has been acquired by internal efforts and occasional requests to these teams for digitising of teaching material can be dealt with on an ad hoc basis.

Sometimes this activity extends to larger collections and the resources are scaled up incrementally to meet the demand, but in other cases the hurdle is too great and no work is undertaken. For relatively small expenditure, one could provide training to staff from the latter situations and empower them to 'start small' with a possible view to later expansion. The skills acquired would serve for ongoing ad hoc tasks, recognising that a time will come when all materials will be 'born digital' and the need for digitisation will fade.

At the other end of the scale are large projects, with over 100 000 items to be digitised, where it is worthwhile to set up a facility with trained staff and specialised equipment and to manage the project on industrial production lines, with production planning and targeting and formal quality assurance. Such projects may be most cost effectively dealt with by outsourcing, because the management requirements are alien to libraries and archives.



### 6.3 Size Distribution of Collections

There is a considerable variation in the, often only estimated, size of the collections, as can be seen in Table 6.2 which shows the numbers of collections in size categories, broken down according to status.

It is no surprise that ‘size unknown’ occurs so often in the description of collections *Not yet Planned* and the peak in the total number of collections in the 100 – 5000 categories reflects an intuitive sense of the size of document collections. The very high numbers of very large collections can be partly explained by the fact that, in such cases, mostly physical specimen *descriptions* are being digitised, and not 3D images of the specimens themselves, although the latter do occur.

Table 6.2 : Collection Size Distribution in Action Categories								
Digitising Status	Unknown	<100	100-1000	1k - 5k	5k-10k	10k - 50k	50k - 100k	>100k
Completed	0	0	6	5	4	2	1	0
Under way	0	3	10	11	7	12	12	1
Planned	3	2	29	16	3	4	7	0
Not yet planned	17	6	3	9	4	5	4	0
<b>Totals</b>	<b>20</b>	<b>11</b>	<b>48</b>	<b>41</b>	<b>18</b>	<b>23</b>	<b>24</b>	<b>1</b>

Discussions with various role-players suggest that the <100 category is the most under-reported.

### 6.4 Collections and their Status - Conclusions and Recommendations

The digitisation projects encountered in the audit covered a wide range in scope, scale, complexity and purpose, and appropriate design and management of the projects varies accordingly. The skills and equipment involved are dealt with in more detail in Sections 9, 10 and 11, and a method for prioritising projects where funds and other resources are limited, as is usually the case, is suggested in Section 7.

---

## 7 Audit Findings - Focus and Prioritisation

---

### 7.1 Prioritisation of Digitisation Projects

The audit reveals a daunting quantity of material, both explicitly and by implication, that is deemed by its custodians to be worthy of digitisation. In view of the constraints on resources available for this purpose, the question of priorities will inevitably arise, both within institutions, and within funding bodies.

In considering priorities one must always start with the reasons for digitising, viz Access and/or Preservation. Moderately important collections that are threatened with imminent decay should clearly have high priority and prioritisation of groups or categories of collections must weigh up these two elements.

In identifying the most urgent or important themes or categories of collections for digitisation one could, in principle, call for a national consensus on priority themes, but the prevailing culture of autonomy of institutions in relation to such activities militates vigorously against this. National consensus may, in any case be an elusive goal. The following approach is therefore proposed:-

- Identification of National Priorities from which themes may be derived
- Formulation of a process that will assist in prioritisation of collections, either within a Theme or within an institution, or both.

(The original plan, viz to identify themes on the basis of frequency turned out to be non-viable because the only available measures were the number of collections and number of items per collection, and neither of these categories was fully broken down into subject categories in the submissions. For the same reason, it was not possible to identify meaningful examples of duplication of effort. )

## **7.2 National Priorities as a Guide to Themes for Prioritisation.**

In the Natural and Engineering Sciences, a natural starting point would be the national priorities as defined in recent government policy documents. Of the two apparently most relevant the National Research and Development Strategy, published by the then Department of Arts, Culture Science and Technology in 2002, is the earliest. (DACST, 2002). This strategy recognised our national *'ability to exploit South Africa's "living laboratories" of local resources and geographic advantage'* Scientific collections that fit in with the relevant Science Missions are clearly part of an important scientific heritage. These Missions were defined as: astronomy, palaeontology, Antarctic and marine science, biosciences, social sciences, earth systems and environmental sciences.

More recently the DST published a 10-year Innovation plan (DST, 2008) that listed 5 Grand Challenges,

1. "Farmer to Pharma" value chain to strengthen the bio-economy – over the next decade South Africa must become a world leader in biotechnology and the pharmaceuticals, based on the nation's indigenous resources and expanding knowledge base.
2. Space science and technology – South Africa should become a key contributor to global space science and technology, with a National Space Agency, a growing satellite industry, and a range of innovations in space sciences, earth observation, communications, navigation and engineering.
3. Energy security – the race is on for safe, clean, affordable and reliable energy supply, and South Africa must meet its medium-term energy supply requirements while innovating for the long term in clean coal technologies, nuclear energy, renewable energy and the promise of the "hydrogen economy".
4. Global change science with a focus on climate change – South Africa's geographic position enables us to play a leading role in climate change science.

5. Human and social dynamics – as a leading voice among developing countries, South Africa should contribute to a greater global understanding of shifting social dynamics, and the role of science in stimulating growth and development.

Collections that are important in pursuing these future directions are clearly of national importance.

There is clearly some correspondence between the Science Missions and the 5 Grand Challenges.

The Department of Arts and Culture, in turn, has a published Strategic Plan for 2008 – 2011 (DAC, 2008) which refers, inter alia, to the policy review on the White Paper on Arts, Culture and Heritage which first saw the light of day in 1996.

Possible themes culled from these two documents include:

- Heritage Sites, possibly including Rock Art, although this is currently receiving attention
- Previously marginalised indigenous languages

The Review of the White Paper, expected soon, may provide more specific themes.

### **7.3 Stakeholder workshop suggestion for themes**

Participants at the Stakeholders' Workshop offered the following possibilities

- Project specific, e.g. Rivonia trial records
- Records at risk: e.g. National or United Party records
- Items on media at risk
  - Nitrate-based photographic negatives
  - Disintegrating documents of lasting value
- Broad categories: Natural history and biodiversity.
  - e.g. Plant specimens of medicinal or other potential (already in train via Aluka)

It was also suggested that the term 'categories' be used instead of 'themes' and the following broad categories were proposed:-

- Art and architecture
- Economic development and trade
- Cultural and political history
  - Private papers of persons distinguished for their insight into their times or their impact on those times
  - Major political parties of yester-year
  - Collections about or by significant individual figures in politics

- Collections by significant commentators
- Key historical events
  - Rivonia Trial
  - Jameson Raid
  - Kimberley Diamond Rush
  - Witwatersrand Gold Rush
  - Records of Antarctic Expeditions
- Archaeology.
  - e.g. 3-D images of key fossils that will create visibility for SA Palaeontologists and global research collaboration

A further consideration is the possibility of exchange of digital copies of certain collections for complementary collections of South African material, physical or digital, held by overseas institutions.

#### **7.4 A Tentative List of Digitisation Themes**

Based on the above, and with a view to incorporating major collections that have been taken up in the database, the following Themes are suggested. In due course it would be advisable to invite stakeholders to comment on these and to suggest alternatives. A panel of adjudicators representing various stakeholder organisations and domains could then make the final choice.

- a. *Department of Science and Technology National Priorities*
  - *“Farmer to Pharma” strengthening the bio-economy value chain*
  - *Space science and technology*
  - *Energy security*
  - *Global change science with a focus on climate change*
  - *Human and social dynamics*
- b. *Department of Arts and Culture Strategic Plan*
  - *Heritage Sites (Note that Rock Art is already being comprehensively dealt with)*
  - *Indigenous Languages*
- c. *From the Stakeholders’ Workshop*
  - *Project specific, e.g. Rivonia trial records*
  - *Records at risk: e.g. National or United Party records*
  - *Broad categories: Natural history and biodiversity.*
  - *Art and architecture*
  - *Economic development and trade*
  - *Cultural and political history*

## 7.5 Justification for Digitising

Once items have been identified as belonging to a priority Theme, resource constraints may still compel selection of higher priority items. Some selection process will thus be needed.

There was an expectation that the response to the ‘Justification for Digitisation’ section of the questionnaire would serve as the basis for creating a set of possible prioritisation criteria. In the event, the majority of respondents made one or more selections from the, possibly over-generous, list of prompts provided in the questionnaire, viz *uniqueness, condition, urgency, deterioration of specimens, importance for research or heritage*.

The most practical addition to this list was the utilitarian ‘*Protection of heavily used/consulted documents*’. Reflecting a different pragmatism was the criterion ‘*No copyright complications*’, which may work well in choosing where to start on a pile of equally important/urgent tasks, but is not a useful means of testing merit.

Reasoning logically, one comes up with the obvious list

1. **Perceived Value:** Clearly this is the most important and the most difficult to assess objectively, but positioning it within a priority theme should help.
2. **Fragility:** the items are likely to fall apart, become illegible, self-combust etc.
3. **Audience:** The items are of interest to far more people than can reasonably visit the archive in person; to which one might add the rider ‘*and are part of the legitimate target group for this service*’.
4. **Reciprocity/Collaboration:** Web publication of the images may give the ‘owners’ access to partnerships in research that would otherwise be difficult.
5. **Funds available or on offer:** Someone has the funds to train our people and/or pay for the work to be done by others. Provided that ownership and copyright are not compromised, this may be a mutually beneficial arrangement. (See Section 7 )
6. **Readiness for digitisation.** It makes a certain amount of sense to focus on this as a relatively minor criterion, but making collections ready for digitisation can absorb large portions of the budget.

Prioritisation in such a heterogeneous field will always be difficult, but if resources are constrained, some form of de facto prioritisation will inevitably take place; best to bite on the bullet and do the best one can. The following evaluation matrix is based on the Think Tools© methodology, which lends itself to such complex comparisons. (The Think Tools© software, which requires a licence, is available without cost to many role-players in the South African system.) Essentially the methodology attaches a weight to each criterion, e.g. Perceived Value – 10, Readiness for digitisation – 4, and then each collection under consideration is evaluated for relevance in terms of each criterion. The total score for any collection is the sum of the products of each pair of the weight and the collection score for that weight. The matrix below gives an example with a set of hypothetical weights and scores.

Other categories, e.g. enhanced access for hitherto inaccessible material, could be added. The list of criteria can be large, but the more there are, the weaker the influence of any single criterion.

**Matrix for Collection Prioritization**

Criterion	Weight	Collections and Scores			
		1	2	3	4
Perceived Value	10	10	6	8	3
Fragility	7	8	3	7	2
Audience	8	7	10	4	10
Reciprocity	5	5	7	10	6
Funding Available	7	8	6	3	7
Readiness for Digitisation	4	7	1	5	3
<b>TOTAL</b>		<b>321</b>	<b>242</b>	<b>252</b>	<b>215</b>

In this case, the most likely candidate is Collection 1, with Collection 4 clearly bringing up the rear and Collections 2&3 difficult to separate in terms of their ‘medium’ priority.

**7.6 Focus and Prioritisation – Conclusions and Recommendations**

The proposal in the brief that frequency should be used to identify priority themes proved unsatisfactory and an alternative, based on National Priorities, has been suggested for review by a panel of adjudicators

---

**8 Audit Findings – Ownership, Copyright and Licensing**

---

Many possible digitisation projects in South Africa are stalled or have been abandoned because of concerns re ‘cultural imperialism’. This arises mainly because many collections that have been and are being digitised in South Africa are destined to be made accessible via websites that are based overseas. There appear to be three levels of discomfort.

1. Access to online materials is easier for overseas users than South Africans
2. Archivists prefer to present full, contextualised packages of material
3. South Africans should manage the process, including the creation of the ‘historical narrative’.

At the lowest level is the concern that access to materials should be at least as easy for SA students/citizens as it is for their overseas counterparts. This is understandable, but refusing digitisation for only this reason has elements of cutting off one’s nose to spite one’s face. Many South Africans enjoy fairly good access to the WWW, while some have effectively none. Is it reasonable to deny web access to SA material to some as long as others don’t have it? Some disparity in user access is likely to persist for as long as the Digital Divide remains; it is surely unreasonable for this to become a total barrier indefinitely. The situation is more complex when one considers that, in terms of the governing Act, material digitised from the National Archive must, by law, be made as

readily available to South African citizens as from the Archive itself. The hypothetical web-disadvantaged student would probably find travelling to Pretoria to consult the Archives as much of a hurdle as lack of web access, perhaps more. A different manifestation of this discomfort is the expressed South African desire to ‘own the channel of distribution’ of SA material, but this can be readily satisfied by a National Portal. (See Section 13)

At a second level is the archivist’s professional desire to present full, contextualised packages of material by preparing sets of documents<sup>5</sup> that present a context for the relevant historical events. This should be possible, in principle, with a finite amount of effort – given sufficient access to the material. One can sympathise with the desire to restrict ‘global access’ while this is accomplished. (Note the irony that the advent of global connectivity plays a dual role here. Global connectivity is part of the disparaged ‘technology power’ of the Northern hemisphere, which enables unreflective global dissemination of whatever material has come readily to hand, Simultaneously, well not quite simultaneously, global connectivity furnishes the means to correct the imbalance by rapid follow-up with a more completely documented account.)

It is interesting to note that one of the US National Archives and Records Administration (NARA) Digitisation Objectives is:-

“Use resources effectively. For example, original records that have been digitized may be relocated to less expensive archival storage locations. Partnerships, where the partner provides resources for digitising, would expand the scale of digitising beyond what NARA itself can do.” (NARA 2008)

The ‘partnership’ notion has obvious advantages and should clearly not be neglected in South Africa either, especially if the partnership is perceived to be balanced between more or less equal local and ‘Northern hemisphere’ partners.

On the third level is the more philosophical question of how and by whom these resources should be managed. This is nowhere better summed up than by Michele Pickover (2008): *‘In the South African context the digitisation of heritage material for publication via the World Wide Web is a site of struggle and the real challenges are not technological or technical but social and political. Digitising archives is more than merely collecting and aggregating documents in cyberspace. What is at stake is the politics of memory in digital form and how what is selected for digitisation projects frames research agendas and plays a role in curriculum strategies.’*

This level of the debate is not readily accessible, in all its nuances, to those not immersed in modern, (or is it post-modern?) theories of history<sup>6</sup>. (Lalu, 2007) Here lies the ownership, and this seems to imply some level of control, of the historical ‘narrative’. Colonial historians are condemned for spending short periods of time in a colony and then presenting their superficial, or misconceived, narrative as authentic, to remain unchallenged by the locals, perhaps indefinitely. There would always have been some resentment of this clumsy paradigm transfer, from the locals’ frame of reference to the alien frame of the colonialist, but modern scholarship is found in the ‘colony’ as well as

---

<sup>5</sup> ‘Documents’ in this context would include published and unpublished historical papers, speeches, videos etc.

<sup>6</sup> This category includes the author, who presents this as a well-intentioned, but possibly rash, endeavour to open up the debate. He is grateful to those archivists and historians who bore patiently with his probing and enabled this perspective to be formed, but they are obviously not accountable for the product.

in the North. The apparent eagerness of the 21<sup>st</sup> C ‘North’ to pre-empt the local, and presumably more authentic, narrative, simply because there is the technology power to do so, attracts the pejorative ‘cultural imperialism’. To those less directly involved, the possibility that this tension might be resolved in a manageable time scale seems remote, not least because of the conceptual difficulty of achieving consensus that the authentic (local) narrative has actually been created. The debate to date appears more focused on characterising the problem than on resolving it.

Simply stated, this implies in the first place, that one should avoid selective digitisation of freely accessible material at the expense of the more complete picture that could emerge if more contextualising effort is expended. It will be more difficult to discern where digitisation should legitimately be held back pending formulation of a widely acceptable consensus on the authenticity of the narrative. It should be possible, however, to identify collections where this can be judged, conservatively, to be an insignificant or acceptable risk and to release them for digitisation. Such decisions will probably need the support of a band of respected stakeholders, inside or outside the organisation.

### **8.1 Bad Experiences**

The situation has been exacerbated by horror anecdotes<sup>7</sup>, of which the following are instructive:

#### **The Nelson Mandela Rivonia Trial speech**

Towards the conclusion of his trial for treason in 1964, Nelson Mandela made a statement from the dock prior to his sentencing that is likely to be quoted by his legions of admirers for a considerable time to come. It has been available in print since that date, but it was also recorded on a magnetic tape of a type no longer used in South Africa and for which no playback machine is available locally. By collegial arrangement with an overseas institution, the recording was transcribed onto CD and returned, with the tape, to SA. Unfortunately it also soon appeared on the website of the transcribing organisation as if they owned the copyright, which strictly in terms of their national governing legislation, they did. This understandably aroused the ire of South African archivists and librarians, and news of the misdemeanour spread like wildfire in their community. The copyright claim was apparently not malicious and has since, on representation, been relinquished. However, this latter fact is much less widely known in SA and the ‘cultural imperialism’ bogey engendered by the incident is alive and well.

#### **The ALUKA case<sup>8</sup>**

Many digitisation projects in South Africa have been funded by the US-based Mellon Institute, a major funder of educational initiatives in South Africa. Much of the material is published on a website – [www.aluka.org](http://www.aluka.org) – which requires subscription access, although this is presently provided gratis to users in research and education institutions in Africa. Copies of all digitised material are also held in SA by DISA,

---

<sup>7</sup> The Afrikaans ‘Skrikstories’ is actually a more graphic term.

<sup>8</sup> Some stakeholders may contest the details of this account, which has been derived by the investigator from conversations with several of the parties involved, including Aluka staff. Be that as it may, the conclusion that damaging assumptions and perceptions are at the root of the problem appears valid and these need to be addressed if progress is to be made.



the digitising body, and Aluka require only a non-exclusive licence to use the material on their website. This does not mitigate a widespread belief that '*Aluka have copyright to our heritage material.*'

In principle, the arrangement suits the Mellon Institute, which includes among its objectives the provision of South African material for educational use in the USA, as well as in African countries, and also suits South African researchers because it satisfies immediate access requirements and is likely to be a more sustainable distributor – and preserver – than any alternative presently available in SA. South African participants were assisted with training and other advice as part of the 'capacity building' activity in DISA. To many, this would appear to be a mutually beneficial arrangement, with a healthy partnership operating between DISA and Aluka.

However, it is also sometimes claimed that material selection, especially as it relates to the Freedom Struggles Archive, is driven by a United States 'agenda', although the selection committee is populated predominantly by South African specialists. As the project has proceeded, the production pressures arising from a need to 'feed the hungry website' have nevertheless created the impression that the partnership has become unbalanced in the Aluka direction, with too great a role being played by funding power. Critics might respond that the local selection committee has been insufficiently resolute in managing the pace of developments, but this is easier to allege in retrospect than it would have been to manage at the time. The recent transfer of Aluka into JSTOR, another Mellon-funded initiative with similar access provisions, has increased the nervousness of the sceptics.

Sceptics have also observed that the 'digital divide' discriminates *de facto* against South African users, especially schoolchildren, but as this is largely a function of the SA government's egregious delay in the bandwidth roll-out in SA; the blame for it cannot be laid at Aluka's door.

Clearly not all of the problems arise from Aluka attitudes, although they tend to get the blame. With the wisdom of hindsight one might call for measures to redress the partnership balance, including

- A greater funding allocation for material assembly. (See 20.7.1.)
- More transparent interactions between Aluka and the local committee.
- More user-friendly arrangements for individuals in South Africa to gain full and free access to the site, i.e. not to have to work through an approved organisation.

.Hopefully this relationship can be repaired to the benefit of all involved.

### **A Successful Collaboration with Elements of 'Historical Heritage'**

It is noteworthy that South African contributors to the African Plants Initiative and the Heritage Sites Initiative report in positive terms about the spin-off benefits of their

interactions with Aluka. An example is the Rock Art Research Institute (RARI) at Wits which is leading an effort to digitise all rock art records in the various holdings around the country. RARI has succeeded in involving most of the custodians, establishing agreed metadata and other standards, satisfying various Khoi/San representative groups of their bona fides and then digitising >100 000 records at a rate of about 5000 per month. There is process learning available in this case study that is a potential asset to the SA community. So too, is this team's experience of consolidating the work on numerous collections of similar material via a single production line.

### **Real problem cases**

Unrelated to the above, there have certainly been examples of thoroughly unprincipled raiding of SA heritage material by overseas publishers, which resulted from naïve local parties signing highly prejudicial agreements under pressure from more sophisticated publishers.

The upshot, a significantly paralysing nervousness on the part of South African archivists about overseas funded digitisation initiatives, is not serving anyone's interests.

## **8.2 A National Policy on Digitisation (especially for Heritage Material)**

Overseas funding sources have supported significant volumes of digitisation in the recent past. The database created as part of this audit reflects at least 300 000 items so digitised. In the order of 15 000 were rescued from permanent loss due to medium destruction. This has given SA sources, and in many cases also SA research achievements, considerable **global visibility** – a precious asset to local workers in the current era. That some of the digitisation has supported overseas agendas is no great loss provided that there have been concomitant local benefits. In some cases, apparent imbalances have arisen because there has been no **formal SA agenda** to position in the negotiation. However, to expect overseas funding to adapt exclusively and unreservedly to SA agendas would appear to be both unrealistic and unnecessary.

Nevertheless there is a widespread feeling, which was also expressed at the Carnegie Corporation's Workshops in January 2008, that some form of authoritative policy guidance is advisable. A March 2007 National Heritage Council *Workshop on the Protection and Promotion of Digital Heritage Resource* (NHC 2007) appointed a Task Team to '*participate in the process of legislative and policy formulation on digitisation of the heritage sector*'. The Task Team has apparently not yet met, and the matter has been handed over to the National Department of Arts and Culture, which set out to appoint a consultant to draft a 'digitisation policy' that will support collection custodians in decision making around the permissibility of digitisation and web-publication. The Department invited tenders for the drafting of a National Policy on Digitisation in 2008, but internal procedural irregularities led to the tender being cancelled and it has been reissued; the closing date for tenders was early March 2009, and the identification of a contractor is awaited.

### 8.3 The Digital Divide

It reeks of exploitation when SA material is digitised and then becomes more readily available to overseas researchers or the general public than to their SA counterparts because of subscription costs or bandwidth restriction. Most digitised material originating in South Africa is currently available free of subscription costs to African users and, for researchers in South Africa, the **digital divide** is narrowing. The imminent landing of the Seacom West Coast cable and the deployment of the SA National Research Network (SANReN) in 2009 will provide much more affordable high bandwidth connectivity between educational and research institutions, both locally and globally. Admittedly, rural schoolchildren in SA are unlikely ever to receive the same online services as have become common in developed countries, but does this justify antipathy to overseas funders provided that their contribution will ensure local access – albeit not comprehensively – and that their ‘agendas’ coincide, or significantly overlap, with the local one(s)?

### 8.4 Possession, Ownership and Rights

Some of the tensions around digitisation arise from different perspectives toward Possession in a physical sense, as distinct from Ownership. Collections are held in protected environments for a multitude of reasons. Often access is restricted because custodians lack the time or other resources to promote the availability of the material and then to ensure monitored access. In other cases it is alleged that ‘owners’ who may be individuals or less readily identifiable entities such as political ‘collectives’ are chary of exposing all the material to possible critical surveillance. Although possession may be 9 points of the law, moral or political pressures on archivists to respect the views of various stakeholders can be extreme, and owner gate keeping is a real factor, even where this is in conflict with the archivist’s professional principles that favour “Open Access to everything”

Once an item has been digitised, the resultant version can be replicated at very low cost and Possession, at least in the sense of the content or appearance of the item, is no longer unique. Fear of unrestrained duplication of one’s material is real, if possibly exaggerated. On the other hand, from a preservation point of view, the LOCKSS<sup>9</sup> approach has much to recommend it.

Furthermore, collections that remain sequestered in South African paper archives provide no value to **anyone** bar the few who can afford to travel to the source and overcome the resistance of collection guardians. Nobody wins here.

---

<sup>9</sup> LOCKSS - Lots Of Copies Keeps Stuff Safe - is a project driven from Stanford University to promote open access by multiple custodianship of items.,

Even if the origins of material can be identified, it may be impossible to identify the current 'owners' and there have been a number of attempts to create ethically responsible,

**SAHA Copyright Notice.**

Copyright Notice: All materials on this site are protected by South African copyright law and may not be reproduced, distributed, transmitted, displayed, or otherwise published in any format, without the prior written permission of the copyright owner. Please note: Every attempt has been made to obtain copyright clearance for the material on this website. However, we have not been able to trace some rightsowners. We have included their material in this digital collection, but invite anyone who can assist us with contact details for the rightsowners, to contact us at pickover.m@library.wits.ac.za so that we can obtain formal permission. Thank you.

and affordable, means of overcoming this. One such was to advertise the known list of (400) correspondents with a former Governor of New Zealand, whose papers were to be digitised, and to invite objections from their descendants. The lack of response to this was treated as *de facto* permission to digitise and

present the material in Open Access. (Cathro, 2007) How useful this specific approach would be in a society with low literacy and much lower access to newspaper advertisements is questionable, but a similarly lateral thinking approach may be useful in developing countries. Indeed the South African History Archive include a similarly directed Copyright Notice with their collections. (See box.)

Of course, heritage material can have special symbolic value that renders the custodians reluctant to allow removal of the material from their premises, or even handling by other than their own staff.

### 8.5 Realism, Royalties and Repatriation.

In technology development circles it is common to remind researchers and technologists that the prospect of any good technology idea generating serious wealth is rated at about 1 in 3000, the point being that ideas have to be shared and developed, at considerable cost, if they are to be commercially successful and that success is never guaranteed. The same is apparently true of other IP – the example of Solomon Linda's 'Mbube' that became 'The Lion Sleeps Tonight' and eventually paid significant royalties to his descendants over 70 years later, is the exception that proves the rule. In fact, most heritage items can yield only trivial amounts, which do not warrant the expense of identifying the beneficiary - the Iris Mjekula example is salutary. (See box) This does not prevent ethnic music composers, performers of music heard at grandmother's knee and especially their well-meaning patrons, from believing that every snippet of a traditional music recording is capable of generating loads of royalties. Repatriation of the music *per se* is a different challenge. Ethno-musicologists feel strongly that music recorded by community members 'belongs to the community', but giving effect to this 'ownership' can be difficult. It certainly helps if one focuses on a **source community** that is more readily identifiable than the heirs of the singer or player. If the music is still played or sung in the community there is obviously nothing to 'repatriate'. If not, one needs to identify a valid, and preferably cost-effective, way of making it available to community members.

**The Iris Mjekula example.**

Boudina McConnachie, a Master's student in Music at Rhodes University, has focused on copyright issues relating to ethnic music. Recordings made many decades ago by Hugh Tracy have since been licensed for sale by ILAM at Rhodes, via a Smithsonian Institute website, on condition that a proportion of the proceeds should be shared with the 'original artists'. As a case study, Ms McConnachie identified a track recorded in 1957 by one Iris Mjekula at a specific location close to Grahamstown and set about locating the singer or her descendants one weekend – she is fluent in Xhosa. Two working days and R656 later she had not succeeded in finding any person or persons to whom the amount due, viz R65.78, could legitimately be paid over, and this would have been one of the larger 'royalties' (McConnachie, 2009)

Protecting contributors to heritage and other collections from exploitation obviously remains an objective, but a dose of realism is surely required. Experience shows that, in any publication of recordings, one should avoid making unachievable commitments to repatriate music or to compensate beneficiaries, especially several generations down the line. The wisdom of hindsight also enables one to adopt better arrangements for recording the origins of a piece so that it can be attributed to an owner if significant income streams subsequently accrue.

### **8.6 Bridging a gap with a moratorium**

Requests for permission to digitise for access purposes may be withheld for valid, but essentially short-term, reasons; the custodians or owners may wish to allow only certain scholars access to the material for a given period, or the archivists may wish to augment the collection with newly-found material, which could be time-consuming. On the other hand, it may be urgent to digitise material whose physical condition is deteriorating, but where immediate full publication is undesirable.

In the publication of research data in pursuance of the principle of access to publicly funded research, it has become conventional to allow the original researchers a period of unique access to the data, so that they may extract reasonable publication value from it before it becomes available in open access for opportunistic exploitation by rival researchers. The period of such moratoria is very domain-bound, with a few weeks being common for genome studies versus a few years in the case sociological research projects where a time series of surveys is expected to yield special insights.

If the real issues are confronted and discussed with an, apparently, over-hasty funder, it should be possible to agree on an appropriate moratorium, which should be captured in contracts and made known to all the parties involved

### **8.7 Ownership, Copyright and Licensing – Conclusions and Recommendations**

The expectation of large incomes to be derived from digitisation of heritage and research material is largely mythical. There are projects where material is selected for its commercial value, e.g. Africa Media Online's *African Image Pipeline* project, but this is

an overt expectation and contracts are drawn up accordingly. Incidental income is so unlikely to accrue from the publication of minor excerpts that it should not be a major issue in licensing material for digitisation. Appropriate protection can always be included in contracts, but revenue expectations among the library community in general need to be rational.

The issue of a National Digitisation Policy is unlikely to be resolved soon and there is a need for guidance that will reassure custodians of collections that are targeted for digitisation, especially where donor funding is involved, that they have taken adequate steps to protect the rights of the relevant parties and not laid themselves, or the 'owners', open to exploitation. The bad experiences of the recent past have engendered a professional consensus best described as 'fearful wariness'. It should be possible for professional groups to identify and agree on a set of interim guidelines that would reassure local custodians that their acceptance of donor funding

- Complies with current best practice
- Would be defensible to and supported by local authorities

This rationale has informed Section 16: Recommendation: Permissions to Digitise: Best Practice Guidelines for South African Practitioners. These Guidelines were drafted by the author and augmented after inputs from the participants in the Stakeholders' Workshop on 18 March 2009.

---

## **9 Audit Findings – Staff**

---

The task of preparing a collection for digitising, including sourcing the items from multiple custodians or owners, obtaining the necessary permissions to digitise and publish them and preparing a digitisation protocol with appropriate processes and metadata, is demanding professional work that can involve a wide range of domain and digitisation specialists. (See Section 20. Appendix B: Digitisation and Preservation Overview). By contrast, the labour involved in repetitive capture of digital images and their presentation, while very demanding in terms of careful and systematic operation and often requiring the greater bulk of the man-hours devoted to the work, can often be undertaken by lesser skilled, but suitably trained, individuals. In start-up digitisation programmes the whole process chain is often undertaken by whoever is available and willing, usually librarians or museum curators, and this may remain the practice where digitisation is a minor activity in an institution. Where very large collections must be dealt with, however, a more production-oriented approach is required. Both approaches were encountered in the survey.

While only 29 organisations reported on staff, one can deduce that nearly 180 people are employed, full-time or part-time, in digitisation activities, with a wide spread of functions.

(Details in Section 24, Appendix F, summarised in Table 9.1) All the expertise categories requested were well represented, and there were significant numbers of additional specialists.

<b>Table 9.1 Summary of Staff Expertise</b>	
<b>Expertise Type</b>	<b>Number - Full-time and Part-time</b>
IT Applications	39
Metadata	51
Conservators	22
Equipment Operators	43
Other	24
<b>Total</b>	<b>179</b>

There are several viable, specialised groups, including

- Sound at ILAM
- Rock Art at Wits RARI
- Routine ongoing digitisation at National Library in Cape Town
- General items – mainly documents - at University of Pretoria, DISA, Wits University.

Much of the digitisation-related current work is done on a part-time basis by staff with other main responsibilities, often given an initial impetus by major project funding, such as the Mellon Foundation’s support of DISA 2. However the ‘line’ of dedicated staff is perilously thin. Several organisations (e.g. UCT Library) are moving towards having formal teams dedicated to digitisation and as more funding for projects becomes available the need for trained staff will increase. Any fully-fledged librarian or archivist of the future is likely to need familiarity with, or at least have access to, best practice in digitisation and the capacity building demand will be high for some time to come.

In this connection it is important to ensure that University Departments teaching the next generation of librarians ensure that they at least have an option to master the elements of digitisation and preservation.

Specialised digitisation work is also undertaken by researchers, but usually on a selective basis driven by a particular research need rather than as a comprehensive effort. Some examples are given in Section 11: Audit Findings – Technology and Physical Infrastructure.

### **9.1 Spectrum of skills needed**

The following is an attempt to illustrate the range of skills required in digitisation, and especially the fact that they can reside in persons of different levels of education and seniority. The graphic is focused around the Collection Champion, being the person who has overall management of the digitisation project or group, although some processes may be supervised by a subordinate and some inter-organisational negotiations may be carried out by a superior.

Table 9.2 Skills Spectrum and Role Players in Digitisation					
Function	Role Players				
	Senior Manager	Collection Champion	Colleague/outsourced	Data/Image Manager	Operator
Collection Assembly			Domain Specialist		
Permissions					
Standards Selection			Domain Specialist		
Equipment Selection			IT Specialist		
Metadata			Cataloguers		
Material Prep Protocol			Conservator		
Image Capture					

Table 9.2 (ctd) Skills Spectrum and Role Players in Digitisation					
Function	Role Players				
	Senior Manager	Collection Champion	Colleague/outsourced	Data/Image Manager	Operator
Image Management					
Project Management					
Marketing and Commercialisation					

It is easy to appreciate from the diagram that the Collection Champion needs a spectrum of skills that is unlikely to be found in a conventional librarian or museum curator. The effective current performers have expanded their skills by application and/or specialised training. It also reflects how a single individual can deal with most of the demands where collections are small and the situation is not complex. As size and complexity increases, the demands on the Collection Champion will also increase and s/he may need further training, collegial support and/or management support.

(It is understood that additional funds for skills upgrading have been made available by Carnegie Corporation to a group of SA university libraries, but how they map onto this picture was not revealed in any of the responses.)



## 9.2 Staff -Conclusions and Recommendations

There is a need for ongoing capacity building activities in the research and heritage communities to create a broader base of skills and awareness. Isolated practitioners are easily daunted by the unfamiliar. User friendly advice and training is required, especially focused on persons who have an active need to undertake digitisation on some scale, i.e. rather than as an academic exercise. They can then go on to more specialised and large-scale initiatives if these are warranted.

Technology needs and availability, both software and hardware, are likely to improve rapidly in the short term, and a watching and advising brief should be created for suitably positioned players in the system. This can be an arid assignment as a full-time commitment, but very fruitful if it is given to an active practitioner, or group of practitioners, who can explore and examine new opportunities and make informed and regularly updated recommendations to the digitisation community. The same approach should apply to discussions and developments with respect to standardisation and systems (See Section 10: Audit Findings – Systems and Standards). Part-time (Virtual) associates of the National Digitisation and Preservation Support Centre (See Section 14) could be an effective resource in both these connections.

Material developed for these purposes could also be made available to Universities for teaching purposes.

---

## 10 Audit Findings – Systems and Standards

---

‘Lack of standardisation’ is a regular cry in the digitisation field in SA, but it is necessary to unpack the term to appreciate the implications. Standardisation is important for the sake of

- Discovery of the digitised objects on the Web
- Reusability of objects, to minimise repeat work
- Interoperability of systems

### 10.1 Discovery of the Digitised Images on the WWW Metadata

The effort of digitising a collection to make it accessible via the WWW is in vain if the items cannot subsequently be found. If the digital images or information are to be ‘discovered’ by search engines within collections or harvested by federated discovery services operating up to and including the entire Internet (e.g. OAIster, Google) they need to be appropriately described. Various systems for doing this and promoting standardisation and uniformity are available. The data used in the description are called *metadata*: ‘data about data’.

The dominant harvester in archival practice is OAIster, a union catalogue of digital resources that provides access to digital resources by "harvesting" their descriptive metadata using OAI-PMH (the Open Archives Initiative Protocol for Metadata

Harvesting). (OAIster, 2009) It is important to note that OAIster needs to be instructed to harvest a database, it does not 'trawl' the Web as Google does.

For Resource Description, OAIster promotes the use of the Dublin Core, a well-known 15 element standard set, which is widely used in South African collections. An optional additional 3 elements convert this to Dublin Core - Qualified. Other standard sets of resource descriptors are VRA Core, designed for the cultural heritage community and a set widely used for describing biological specimens named Darwin Core, which has variations designed for Ocean Biogeographic Information Systems (OBIS) and PaleoPortal, used in palaeontology. (Darwin Core, 2008) The Darwin Core set enables recording of the geographical location of an item. All these systems are encoded using XML (eXtensible Mark-up Language)

The audit revealed widespread adherence to Dublin Core (63% of respondents) or Dublin Core Qualified (18%) with XML Encoding Language for individual items. The museum/biodiversity fraternity favour Darwin Core (5%) and the HSRC are using the Hasset Thesaurus and Data Document Initiative (DDI) - an international standard for describing social science data. Two groups use customized systems of their own devising whose accessibility to harvesting via OAIster, for instance, is not known

### **10.2 Reusability of images, to minimise repeat work**

Reusability relates more to format and .pdf has been widely used in South Africa. Most users can read documents in pdf and it is a reasonably well compressed format. However not all systems can search pdf's so the full text can remain opaque and further management of the image is difficult if not impossible. For this reason, many contributors reported that they created images in .tiff and archived them thus, but distributed pdfs.

### **10.3 Interoperability**

This enables items in different image management packages to be shared easily. In SA there are users of Greenstone, DSpace and Fedora, with no current evidence of standardisation. However all these systems use xml encoding which will support harvestability.

### **10.4 Overall**

There is no evidence of a national standard, but where standardised methods are sought, there is a strong tendency to adopt global good practice for the relevant domain. The DISA practices, well described in an online manual, provide a good example of standardisation of contributions from multiple sources to a common collection or aggregation of collections. [http://www.disa.ukzn.ac.za/index.php?option=com\\_docman&task=cat\\_view&gid=62&Itemid=88](http://www.disa.ukzn.ac.za/index.php?option=com_docman&task=cat_view&gid=62&Itemid=88). Indeed much of the common choice of standards reflected in the above appears to have been led by participation in the DISA initiatives.

The topic of standardisation of systems continues to attract attention at international level. In November 2008, the Europe-based consortium DRIVER (Digital Repository

Infrastructure Vision for European Research) published its second set of *Guidelines for Content Providers - Exposing textual resources with OAI-PMH*. (Driver, 2008) This appears to be an excellent and experience-matured guideline for SA to follow.

### **10.5 Systems and Standards – Conclusions and Recommendations**

Future work in South Africa could well piggy-back on the DISA example. A sound approach might be to identify user groups, and the Audit Database should be useful here, to explore options for various purposes. The user groups could be convened by the National Digitisation and Preservation Support Centre (NDPSC), which should also lead the ongoing scrutiny of overseas developments.

---

## **11 Audit Findings – Technology and Physical Infrastructure**

---

The basic requirements for digitisation of simple, unbound, ≤A4, flat documents are modest: a PC, scanner and open source software for simple content management will get you going. For larger documents, or irregular shapes, or ‘non-flat’ bound items, the complexity escalates, with a camera with some depth of field becoming necessary.<sup>10</sup> Special software is needed for Optical Character Recognition (OCR) to enable word search.

This section was not very well completed. Even those institutions who reported on their equipment failed to provide any information on specifications and reasons for choices. Of those who did report, many seemed to have the basics, and a good few have more sophisticated cameras. Robotic cameras and Cruse Scanners are coveted by numbers of users, but the most sophisticated unit reported was the Digibook Suprascan at the University of Pretoria. There are a number of systems for digital imaging from microfiche and spooled film.

There is considerable interest, especially among the museum researchers, in 3-dimensional imaging of fossils and artefacts. Dr Francis Thackeray, formerly of the Transvaal Museum and now at the Institute of Human Evolution at Wits, sent a CAT scan of Mrs Ples’ skull, obtained at a local hospital, to a French correspondent at the University of Toulouse, who constructed a 3D model of Mrs Ples’ brain from it. (Thackeray, 2002). IZIKO researchers have used digital microscopy solutions that enable multiple view microscope images, that typically have very low depth of field, to be ‘stitched’ together to create a 3D effect. This is a handy technique for special purposes but probably too labour intensive (~1hr per specimen) for mass digitisation. Holographic techniques are another possibility.

No organisation reported an ability to make full 3D images of their collections, although Heinz Ruther of the Dept of Geomatics at UCT has a laser scanning facility that has been used to make accurate 3D images of heritage structures and sites in Africa (available on Aluka) and this could, in principle also be used for much smaller objects, e.g. fossils or

---

<sup>10</sup> This is particularly important for precious bound volumes where the tendency to bear down on the volume, placed face down on the platen, threatens fragile bindings.

plant specimens. The Advanced Manufacturing Technology fraternity in the universities and science councils routinely create 3D images of objects to enable copy machining.

Only two organisations, RARI (Rock Art) and DISA, reported a ‘digitisation production-line’; in most cases routine work seems to be done by low-skilled employees or students, with library and museum professionals ‘getting their hands in’ on a trial basis. This has doubtless given rise to considerable capacity building, but the task of routine scanning and processing of many thousands of images seems more fitted to a factory type production environment than the professional environs of a library or museum. This is one reason why outsourcing of certain production runs is an attractive proposition.

Digitisation of sound material is well established at several institutions, especially Rhodes University (ILAM), the Mayibuye Institute and the SABC.

### **11.1 A National Portal**

Of particular importance is the availability of portals. The Aluka portal, which is focused on African collections, has been mentioned and the University of Pretoria team also has experience of portal design and use. In the South African context a National Portal would make a unique contribution; in addition to providing the conventional ‘one-stop’ access to a multiplicity of linked South African collections, with federated search of all of them, a South African portal would be a ‘nationally owned channel of access’ to this South African material, and mitigate some of the reservations felt by librarians and curators about access to ‘our’ material being via foreign-based portals. Of the 185 collections listed in the audit, only 43, representing 29 organisations, provided web addresses for their collections. This suggests that there would be a base to start from and considerable opportunity to host many more of the collections. This latter feature would also enable custodians in institutions that are difficult to access via the internet, to position a mirror site on the server that carries the National Portal.

Any shared servers or portals need to be carefully located for optimal performance and access.

### **11.2 Sharing of Specialised and Sophisticated Equipment**

Highly specialised equipment, e.g. for 3D imaging or high-production scanning can, in principle, be shared, although efforts to create such sharing arrangements have a poor history in the Science Sector. One reason is that sophisticated equipment needs to be under the control of well-trained and sensitive operators, who are often the critical resource. In practical digitisation of large collections the equipment has little slack time and then usually because of staff shortages. Arrangements could presumably be made for short runs to be undertaken by the long-term ‘owners’ of such equipment, against some form of compensation for the loss of productivity to the owning team. This will require sensitive negotiation and management, probably lubricated with special funding.

Where a small collection (say <1000 items) needs to be digitised and this is the sole likely digitisation project in an institution, it may be more sensible to have the work done by an established team rather than to train staff for a short-term exercise. Where

equipment that requires relatively little training for effective use is known to be under-utilised it can be made available to other users, and indeed this is happening in the system at the moment. However, the benefits of maintaining records and databases to manage such sharing processes are unlikely to warrant the effort required.

### **11.3 Back-up**

Several organisations make provision for off-site back-up copies – ILAM material is backed up in the heart of a mountain in Norway! It is more common for back-up to be sought within the host organisation and often in vain. Long term storage capacity (>10 years) is a theoretical concern for most at this stage. The threat of lack of capacity to store the millions of images that could be created is discounted by IT futurists who believe that improving packaging and compressing technology will keep up with the ongoing explosion of digitised material.

### **11.4 Software Selection and Innovation Management**

A variety of Open Source and proprietary software is in use. A particular problem experienced by new entrants into digitisation activity is the selection of appropriate software, because once a significant amount of scanning and preparation of images has taken place, teams feel ‘locked in’ by this investment of time and effort and are reluctant to change. Early advice from experienced individuals (or teams, to avoid purely personal preferences) would be valuable. There will always be a tension between the tried and trusted and the latest innovation and deciding to migrate from outdated software can be a headache. Again a team approach is valuable and South African operators can take advantage of the intensive development work in the UK and Australia, where shared initiatives, funded by government, are the order of the day. Such networking could be coordinated by the National Digitisation and Preservation Support Centre.

### **11.5 Technology and Physical Infrastructure – Conclusions and Recommendations**

Basic digitisation equipment for flat specimens  $\leq$  A4 in size, is widely available, but specialised or high production equipment is rare and then tends to be fully occupied. Many teams have large scanners (A3 or larger).

Sharing of sophisticated equipment is taking place on a limited scale, but is unlikely to be a good long-term solution to resource optimisation. It will be better to subcontract small projects requiring specialised equipment to the present users than to share equipment among a group of users.

There is interest in 3D techniques, but only one large scale facility, which is presently used for site-scanning although it could, in principle, be used for smaller objects. Facilities for object copying in the automotive industry are also potentially available.

There is enough skill and experience in the system to set up a South African portal, which will enable minor players to make their material more readily available and to give a South African ‘brand’ to our heritage and research material. Sustainable funding could be a problem if open access (which is obviously desirable) is to be maintained. Up-to-

date IT infrastructure is essential if fast, convenient access and effective digitisation are to be achieved. This aspect is further explored in Section 13: Recommendation: A National Portal to Heritage and Research Collections.

Software choices abound, and some standardisation can have benefits. More important is the availability of good objective advice, although the history of software in most application areas inclines towards diversity rather than standardisation.

---

## 12 Audit Findings – Funding and Sustainability

---

### 12.1 Funding

The detailed funding information provided in the submissions was scant. There are several probable reasons for this.

- Only 30 of the 185 collections, representing 25 organisations, actually reported on funding.
- For all categories except the collections already digitised, the ultimate total cost is speculative – especially where the activities are ongoing
- Most of the organisations represented are not project driven, viz with all activities (including staff time and cost) attributed to registered projects. Internal funds are spent according to line categories such as Staff, Equipment and Consumables. Where the digitisation activities form only part of the duties of staff members, especially management staff, their costs are typically not attributed to the project. Unscrambling this to report costs by collection (to be) digitised seems to be beyond the capabilities (or willingness) of most respondents.

Table 12.1 gives a breakdown of funding sources reported. The category ‘Not stated’ represents those collections for which the responses were ‘Yes’, but with no further detail provided. In effect, these are equivalent to ‘Internal’.

<b>TABLE 12.1 Reported Funding Details</b>					
<b>Source</b>	<b>Number of Projects</b>				
	<b>Total</b>	<b>Digitised</b>	<b>Digitising Under Way</b>	<b>Planned</b>	<b>Not Planned</b>
Not Stated	16	1	9	3	3
Internal	6	0	3	2	1
Donor	6	1	3	2	0
Government Dept	2	2	0	0	0

There are no real surprises in this pattern. Digitisation of collections is a relatively new field. In most organisations it is in ‘start-up’ phase with enterprising individuals taking the lead using such resources as they can muster. Coordinated activities, such as DISA

Phase 2 and the National Rock Art project have attracted significant funding because of their scale and perceived significance, but the mobilising efforts of champions who create visibility and teamwork are crucial. ILAM at Rhodes has funds from at least 4 sources for different projects, a tribute to the fund-raising zeal of the current Director. Fruitful collaborations have developed between the interested and affected parties in the environmental sector and local researchers in several fields collaborate with overseas counterparts, some of whom are able to mobilise support in cash or kind.

The notion of coordination of these efforts with a view to making better use of the available sources is attractive to very few of those consulted. There is a widespread fear that this may blunt rather than augment the efforts of these champions, especially when they are focused on their own organisational management, either to provide internal funding or to give support in attracting external funding. Many interviewees spoke despairingly of their failure to muster **national** funds for heritage and other projects, and a large national fund for this purpose would be welcomed, provided that it employs enlightened management to promote activity, rather than a 'dead, bureaucratic hand' to control it. Fears of such bureaucratic control may have contributed to the lack of detail provided in the responses.

## **12.2 Sustainability**

Sustainability is a particularly intractable problem in any digitisation activity and the South African pattern of reactions is typical, with a growing awareness of the problem being largely unaccompanied by long-term planning to solve it. Most participants seem to believe, reasonably enough, that a short-term digitisation horizon has significant benefits for access, and also helps with preservation of the originals in those cases where regular handling presents a threat or the material is fragile for reasons of age or poor protection. Obviously preservation 'in perpetuity' seems desirable but the lack of a sustainable preservation plan is not seen as a barrier to undertaking digitisation in the first place. One might also reflect on the fact that the material to be digitised does not have an indefinite lifetime either, so that the introduction of digitising does not confer an additional risk. As far as funding is concerned, the focus is on digitisation, with no evidence of ring-fenced provisions for sustainable preservation.

## **12.3 Funding and Sustainability: Conclusions and Recommendations**

From the relatively sparse information provided, it appears that a good deal of start-up digitisation is under way or contemplated, with much of it being funded by existing budgets or small increments internally procured. The large collaborative projects have attracted significant funding, predominantly from overseas donors. It is also clear that the projects presently visible in the database represent a small proportion of the total.

Successful procurement of funds is seen as a key competence for many project managers and they cherish their sources, while welcoming the prospect of funds becoming more readily available, but without (more) bureaucracy. Better coordination of funding will probably assist the less successful applicants.

Sustainability of funding is seen to be important but not the top priority at present.

---

### **13 Recommendation: A National Portal to Heritage and Research Collections (including Development of a Collection Database)**

---

Web access to South African collections is presently via a multiplicity of web-pages on organisational websites, some of which are located in overseas organisations, e.g. the Aluka website. The proposal that there should be a national website for South Africa is widely favoured, provided that it satisfies the criteria which are listed below. A rationale behind each criterion is given *in italics*.

#### **13.1 Criteria for a National Portal**

- a. The portal should be hosted at a very up-to-date South African site, with broadband connectivity at the best level available in the research system, be virtually always online and provide excellent user support.
- b. The portal should have a South African ‘look and feel’. (*These two features will reassure contributors and users that the channel of access to the collections is ‘in South African hands’*)
- c. The portal should act as a one-stop access point to South African collections.
- d. In most cases, material should be available in Open Access mode. In those cases where selective access is warranted, e.g. to locations of Rock Art which needs to be protected from vandalism, a user authentication system should be available.
- e. Where possible, users should be routed, via the portal, to websites that are managed and updated on their institutional websites. (*This leaves responsibility for the content of the website in the hands of the owners or custodians of the collection, who are most likely to have the interest and resources to update the collection.*)
- f. In those cases where institutional IT is inadequate to manage the expected traffic to collections, they may be hosted at the portal site.
- g. The portal should provide functionality that is user friendly to both professional users and lay-persons, especially learners. In particular it should enable effective search of South African collections.
- h. The facilities available via the portal should be actively marketed.
- i. An Advisory and Support group comprising leading portal managers in South Africa, preferably with good networking with lead practitioners overseas, should be established.

#### **13.2 The Maintenance and Expansion of the Audit Database**

Considerable effort has gone into the creation of the audit database. This effort, and the database itself, have served as inputs to this report. If the database can be improved, with the experience of those who have created and used it, it can be a useful ongoing resource, with the caveat that it will never be ‘complete’.

The most important actions are:

- Build a user-friendly, web-accessible database from the basic data and invite contributors to check their entries for accuracy and completeness.



- Improve the submission process by focusing more specifically on readily obtainable and relevant information, e.g. collection detail and expertise availability.
- Encourage other collection champions to contribute data, perhaps by demonstrating that visibility in the database attracts donor funds.

The logical location for the database would be within the Portal, and it should be managed by the NDSC.

### **13.3 Resources for Establishing the Portal**

Collaboration with an established overseas portal team is a potentially valuable asset, where the team concerned has technology and expertise of high calibre and an apparent willingness to share it with South Africa. Issues to be addressed in entering into any working agreement with such a partner should include:-

- Ownership of the initial technology/software and availability of updates
- Duration and cost of any formal licence available from the partner
- Availability of support in installing outsourced technology
- The partner's assessment of staff and infrastructure needed to make the portal work effectively for South African users
- Availability of funding for the costs to be incurred in installing and commissioning the Portal, as well as providing a user support service, and a Business Plan that will ensure sustainability.
- Local IT support that can be outsourced if the Portal team should need it.
- All the above to be assessed in relation to the resource implications of creating a South African portal *de novo*.

It would be advantageous to establish the Advisory Group mentioned above before finalising any working agreement. .

---

## **14 Recommendation: A National Digitisation and Preservation Support Centre (NDPSC)**

---

### **14.1 Context emerging from the study**

The principal obstacle to smooth and rapid development of capacity to digitise the backlog of collections and establish a sustainable system for preserving and marketing the digital collections is lack of resources. As usual, the fundamental resource is money, but a greater supply of money will not immediately solve all the problems. Certainly more equipment will be needed, but the key resource is informed and experienced practitioners of all types.

Pockets of expertise exist in various institutions, but most are sub-critical, with precious expertise bound up in only one, or at most two, individuals; often the ones whose

enterprising spirit has established the digitisation programme. These persons are critical to success in their current environments.

It is also true that the scale and complexity of the digitisation projects vary enormously: from 'routine' EDT repository creation at the 'simple' end of the spectrum, to digitisation of tens of thousands of documents and artefacts of diverse shapes and sizes, sourced from half a dozen owners, each with a unique view of value and permissions, at the 'complex' end.

The first category is likely to involve many existing staff in institutions that will be running digitisation projects in the short term to catch up with back-logs. These staff will require guidance and training that will assist them to upgrade their skills. Training programmes aimed at these candidates are likely to attract large numbers of trainees in the first few years, but demand will tail off as the skills spread and the backlog of digitisation projects is worked off, and this should be taken into account in planning.

However some of the issues encountered, especially in the 'complex' category, can be highly technical, and expertise relating to them is hard-won – usually by experience. They include:-

- Obtaining the permissions required by law and ethical practice in contested situations
- Choosing an appropriate set of systems and standards , where a conventional solution is inadequate
- Selecting an appropriate metadata system and/or thesaurus and applying it consistently
- Identifying appropriate software
- Making effective use of IT systems to store and structure images, sometimes with compression, as well as to manage their migration to more stable media from time to time.

Access to funding across the full spectrum of activity is a major obstacle, in part because investments in training and the development of higher skills are required as well as funding for operations.

A National Digitisation and Preservation Support Centre would be a valuable resource investment for South African Heritage Research and eResearch if it can respond to all these levels by providing or improving access to expertise and training. In principle, one might advocate the assembly of a centralised team of the top specialists available in the system and set them up to provide guidance and training to both new entrants and existing professionals. In practice, this would deplete the distributed resource significantly and may actually hinder acceleration of the effort, even if the individuals involved could be persuaded to leave their current posts and relocate for this purpose. (The option of appointing a central team of promising but inexperienced staff and investing time and money in training them up is regarded as a recipe for failure and has thus not been considered.)

## **14.2 Preferred Option – A Virtual Centre of Excellence in Digitisation and Preservation**

Some of the functions will be concerned with relatively routine training and maintenance which will require full-time staff. This applies particularly to entry level training and IT support to manage the portal, as well as for maintenance of a collection database.

The more high level advisory functions, and possibly some research management, would need only part-time attention, which suits the model of recruiting highly experienced and able specialists on a part-time basis from within the systems to attend to these aspects.

A team is therefore proposed that could provide support services, call upon expertise to develop solutions – via consultation and research – and promote innovation and alignment with global practice. Much of the work would be subcontracted to specialists in the system, either by part-time appointments or retainers.

### **Outputs and Structure**

Two portfolios should be considered.

#### **Information Resources and Education Portfolio**

- Provide or convene entry level training via an online manual and short, hands-on, courses. (The equipment used for this could conceivably be made available to groups to undertake digitisation projects of manageable scale)
- Maintain the database of projects and expertise
- Coordinate advice on technology and infrastructure
- Manage the National Portal, including user support and upgrading.
- Manage shared resources (e.g. diversion of small contracts to big teams)

There would need to be a sustainable training and portal team if these activities are kept in-house, recognising that the training load will peak and then diminish, whereas the portal activity is likely to expand.

#### **Solutions and Research Portfolio**

- Develop innovative solutions to current and new problems (including permissions)
- Manage peer group consultations on systems, standards and software
- Manage R&D contracts to solve problems

The Centre Director would be full-time and could double as a Portfolio Manager. The Education Portfolio Manager would need to be full time and both Portfolio Managers would need to be experienced professionals. Specialists would be involved, by the Portfolio Managers, as required.

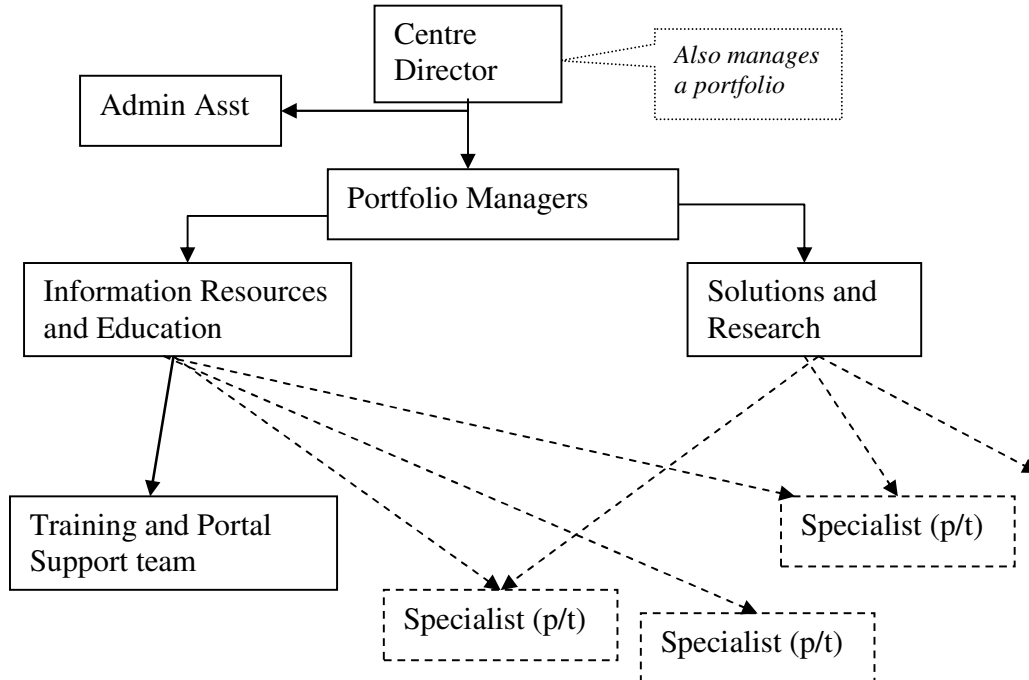
This approach provides for the Solutions and Research Portfolio Manager to identify the most informed specialists and enable them to spend time on assisting others to become proficient, without depriving their current environment of the energy and resourcefulness that made them pioneers. The Centre would thus have high quality output, at least in the form of best practice advice, almost from scratch. Starting small, it could grow in response to need.

Such a Virtual NDPSC would need a host organisation to provide administrative governance and support services. Important characteristics of this organisation would be:-

- Sound Governance and Administrative Support
- Ability to Establish and Maintain Capacity
- Attractiveness to Funders (Ability to procure and manage funding)
- High quality IT support

Technical and performance governance would be provided by a Board representing the users.

The success of such a new venture would be critically dependent on the competence of the first Director, who should be chosen carefully for the following characteristics:



- Credibility as knowledgeable and experienced in Digitisation
- Innovative and willing to learn from counterparts, local and international

- Able to manage a diverse team of part-timers.

Once again a decision matrix would support good decision-making

---

## **15 Recommendation: Funding for Sustainability**

---

From the relatively sparse information provided, it appears that a good deal of start-up digitisation is under way or contemplated, with much of it being funded by existing budgets or small increments internally procured. The large collaborative projects have attracted significant funding, predominantly from overseas donors. It is also clear that the projects presently visible in the database represent a small proportion of the total.

Successful procurement of funds is seen as a key competence for many project managers and they cherish their sources, while welcoming the prospect of funds becoming more readily available – but without (more) bureaucracy. Better coordination of funding will probably assist the less successful applicants.

Sustainability of funding is seen to be important but not the top priority at present.

An increase in the flow of government funding for digitisation would send a valuable message to the community, but this may be a vain hope in the current economic climate.

---

## **16 Recommendation: Permissions to Digitise: Best Practice Guidelines for South African Practitioners**

---

### **16.1 Context**

Proposals are frequently made for the digitisation and WWW publication of archived collections of material that is perceived to be valuable or especially interesting to a research or teaching community, or even to the general public. Because of local shortages of funds, the only available funding may come from overseas organisations. In recent times, the approach and actions of some such overseas funders have aroused suspicion among local practitioners, and cries of *cultural raiding* or *imperialism* have been raised. Not all of these are justified, but the effect has been to discourage many South African role-players from so much as contemplating giving permission for such digitisation. This will typically prevent local students and researchers from gaining easy access to the material. Also, because some of the material is decaying and needs urgently to be rescued by digitisation, such paralysis in relation to permissions does not serve any interests.

These guidelines are designed to assist decision-makers who are confronted with partnership negotiations for the digitisation of collections to ensure that they have taken the necessary steps to balance local and overseas interests in a responsible manner. They are based on

- Inputs to the Digitisation Audit

- Interviews with appropriate role-players during the Audit
- Feedback from participants in a Stakeholders' Workshop arranged to review the draft report on the Audit.

They can be used in their present form, but a better approach would be for professional groups to use them to develop their own consensus on the approach to permissions.

## 16.2 General Principles and Definitions

Before outlining the Decision Steps that must be taken before permission to digitise is given, the following principles are set out.

### 16.2.1 *The 'Digitisation Agenda'*

It may happen that a potential funder or licensee has a very carefully formulated set of objectives (an 'agenda') for digitising the collection. Any combination of the principles listed in this section may feature on such an agenda. If the South African holding archive has no 'digitisation agenda' of its own, it will not be surprising if the applicant's agenda comes to dominate –a situation often characterised as 'the tail wagging the dog'. When such an 'agenda-imbalance' arises, it may be opportune to motivate an organisational debate aimed at creating an 'own' agenda. This will, at least, ensure that one's decisions are referenced to a supportive base of organisational objectives.

### 16.2.2 *Reasons for Digitising*

Collections are usually digitised for one, or a combination, of the following reasons:

**Access:** The digital images will be made available on a website so that connected researchers anywhere in the world can access them. This obviously also helps interested parties in South Africa, who might otherwise find it expensive to travel to the archive for the purpose of perusing the documents or other items.

**Protection:** Some archived items are heavily used, e.g. by students in the relevant university or other groups. In most cases access to digital copies is adequate for such purposes and the originals will thus be better protected.

**Preservation:** Extending the protection concept, most paper and many other materials will inevitably deteriorate with time, and the creation of digital copies, possibly augmented by storage at various sites, will secure their appearance and content more effectively for later generations.

**Prestige or Reciprocity:** In many cases, **visibility** on a well-harvested website confers advantages on the author or the holding archive. This is especially true in scientific collections, where the owners of specimens will benefit from **collaboration** with counterparts elsewhere and the website is an economical way of achieving high visibility. There may also be prestige attached to holding collections belonging to **prominent personages**.

**Income generation:** In some cases there may be a perception that large amounts of money will be made from the digitised images. This is rarely true for collections as a whole, although it may be true for specific items. (See next Section)

### *16.2.3 Intellectual Property and Income Generation*

The term ‘Intellectual Property’ occurs commonly in the digitisation debate and is often used misleadingly. Strictly speaking, Intellectual Property refers to those subsets of Intellectual Capital that can be traded – hence ‘Property’- and is confined to Patents, Trademarks, Designs and Copyright, of which the latter is the only one likely to be relevant in digitisation. The web-publication of some selected heritage items may have some commercial value, but this is rare. Perceptions that the custodians or the licenses for digitised collections are likely to earn worthwhile sums from their publication are thus usually unfounded, although it is wise to be alert for possible ‘cherry-picking’. When an applicant expresses interest in digitising only a select few items custodians should be on their guard, not only because this suggests that these items are seen to have commercial value, but also that their selective publication may devalue the collection as a whole.

### *16.2.4 Permissions*

Permission for digitising is required

- When the holding archive owns the **copyright** for the material, or when copyright still resides with the author.
- And/or when the holding archive has the **right to restrict access** to the material
- And/or when the original owners, explicitly or implicitly, **excluded permission for publication** of the material in donating the material to the archive. (In most cases, archivists adopt a conservative approach to this aspect.)

**Copyright** in South Africa resides in the original author and persists for 50 years after his/er death, or the date of first public presentation, whichever is the latest.

Copyright holders, who may include the heirs of the author, can

- **license** material for publication on any conditions they choose (e.g. for single publication, publication only in a particular medium, etc). Such licensing arrangements need not (in most cases should not) allow for onward licensing. They should be made exclusive only under extraordinary circumstances, and then only for a fixed period, not ‘in perpetuity’.
- **transfer copyright** to a new holder.

In either case a **fee or royalty** can be required, such arrangements typically being embodied in legal contracts.

**South African History Archive Copyright Notice.**

All materials on this site are protected by South African copyright law and may not be reproduced, distributed, transmitted, displayed, or otherwise published in any format, without the prior written permission of the copyright owner. Please note: Every attempt has been made to obtain copyright clearance for the material on this website. However, we have not been able to trace some rightsowners. We have included their material in this digital collection, but invite anyone who can assist us with contact details for the rightsowners, to contact us at pickover.m@library.wits.ac.za so that we can obtain formal permission. Thank you.

**Identification of the copyright holder** can be difficult, when the identity of the author(s) of an item was not recorded. Intending digitisers have a number of options in such cases, including publication of notices in the popular press or attaching disclaimers, e.g. of the kind given in the box, to the digitised image or collection of images.

### **16.2.5 Partnerships**

The nature of the partnership between the funding agency, which may desire a licence to use some of the images in return for providing funding, and the archive should be carefully considered. At the root of the partnership discussion should be the understanding that both parties have something to gain and that the partnership should be, and be experienced as, at least approximately, equal. The formulation of the partnership and the operational ethos should emphasise collaboration and sharing. (Legal advice can be useful, but lawyers who have had no involvement in similar projects may give misleading advice; the South African legal system is adversarial in nature, and collaboration for the common good is not a common driving principle in legal documentation.)

### **16.2.6 Controversial and Non-Controversial Material**

Where controversy exists or may arise around the digitisation of a collection, special care will be required. Distinguishing between the controversial and the non-controversial may be challenging, but some guidelines can be identified.

**Controversial Material.** Much **political material** is controversial, either because its publication may make known hitherto concealed facts or interpretations, or because stakeholders may wish to lay down conditions relating to who may digitise and where they may publish. In some cases the publication of South African Heritage material via an obviously overseas-based website may give offence, especially if there is cost or other barriers to access by South African citizens. In these cases the licence conditions will need to be carefully thought through.

Donors of **personal material** – e.g. love-letters or other personal correspondence – may feel that limited access via an archivist is acceptable, but balk at global publicity via the WWW.

**Non-Controversial Material.** This is easier to categorise. If the material has already been published or displayed in print, or the publication of the material was clearly the intention of its creation – e.g. a newspaper or scientific article, copyright clearance may still be required but is unlikely to be problematic.



### **16.2.7      *Open or Restricted Access***

It is clearly desirable that the digitised material should be as accessible as possible, especially to South African users, and an Open Access approach is usually the preferred option. However, there may be cases where the owners and operators of a website or portal may wish to restrict access to certain subsets of metadata, e.g. the GPS location of a rock art picture which is vulnerable to damage by uncontrolled tourist activity.

Researchers, in particular, may request a moratorium on access to give them space to extract the maximum amount of value from their data before open access is allowed. The term of the moratorium should be taken from best practice in the domain – e.g. Molecular Biology – a few weeks, Sociology – several years.

Some portal owners may wish to recoup some or all of their costs by the sale of access licences to educational and research institutions. They can usually be persuaded to make licences available to South African institutions at no cost, but the need for access by unattached individuals, although probably not widespread, is difficult to address.

### **16.2.8      *Preservation***

It is widely recognised that

1. Digitisation has definite short-term value, especially in terms of access and rescue of decaying items from oblivion
2. Preservation in the long-term is desirable for the sake of future users.

However, there will be future preservation costs, including storage media migration, provision of back-up copies, augmenting and consolidating collections and weeding of low-relevance material and duplicates. The actual scale of these costs will depend on decisions that can be made only in the future. Such costs cannot reasonably be provided from current budgets, especially in institutions using government policies with respect to capitalisation which do not generally allow for provision of reserves to pay for future expenses. Furthermore, the problem also occurs in decision-making in relation to current and future research data that are born digital, where it is proving equally intractable. (ICT specialists believe that the technical challenge of preserving large volumes of material will be less daunting than the financial.)

If the lack of a clear plan for future preservation of a collection is viewed as a veto for digitisation, the entire enterprise founders fatally: better to do the best you can with present knowledge. In making decisions re prioritisation, it will be prudent to confront this issue and make an assessment, however speculative, of the relative scale of the preservation cost; to attempt more than this will be a waste of effort.

### **16.3      *Decision Steps***

If taken in the context of the above, the following decision steps should ensure attention to all relevant details

- Step 1:** Decide or determine whether the collection is non-controversial. (See 17.2.6.) If so, proceed to Step 4

**Step 2:** For controversial material, satisfy yourself that the material is digitisation-ready – viz that it is appropriately catalogued and that an informed specialist has verified that appropriate contextual material is available, preferably in digital format, to justify the publication of the collection. Beware of requests for ‘selected’ items, because these are often selected for commercial value and their publication might be misleading or undermine the viability of publication of the full collection.

**Step 3:** Confirm that digitisation of this collection is a priority in your own organisation and that the exercise is not simply being driven by some ‘overseas agenda’. (See 17.2.1)

**Step 4:** Check – as far as possible- that the material has not been digitised already. This is especially important for published items, e.g. newspapers. Use the NRF/Carnegie database. If related material is known to be archived overseas, use the planned project to leverage return of the expatriate material or, at least, the repatriation of digital copies.

**Step 5:** Obtain permission to digitise and publish from interested parties by identifying the likely author/owner, assuming that s/he is still alive, or his/er heirs. (See 17.2.4)

**Step 6:** Ensure that the Digitisation/Preservation budget covers all reasonable costs, including the direct costs of digitising as well as any assembly/contextualization costs.

**Step 7:** Claim all future use and disposal rights to the original and digitised versions of the collection for your organisation (and any identified owners). Keep a digitised version. If licences are being given, make them non-exclusive and non-transferable.

**Step 8** Ensure Open Access to the published material for members of your organisation, (and preferably all SA citizens) for as long as the life of any licences. Consider a moratorium on access for highly topical material, e.g. to enable digitisation but ensure that, local scholars have a period of privileged access.

**Step 9:** Contracts of all kinds should be based on South African legislation

**Step 10:** Resist proposals for the collection to be shipped overseas for digitisation except in those rare cases where only a small number of items are involved or very elaborate techniques are needed. In such cases the items should be accompanied by a SA custodian and specialist advice should be obtained to ensure their repatriation through customs etc.

**Step 11:** Check technical issues, including

- Whether planned metadata will be compatible with South African usage and enable effective harvesting
- Licence implications for software to be used, OSS applications to be preferred.

- Provision for long-term preservation, including media upgrading and backup. (See 17.2.8).

---

## **17 Recommendation: Focus and Prioritisation**

---

Experience gained from the survey, augmented by inputs from the participants in the Stakeholders' Workshop, identified the National Priorities as a better basis for establishing important Themes for digitisation. These have been used to generate a tentative list. (copied below)

- d. *Department of Science and Technology National Priorities*
  - *“Farmer to Pharma” strengthening the bio-economy value chain*
  - *Space science and technology*
  - *Energy security*
  - *Global change science with a focus on climate change*
  - *Human and social dynamics*
- e. *Department of Arts and Culture Strategic Plan*
  - *Heritage Sites (Note that Rock Art is already being comprehensively dealt with)*
  - *Indigenous Languages*
- f. *From the Stakeholders' Workshop*
  - *Project specific, e.g. Rivonia trial records*
  - *Records at risk: e.g. National or United Party records*
  - *Broad categories: Natural history and biodiversity.*
  - *Art and architecture*
  - *Economic development and trade*
  - *Cultural and political history*

It is recommended that this list be circulated to stakeholders with a request that they

- Add to the list, if warranted, and
- Motivate their preferences from the list.

The outcome can be reviewed by a panel appointed by the NDPSC, to yield a selection of top priority themes that address broad areas of, say, Science and History, for example.

Once the Themes have been identified, one can use a prioritisation process of the type given in Section 7.4 for prioritisation within the Theme, or for prioritisation within an Institution.

---

## 18 References

---

- Aschenbrenner, A. et al. 2008. The Future of Repositories, Patterns for (Cross-) Repository Architectures. *D-Lib Magazine*, November/December 2008. Viewed 27 January 2008. <http://www.dlib.org/dlib/november08/aschenbrenner/11aschenbrenner.html>
- Beall, J. 2005. Metadata and Data Quality Problems in the Digital Library. *Journal of Digital Information*, 2005. Viewed 2 January 2009. <https://journals.tdl.org/jodi/article/viewFile/jodi-171/68>
- Cathro, W. 2007. Digitisation in Australasia. *Serials*, vol 20, no 1, March 2007.
- DAC, 2008. Strategic Plan for 2008 – 2011. Department of Arts and Culture. 2008. [http://www.dst.gov.za/publications-policies/strategies-reports/reports/sa\\_nat\\_rd\\_strat.pdf](http://www.dst.gov.za/publications-policies/strategies-reports/reports/sa_nat_rd_strat.pdf)
- DACST, 1996. White Paper on Arts, Culture and Heritage. Department of Arts, Culture, Science and Technology, Pretoria, 4 June 1996. [http://www.dac.gov.za/white\\_paper.htm](http://www.dac.gov.za/white_paper.htm)
- DACST 2002. South Africa's National Research and Development Strategy. Department of Arts, Culture, Science and Technology. 2002. Viewed 27 March 2009. [www.dst.gov.za/publications-policies/strategies-reports/reports/sa\\_nat\\_rd\\_strat.pdf](http://www.dst.gov.za/publications-policies/strategies-reports/reports/sa_nat_rd_strat.pdf)
- Darwin Core. 2008. Biodiversity Information Standards, Taxonomic Databases Working Group. Viewed 29 December 2008. <http://wiki.tdwg.org/twiki/bin/view/DarwinCore>.
- DRIVER. 2008. Guidelines for Content Providers - Exposing textual resources with OAI-PMH.V2. *DRIVER (Digital Repository Infrastructure Vision for European Research)* November 2008. Viewed 22 January 2008. [http://www.driver-support.eu/documents/DRIVER\\_Guidelines\\_v2\\_Final\\_2008-11-13.pdf](http://www.driver-support.eu/documents/DRIVER_Guidelines_v2_Final_2008-11-13.pdf).
- DST, 10-year Plan. 2008. The 10 Year Innovation Plan for Science and Technology, Department of Science and Technology. 2008. Viewed on 27 March 2009. <http://www.dst.gov.za/publications-policies/strategies-reports/The%20Ten-year%20Plan%20for%20Science%20and%20Technology.pdf>
- Hillman, D. 2005. Using Dublin Core. *Dublin Core Metadata Initiative*. Viewed 29 December 2008. <http://dublincore.org/documents/usageguide/#whatismetadata>.
- Jones, S. et al. 2008. The Data Audit Framework: A First Step in the Data Management Challenge. *The International Journal of Digital Curation*. Issue 2, vol 3, 2008. Viewed 12 January 2009. <http://www.ijdc.net/index.php/ijdc/article/viewFile/91/62>
- Lalu, Premesh. 2007. The virtual stampede for Africa: digitisation, postcoloniality and archives of the liberation struggles in Southern Africa. *Innovation*, No.34, June 2007

Library of Congress. 2007. Sustainability of Digitisation Formats. *Planning for Library of Congress Collections*. Viewed 29 Dec 2008

<http://www.digitalpreservation.gov/formats/sustain/sustain.shtml>

McConnachie, B. 2009. MMus thesis, Rhodes University, to be published.

NARA. 2008. Strategy for Digitising Archival Materials for Public Access, 2007 – 2016. National Archives and Records Administration, USA. May 2008. Viewed 30 January 2009.

<http://www.archives.gov/digitisation/strategy.pdf>

NHC. 2007. Report on the Workshop on the Protection and Promotion of Digital Heritage Resource: Towards a National Strategy. National Heritage Council. 12 March 2007.

OAIster. 2009. OAI-PMH - the Open Archives Initiative Protocol for Metadata Harvesting. Viewed 22 January 2009. <http://www.oaister.org/>

Pickover, M. 2008. The DISA Project. Packaging South African heritage as a continuing resource: content, access, ownership and ideology *IFLA Journal* 2008; vol 34; p192. Viewed 29 Dec 2009. <http://ifl.sagepub.com/cgi/content/abstract/34/2/192>. .

Resource Description and Access (RDA) 2007. RDA Brochure . *Joint Steering Committee for Development of RDA*. 17 November 2008. Viewed 28 January 2009

<http://www.collectionscanada.gc.ca/jsc/rda.html>

Rog, J. & van Wijk, C. 2008. Evaluating File Formats for Long-term Preservation. *National Library of the Netherlands*; The Hague, The Netherlands. Viewed 28 January 2009.

[http://www.kb.nl/hrd/dd/dd\\_links\\_en\\_publicaties/publicaties/KB\\_file\\_format\\_evaluation\\_method\\_27022008.pdf](http://www.kb.nl/hrd/dd/dd_links_en_publicaties/publicaties/KB_file_format_evaluation_method_27022008.pdf)

Swan, A. & Brown, S. 2008. The Skills, Role and Career Structure of Data Scientists and Curators: An Assessment of Current Practice and Future Needs. *Report to the JISC*. July 2008. Viewed 2 January 2009.

<http://www.jisc.ac.uk/publications/publications/dataskillscareersfinalreport.aspx>

Waring, P. 2001. Edmund Barton digitisation project: *Staff Papers, National Library of Australia*. 2001. Viewed 28 January, 2008

<http://www.nla.gov.au/nla/staffpaper/2001/waring1.html>

---

## 19 Appendix A: The Context for the Audit

---

In preparing this report, it was necessary to make a study of the processes and issues involved in the digitisation of heritage and research collections. For the information of interested readers, a summary of these has been included in Appendices 19 and 20 of this report

### 19.1 eResearch and Digitisation

Recent developments in information and communication technologies (ICTs) have presented opportunities for '*new, better Research*' which has come to be described in South Africa as eResearch. eResearch includes the rapid production of new knowledge especially by the

- online sharing of research results
- enhanced computational capability,
- systematic capture and sharing of research data, and the
- long-term preservation and curation of data and information beyond the lifetime of the generating project.

Ongoing research now focuses strongly on the use of eResearch techniques in new projects. However, the access and sharing networks that enable this new knowledge generation also enable the capturing and sharing of material that constitutes a precious heritage store of intellectual capital. In particular, innovative digitisation or digital imaging technologies have been used, mainly in developed countries, to capture rare manuscripts, photographs, slides, journals, paintings, monographs, and promote their long term preservation and wider accessibility. Digitisation and Preservation are inevitably linked, especially where vulnerable collections are concerned. This has given rise to new processes, behavioural disciplines, know-how, skills and facilities that are required to support the digitisation initiatives.

In South Africa, there are national level activities related to many aspects of eResearch, including

- the creation of Open Access repositories for research outputs at several universities,
- the High Performance Computing Centre (HPCC), which is part of the Meraka Institute and housed at UCT
- the current roll-out and commissioning of the SA National Research Network (SANReN) that will provide affordable high bandwidth connectivity between local and global research organisations (including HEIs)
- establishment of a National Electronic Dissertations and Theses Central Repository, hosted at the NRF for the harvesting of ETD Repositories at SA Universities.
- management of a National Scholarly Publication Programme by the Academy of Science of South Africa (ASSAf)

However, there is no coordinated initiative to support Digitisation and Preservation of research material.

Whether organizations, institutions and countries pursue digitisation initiatives individually or collectively depends, inter alia, on availability of resources, both material and financial; the organizational or national capacity, including human and infrastructural; the nature of the knowledge resources or content that need to be digitized, and the scope and scale of the sharing that is envisaged, especially where this crosses individual or organizational boundaries.

### **19.2 Some established digitisation initiatives**

In South Africa a variety of digitisation initiatives have been undertaken, including

- The Digital Imaging Project of South Africa (DISA) – a collaborative effort by a consortium of South African Higher Education Institutions (HEIs) that has been partially funded by the Andrew Mellon Foundation and has focused on historical material, especially related to the Freedom Struggle.
- DISA has become a centre of digitisation expertise in South Africa, and provides training and support across South Africa and elsewhere in Africa. Collections are accessible from the DISA ([www.disa.ukzn.ac.za](http://www.disa.ukzn.ac.za)) and Aluka ([www.aluka.org](http://www.aluka.org)) websites.
- Sabinet has embarked on the digitisation of African journals. ([www.sabinet.co.za/open\\_access.html](http://www.sabinet.co.za/open_access.html))
- The South African Rock Art Digital Archive ([www.sarada.org](http://www.sarada.org)) is compiling a digitised record of the major South African Rock Art collections
- Many HEIs, including those at the Universities of the Witwatersrand, Cape Town, Pretoria, Johannesburg, Rhodes and Stellenbosch are digitising their library and other collections. Some activities use in-house resources; others are wholly or partially dependent on donor funding.
- The National Library of South Africa, the Nelson Mandela Foundation and the Desmond Tutu Foundation are contemplating digitisation of some of their holdings.
- Activists such as the South African History Archive promote access to material, inter alia, via digitisation and web publication.

### **19.3 Challenges to the Digitisation Community**

Together and individually these organisations have been facing the challenges of digitisation, curation, and the management and preservation of digital records. Key amongst these are:

- Many large and small collections have yet to be catalogued and organised to render them suitable for digitisation.
- Development and maintenance of knowledge and skills for digitisation; and the training and retention of skilled people
- Resources: finances, technology infrastructure (software, hardware and systems) and connectivity
- Identification and promotion of standards for the creation and management of the digital information resources;
- Policies on intellectual property rights, ownership, access licensing and prioritisation.
- No national long-term preservation strategy is in position.

Lack of resources and of capacity have kept other institutions from embarking on digitisation projects.

#### **19.4 A Collaborative Digitisation and Preservation Initiative**

There is therefore a perceived need for a **collaborative initiative** to:

- provide technical digitisation support, and services to those institutions that are not able to wholly or partially do it themselves;
- provide or mobilize experts to conduct training and support for persons and institutions that would like to embark on digitisation and digital preservation initiatives;
- coordinate collaborative digitisation and digital data preservation initiatives among HEIs and, as far as possible, NGOs and other organizations that are keen to collaborate and or provide digitisation support to others.
- facilitate the sharing of knowledge

The initiative should be accessible to all stakeholders and neither be nor perceived to be in competition with institutions and organizations active or potentially active in digitisation. This report recommends (Section 16) the establishment, with the financial support of the Carnegie Corporation of New York and in collaboration with selected South African stakeholders, a national digitisation and digital preservation centre.

#### **19.5 Actions that gave rise to the Audit**

In January 2008, the Carnegie Corporation of New York (the Corporation) through the office of the Programme Officer: Libraries and Information, Dr Rookaya Bawa, convened a meeting in South Africa with South African digitisation stakeholders. The meeting was partly prompted by the fact that the Corporation had recently received a number of proposals from different organizations and institutions in South Africa for funding to digitize South African resources as well as South African material at the United States Library of Congress. It was pointed out that in examining the South African digitisation proposals, the Corporation had identified three challenges. These were:

- gaps in technological know-how and lack of a standard approach to digitisation tools, which might inhibit the ready exporting, importing or general sharing of digital content;
- overlap of content;
- differences in approach to metadata and material packaging. In the latter case, some interpretation of content is involved, which may give rise to criticism and controversy.

The meeting of the Digitisation Stakeholders was asked to share thoughts on the priorities and needs for digitisation in South Africa and to generate a “wish list” of initiatives that could be supported, individually and/or collaboratively, by the SA national government, by foundations such as the Carnegie Corporation, and by other funders or supporters. Stakeholders who attended the meeting represented, inter alia, HEIs; Non-Governmental Organizations (NGOs) such as the Nelson Mandela Foundation; government departments e.g. the National Library of South Africa (NLSA), the National Archives, the National Indigenous Knowledge Systems Office (NIKSO) at the Department of Science and Technology; some provincial governments; and science councils e.g. NRF and the Human Sciences Research Council (HSRC).



A key discussion at the meeting centred on ethical issues, especially so-called "cultural imperialism" in which African archival collections have allegedly been "raided" for digitisation by foreign "donors" and at times, ownership has been transferred. Participants in the meeting agreed that the South African community needed guidelines for making South African content accessible to researchers without losing ownership to, or having priorities dominated by, well-endowed foreign institutions.

Consensus was reached that the Carnegie Corporation of New York should be approached to support an investigation of South African digitisation policies and technical specifications. The NRF was identified as being well-placed to coordinate the investigation task team, which should comprise a wide range of role-players. The outcome could be a national digitisation institute or centre. Other points of agreement were

- that an audit of existing digitisation projects was needed and
- that further research into the impact of digitisation on print collections and on users is desirable.

This process should contribute to the identification of a national digital archiving project with appeal for ordinary South Africans.

The brief in **Section 2** of this report is an outcome of this process.

---

## 20 Appendix B: Digitisation and Preservation Overview

---

During the course of the audit, a picture was formed of what is required to make Digitisation and Preservation happen in South Africa. The following is a description of the Players, the important distinction between having Rights to Digitise and Means to Digitise, Preservation matters, the importance of Marketing and Commercialisation of digitised material and the expected spectrum of skills and personnel that are required for success.

### 20.1 The Players

Before considering the political, technical and resource issues that arise in Digitisation and Preservation, it is useful to study the dramatis personae, because it is their skills, energy and commitment, as well as their professional perspectives that create, or resolve, important issues. The wide range of skills required in these activities is also a source of complexity and the categories listed below often overlap significantly.

Much of the material is under the management of **archivists and museum curators**, whose principles strongly favour **open access** to **complete** collections, although 'complete' may often be difficult to define. They thus cherish all the material, usually without attempting to rank items in importance, and are doughty defenders of the rights of the owners, however defined. In the interests of creating a valid context, archivists also seek to build collections by aggregating and integrating scattered 'packets' of material from various sources. The work involved in this, especially when it is proposed that the scattered collection should be digitised, can be enormous and the champions of digitisation need to have or acquire negotiation and contracting skills to set the project up, and production management skills to sustain it efficiently.

**Historians** are usually concerned with how records of past events are used to interpret the past, often with conflicting objectives. This is particularly important when priorities have to be set. Some of them are **researchers** who prize both their own material and any to which they have privileged access for its potential to yield the citable publications which are essential to their reputations. This may render them unwilling to share their data, although the modern drive for open access to research from public funding presents them with some incentives to share data. Once they have extracted as much meaning from their data as possible, they tend to have no interest in its further preservation.

**Political** and other **activists** may pursue freedom of access by focusing their efforts on extracting protected archives from their custodians or owners.

**Librarians** are often the closest professionals to the need and thus by default find themselves taking charge of digitisation. Their professional commitment to access and preservation, as well as their cataloguing skills, are significant assets in Digitisation and Preservation, but they usually need to learn about capture and presentation software as

well as production and project management. They are frequently the first line promoters of digitisation as well as being sticklers for principled action in relationship to ownership and permissions.

**IT professionals** are also frequently involved, because of the need to manage output files, servers and connectivity issues, but they tend not to take over long-term management of the process. Good relationships between IT support staff and the digitisation team are essential and it helps if the latter are well informed on IT issues. IT specialists should not be expected to have the same grasp of important factors in digitisation as will be found among archivists and librarians.

**Collection owners and custodians** (the latter often including librarians) often initiate the Digitisation and Preservation process, but sometimes need to be persuaded into it. Their permission is essential to get things under way and their domain-specific knowledge may be needed for guidance.

## 20.2 The Principles

The two main drivers for Digitisation and Preservation are

- improved Access and
- long-term Preservation

of the relevant material, and it is usually the case that these are linked and sometimes overlap. For example, documents or other materials that are in good condition may be digitised initially to enable greater accessibility, usually online. The production of back-up copies will improve the chances of survival, in the first place, and provide back-up in a different sense if the physical condition of the material deteriorates. Of course, fragility of material may be the primary cause for digitisation, but many archivists express concern that digitisation – especially of artefacts – may lead to neglect of the originals.

This overview is designed to provide a framework for the discussion of the issues that must be resolved, individually and collectively, if digitisation and preservation are to be practiced effectively and efficiently. The findings of the study are dealt with in Sections 5-12 of the report.

The issues also fall into two categories viz

- Rights – who may digitize, under what circumstances and controls, and how may the digitised material be used, and
- Means – what resources and support systems will assist in achieving the Digitisation and Preservation objectives and how they should be managed, organisationally and nationally.

## 20.3 Rights to digitise

If a team or individual has identified a need in the Access/Preservation driving framework, what gives them the right to proceed with digitisation? This situation is commonly described as ‘copyright’, but in practice the legal copyright applies in relatively few cases. Older material may technically be in the public domain, but is closely guarded by the families of the authors or original owners, or by the archives to

which they have been entrusted. In either case there may be a moral obligation to consult interested parties for permission to digitise, usually because this involves a much higher level of public accessibility than is the case in a print archive. Furthermore, copyright owners can grant exclusive or non-exclusive licences to users. The collective term ‘permissions’ may better describe this deployment of ‘rights’.

### **20.3.1 Intellectual Property Rights - Copyright**

In most cases the only intellectual property per se is **copyright**<sup>11</sup>, which is governed by legislation that aims to prevent exploitation of the rights of creators of material – e.g. authors and composers (or their heirs) – much more than the rights of performers. Once the material has been published, i.e. made available in public by performance or distribution in printed form, there is little point in attempting to prevent any person from singing a heard piece or quoting a statement or passage. In many cases, this would have been the intention of publication. However, where the copyist stands to benefit materially without attempting to compensate the composer or author, this is exploitation. Copyright is essentially a private right and the onus is on the offended owner to sue the infringer, in a civil court. Only when copyright is infringed on a large scale (piracy) can the law be invoked and criminal proceedings initiated. For copyright damage claims to be warranted, material harm to the creator or current owner of the copyright would need to have occurred for either plaintiff to go to the trouble of suing the offender, or for the criminal justice system to be activated.

Periods of validity of copyright differ in different jurisdictions; in SA copyright persists for 50 years after the death of the author of the material, or the date of first public presentation, whichever is the latest. In most parts of the world the term is 50 – 70 years but in United States copyright duration varies according to the date of creation and also with the type of material. (Note that “private” material, published only after the demise of the author, may remain copyright-protected well beyond his/er lifetime.) Heritage material that has not been published previously is thus likely to be covered by copyright – but some of it may have entered the public domain.

It is of particular importance that the act of digitisation is regarded, in many jurisdictions, as a new creation and the digitising body acquires copyright to the digitised version, in just the same way as the photographer of an item of Rock Art acquires copyright to the photograph, rather than the original work, which may in time disappear due to weathering or other assault, leaving the photographer as the sole ‘owner’ of the work.

### **20.3.2 Ownership and licensing**

Copyright is most important when the material is accessible, e.g. in a library. For many collections, however, the important consideration is ownership, viz who controls access to, and hence copying/digitisation of, the material, regardless of whether age or other criteria place it in the public domain. In practice, many archivists make a practice of consulting owners or donors of collections in their custody before agreeing to the promotion of public access to the digitised material by display on a website, whether or

---

<sup>11</sup> Intellectual Property, a distinct subset of Intellectual Capital, is a product of the mind that can be *traded* – hence ‘Property’.

not it is strictly in the public domain. This is very much the case where private correspondence or diaries are involved. Archivists may also delay the piecemeal release of material pending the accumulation of sufficient related material to create a valid context.

When permission to digitise material is given by an owner, s/he (or it, if a committee is involved) enters into a **contract** with the digitiser as to what use may be made of the material. This arrangement is best described as **licensing** although ‘copyright’ is often used as a catch-all description.

### **20.3.3**        *Heritage Ownership perceptions*

In practice, moral or ethical constraints on the **exploitation** of the creators or owners of the material may play a significant role, especially when these parties are relatively unsophisticated and hence vulnerable. Furthermore, certain materials may be perceived by their custodians and other stakeholders, e.g. communities, to have national or group ‘heritage’ value far in excess of any commercial value. These perceptions evoke strong feelings and failure to address them appropriately can be a significant obstacle to the progress of a digitisation initiative.

There is also a strong sense of ownership of heritage material in some **communities** – especially of material relating to difficult times, e.g. the Freedom Struggle in South Africa. This may extend to a national level where, for example a perception develops that ‘Northern Hemisphere interests’, however philanthropic their intentions, are ‘owning’ the presentation channel for South African heritage material. (See Section 8: Audit Findings: Ownership, Copyright and Licensing.)

## **20.4**        **Means to Digitise**

For any material to be digitised and made accessible via the web, there are certain essential requirements of capture and processing that need attention, some of which also impact on preservation in the long-term. Technicalities of obtaining an image, characterising it and manipulating it can be extremely complex, and full understanding of all the complexities is likely to elude most digitisation practitioners. However, if one is in touch with good practice, detailed technical knowledge can be bypassed, provided that practitioners become familiar with standard approaches. The need for standardisation is thus considerable. (See also 20.5. Preservation Issues)

### **20.4.1**        *Assembly of the material.*

The planned digitisation project may include in its aims the packaging of a large collection, or the accumulation of a family of scattered collections, e.g. the papers of a celebrated individual who has lodged material at various institutions. This may involve organising a variety of contemporary items so that the context is recreated and the meaning that is derived from the items involves minimal speculation.

Once the necessary **permissions** have been obtained, the logistics of assembly will depend, inter alia, on

- the **scale** of the collection; the smaller it is, the easier it will be to move to a central point for digitisation or other treatment. Items in ‘National’ ownership should be

easier to transfer to a single point for processing than where a number of private individuals are owners.

- **Complexity** of the digital capture process (see below); high resolution scanning of small (<A4) items can be carried out in many locations. Large 2D items or 3D items may need to be taken to the scanner.
  - **Perceived value or fragility** of the material, which may render the custodian reluctant to move it or allow others to handle it. This is especially true of heritage items that may have special symbolic value.

#### **20.4.2**      *Preparation of the material.*

Special treatment or procedures may be required, e.g. for fragile documents or artefacts, creased or folded papers, potentially contaminated specimens, or worn and/or damaged sound or film recordings.

#### **20.4.3**      *Description of the material (Attribution of metadata).*

The effort of digitising a collection to make it accessible via the WWW is in vain if the items cannot subsequently be found. If the digital images or information are to be ‘discovered’ by search engines within collections or harvested by federated discovery services operating up to and including the entire Internet (e.g. OAIster, Google) they need to be appropriately described. Various systems for doing this and promoting standardisation and uniformity are available. The data used in the description are called *metadata*: ‘data about data’.

The dominant harvester in archival practice is OAIster, a union catalogue of digital resources that provides access to digital resources by "harvesting" their descriptive metadata using OAI-PMH (the Open Archives Initiative Protocol for Metadata Harvesting). (OAIster, 2009) It is important to note that OAIster needs to be instructed to harvest a database, it does not ‘trawl’ the Web as Google does.

For Resource Description, OAIster promotes the use of the Dublin Core, a well-known 15 element standard set, which is widely used in South African collections. An optional additional 3 elements convert this to Dublin Core - Qualified. Other standard sets of resource descriptors are VRA Core, designed for the cultural heritage community and a set widely used for describing biological specimens named Darwin Core, which has variations designed for Ocean Biogeographic Information Systems (OBIS) and PaleoPortal, used in palaeontology. (Darwin Core, 2008) The Darwin Core set enables recording of the geographical location of an item. All these systems are encoded using XML (eXtensible Mark-up Language)

#### **20.4.4**      *Digital Capture*

The basic requirements for digitisation of simple, unbound, ≤A4, flat documents are modest: a PC, scanner and free software for simple organisation will get you going. For larger documents, or irregular shapes, or ‘non-flat’ bound items, the complexity escalates,

with a camera with some depth of field becoming necessary.<sup>12</sup> Special software is needed for Optical Character Recognition (OCR) to enable word search.

For high speed production using suitably manageable originals, semi-automatic handling can be introduced. Robotic devices that do this are becoming more common, but remain extremely expensive

There are a number of microfiche and tape scanners that use special positioning mechanisms for originals.

Significant attention is now being paid to 3-dimensional scanning, which can range from software that integrates multi-view photography (e.g. QTVR), through full laser scanning that allows rotation of the image to any angle, or computerised tomography that allows reconstruction of a 3D image of the internal layers of an object such as a skull. (See Section 11.)

Audio-visual material is also a major focus of digitisation in South Africa, and imposes its own demands on equipment and standards. The key is to have good equipment for converting analogue sound or film to digital, because further equipment for processing the digital version is easily sourced. Metadata is particularly important because no readily available means exists for ‘full track harvesting’ of a sound- or film track.

#### **20.4.5            *Resolution, Format, File Size and Storage***

The more resolution available in the scanning or photographic process, the more possibilities there are for manipulation of the resulting image, e.g. to remove dirt or creases, reduce blurring, correct or adjust colour, compensate for bending (e.g. the roll at the gutter in a large book) or even removal of the image of the fingertips of the operator who held the page flat!

As resolution increases, so does file size and hence required storage capacity and web-viewing bandwidth, and it may be necessary to compress a final image or simply reduce resolution, once processing is complete.

Final document formats are usually Portable Data Format (pdf), valued for good compression and limits on subsequent interference with the image but intermediate versions in Tagged Image File Format (tiff) are common because of the latter’s richness in detail, although this is accompanied by very large file sizes.

Sound archives tend to use MPEG-1 Audio Layer 3 (mp3) or Waveform Audio Format (wav)

#### **20.4.6            *Capture software***

The software capturing the image is important because this influences the further management of the digitized document. Useful features include

---

<sup>12</sup> This is particularly important for precious bound volumes where the tendency to bear down on the volume, placed face down on the platen, threatens fragile bindings.

- Recognition of a document's logical structure and formatting
- Optical Character Recognition (OCR) that renders graphic images of words (from scanners or cameras) into searchable text.
- Document conversion – especially from pdf to text.

Great emphasis is currently placed on the use of Open Source Software (OSS) because it comes without initial payment, but installation, customization and trouble-shooting is likely to challenge those without considerable technology expertise and this may prove expensive. One is safer if one has a network of helpful colleagues using the same packages.

#### **20.4.7 Database Management /Presentation software**

The captured images need to be organised and made accessible in a structured fashion. Two open source software suites, DSpace and Greenstone, designed for this purpose, are widely used in SA but there are others. Collections can also be housed in the library online OPAC.

Various portals have been developed for the user-friendly presentation of the digitised material and a good deal of work has been put into a federated approach to search various databases. This is an intensively studied and specialised technology. (Aschenbrenner, 2008)

#### **20.4.8 Project Management and Quality Assurance**

The above list of process steps reminds one that the effective achievement of efficient outcomes will require expert management of the complex processes involved. Both product quality and the need to preserve the value of the originals without damaging them require attention also to quality assurance. In principle, these processes have the same characteristics as will be found in any industrial production process. Beall (2005) gives a useful overview of the issues, complete with extensive references to more detailed information.

### **20.5 Preservation Issues**

Two IT issues predominate in the Preservation arena. Obviously the intention in 'preserving' material is for it to last indefinitely. However, the material in its present form is unlikely to be thus protected and it is not reasonable to demand indefinite lifetimes for digitised material. It is prudent to pay attention to plans for media migration and media for repatriation of digitised material beyond the licensing date.

The second issue is storage. The best images contain many megabytes and normal servers soon become overloaded. Final working versions usually comprise smaller files, but decisions will have to be made as to what gets saved and for how long, and when and how to migrate. Obviously migration should precede obsolescence of the current medium. Concerns over an ultimately unmanageable volume of digitised material are discounted by specialists, who believe that improvements in technology will outstrip digitisation.



There are also other technicalities, as summed up by the **seven sustainability factors** of the United States Library of Congress (Library of Congress, 2007) that require attention if preservation is to be successful. They are repeated in the box on the next page. These are experience-based guidelines that will serve as a checklist for anyone setting up a digitisation project. They also highlight some of the skills issues (see Section 20.7) that will come up.

## **20.6 Marketing and Commercialisation**

Research material will usually be sought out by potential users, provided that the collection has been well-described and harvested. Much of the digitised material may have more popular appeal, in which case it is essential that it be marketed; i.e. have attention drawn to its existence by advertisement in some form, with user friendly access routes and possibly a modest price, often designed to cover servicing costs. Some digitised items may have significant commercial value, especially photographic and sound material, and there may be a case for marketing specific items. Collections to be marketed in this way are usually carefully selected for their commercial appeal, and arrangements are often made for some of the proceeds to be paid over to 'owning communities' or descendants of a former copyright holder whose rights have lapsed.

## **20.7 Skills and People**

The key resource for Digitisation and Preservation is a **critical mass of suitably informed and skilled people**, to direct and to manage as well as to execute the work. In practice, organisations may find it most appropriate to share certain specialised skills, or even to outsource them.

### ***20.7.1 Domain Specialists for Assembly/Collection***

Many of the candidate collections for digitisation are already in existence, having been assembled over years of research and/or collection. The people involved are usually **domain specialists** of some kind. They may also be involved in identifying the location of complementary collections and in selecting key items if the size of a digitised collection must be limited because of shortages of space or funds.

### ***20.7.2 Curators to prescribe Specimen Preparation and Conservation***

Where items are old, fragile or otherwise difficult to handle, **curators** may be required to devise handling procedures that will minimise damage.

### ***20.7.3 Metadata Attribution Specialists***

This work may require specialised skills, but where the material has been catalogued in some known schema, the selection of appropriate data can be carried out by suitably briefed librarians, especially cataloguers. An emerging career path for librarians is the data librarian option.

#### **US Library of Congress 7 Sustainability Factors**

1. **Disclosure**. Degree to which complete specifications and tools for validating technical integrity exist and are accessible to those creating and sustaining digital content. A spectrum of disclosure levels can be observed for digital formats. What is most significant is not approval by a recognized standards body, but the existence of complete documentation.
2. **Adoption**. Degree to which the format is already used by the primary creators, disseminators, or users of information resources. This includes use as a master format, for delivery to end users, and as a means of interchange between systems.
3. **Transparency**. Degree to which the digital representation is open to direct analysis with basic tools, such as human readability using a text-only editor.
4. **Self-documentation**. Self-documenting digital objects contain basic descriptive, technical, and other administrative metadata.
5. **External Dependencies**. Degree to which a particular format depends on particular hardware, operating system, or software for rendering or use and the predicted complexity of dealing with those dependencies in future technical environments.
6. **Impact of Patents**. Degree to which the ability of archival institutions to sustain content in a format will be inhibited by patents.
7. **Technical Protection Mechanisms**. Implementation of mechanisms such as encryption that prevent the preservation of content by a trusted repository

#### **20.7.4            *Scanner /Camera Operators***

Familiarity with photographic methods is essential in setting up procedures for image capture. Once a handling/scanning protocol has been set up, the labour of scanning may be delegated to routine operators who need possess only enough skill to work meticulously as they have been directed. For large collections, this comprises the bulk of the work-time that will be required and this portion of the work should ideally be undertaken in a routine production-driven environment, rather than e.g., in a university library, where the work ethos is more professional. Students are often used, being both available and cheap, but they do not always have an appropriate production commitment and are seldom available on a full-time basis.

#### **20.7.5            *Data Scientists***

The destination of all digitised material is some form of database and each project will require database development and management skills, related to the server on which the database is held, so familiarity with the relevant IT methods is essential. As the field develops, specialists in data science will be needed, although even developed and well-resourced societies such as the United Kingdom have yet to tackle this effectively. (Swan & Brown, 2008).

#### **20.7.6            *Manipulators of Images for Organisation and Presentation***

Most complete collections are found via portals of varying degrees of sophistication. In getting them there the digitisers may use a variety of processing software, e.g. for OCR and image clean-up. The archivists' criterion is to capture the image of a document 'as is', but for some purposes, especially when the content is the principal focus of attention, some image clean-up is called for, in which case a back-up copy of the uncorrected original needs to be kept.

If the images are to be effectively viewed, their organisation according to some professionally recognisable framework is desirable and some effort to create a visually appealing page is desirable. Various software packages are available for this purpose, but the initial input of a domain specialist and some skill on the part of the operator will create a better product.

**20.7.7      *Project Managers/ Digital Librarians.***

Any digitisation project or working group needs a leader who knows a good deal about most aspects of the work and can manage a diverse group of skills. This can be crucial where external project funds are being used and production rates are demanding. This new role-player will typically have IT and/or Library experience and may best be described as a Digital Librarian.

## 21 Appendix C: Submissions by Organisation

<b>TABLE 21.1: Number of Submissions by Organisation</b>		
<b>Organisation</b>	<b>Department/Unit</b>	<b>No of Submissions</b>
<b>TOTAL</b>		<b>185</b>
<b>Business/Commercial</b>		<b>4</b>
Art Bank Johannesburg	City of Johannesburg	1
Sabinet	Sabinet Gateway	2
	Sabinet Online	1
<b>Government</b>		<b>19</b>
Department of Arts and Culture	Multimedia	1
eThekwini Municipal Libraries	Don Africana Library	2
eThekwini Municipality	Local History Museums	3
	Libraries & Heritage	1
National Archives & Records Service	Department of Arts and Culture	1
National Library of South Africa	Collections Management	8
Robben Island Museum	Mayibuye Centre	3
<b>Museums</b>		<b>15</b>
Iziko Museums of Cape Town	Natural History Collections	10
McGregor Museum	Photographic Department	1
Northern Flagship Institute	Transvaal Museum	4
<b>NGOs</b>		<b>5</b>
Africa South Art Initiative (ASAI)		1
SA History Archive		4
<b>Science Councils</b>		<b>30</b>
Africa Institute of South Africa	Library and Documentation Services	1
Agricultural Research Council -	Pedometric Division	1
Council for Geoscience	Information and Collections Management	2
CSIR	Satellite Applications Centre	1
	R&D Core: Information Services	1
Human Sciences Research Council	Directorate: Information Services	2
Medical Research Council	South African Medical Research Council	3
Mintek	Communications	5
SABC	Media libraries	5
SAEON	Ndlovu Node	4
South African Bureau of Standards	Sales and Promotion	1
South African Heritage Resources Agency	Corporate Affairs	1
S African Inst for Aquatic Biodiversity	Library / Information Systems	3

**TABLE 20.1 Number of Submissions by Organisation (ctd)**

<b>Organisation</b>	<b>Department/Unit</b>	<b>No of Submissions</b>
<b>Universities</b>		<b>112</b>
Cape Peninsula University of Technology	CPUT Libraries	8
North-West University	Library Services, Potchefstroom Campus	1
Rhodes University	Library	15
	ILAM	10
Stellenbosch University	Library and Information Service	7
	Department of Research Support	5
Tshwane University of Technology	Library and Information Sciences	1
University of Cape Town	Special Collections	13
	Centre for Curating the Archive	1
	Centre for Popular Memory	2
University of Johannesburg	Library and Information Centre	11
University of KwaZulu-Natal	Programme of Geomatics (Land Surveying)	1
	DISA: Digital Innovation South Africa,	1
University of Pretoria	Department of Library Services	1
	Dept of Architecture	7
	Veterinary Science	9
	Special Collections,	6
University of South Africa	Department of Library Services, Archives	1
University of the Free State	Library and Information Services	5
University of Venda	Library	1
University of Zululand	Library	1
University of the Witwatersrand	Bernard Price Inst for Palaeontological Research.	1
	Wits Art Galleries	1
	Adler Museum of Medicine	1
	Historical papers - Library	1
	Rock Art Research Institute	1

## 22 Appendix D: Respondents to the Questionnaire and their e-mail Addresses

Organisation	Contact Person		e-Mail Address
<b>Universities - Responded and/or Submitted</b>			
CPUT	Moll	Michiel	<a href="mailto:mollm@cput.ac.za">mollm@cput.ac.za</a>
NWU	Larney	Tom	<a href="mailto:tom.larney@nwu.ac.za">tom.larney@nwu.ac.za</a>
RU	Thomas	Gwenda	<a href="mailto:g.thomas@ru.ac.za">g.thomas@ru.ac.za</a>
TUT	Agyai	Vivian	<a href="mailto:Agyeimv@tut.ac.za">Agyeimv@tut.ac.za</a>
U Ven	Matodzi	TA	<a href="mailto:annah.matodzi@univen.ac.za">annah.matodzi@univen.ac.za</a>
UCT	Hart	Lesley	<a href="mailto:lesley.hart@uct.ac.za">lesley.hart@uct.ac.za</a>
UCT	Skotnes	Pippa	<a href="mailto:pippa.skotnes@uct.ac.za">pippa.skotnes@uct.ac.za</a>
UJ	Latsky	Henriette	<a href="mailto:henriette@uj.ac.za">henriette@uj.ac.za</a>
UKZN	Grossmann	Carl	<a href="mailto:grossman@ukzn.ac.za">grossman@ukzn.ac.za</a>
UNISA	Coetzee	Marie	<a href="mailto:coetzma@unisa.ac.za">coetzma@unisa.ac.za</a>
UOFS	Namponya	Clarence	<a href="mailto:crn.bib@ufs.ac.za">crn.bib@ufs.ac.za</a>
UP	Groenewald	Ria	<a href="mailto:Ria.Groenewald@up.ac.za">Ria.Groenewald@up.ac.za</a>
US	Dietrich	Keith	<a href="mailto:kd2@sun.ac.za">kd2@sun.ac.za</a>
US	Klapwijk	Wouter	<a href="mailto:wklap@sun.ac.za">wklap@sun.ac.za</a>
Wits, Rock Art Res Inst	Smith	Benjamin	<a href="mailto:benjamin.smith@wits.ac.za">benjamin.smith@wits.ac.za</a>
Wits	Rankin-Smith	Fiona	<a href="mailto:Fiona.Rankin-Smith@wits.ac.za">Fiona.Rankin-Smith@wits.ac.za</a>
Wits	Pickover	Michelle	<a href="mailto:Michele.Pickover@wits.ac.za">Michele.Pickover@wits.ac.za</a>
UWC Library	Sello	Stanley	<a href="mailto:ssello@uwc.ac.za">ssello@uwc.ac.za</a>
UKZN	Buchanan	Nora	<a href="mailto:buchanan@ukzn.ac.za">buchanan@ukzn.ac.za</a>
UCT	Ruther	Heinz	<a href="mailto:heinz.ruther@uct.ac.za">heinz.ruther@uct.ac.za</a>
UFH	Soul	Yoli	<a href="mailto:ysoul@ufh.ac.za">ysoul@ufh.ac.za</a>
UZUL	van Wyk	Brenda	<a href="mailto:bvanwyk@pan.zulu.ac.za">bvanwyk@pan.zulu.ac.za</a>
<b>Universities - No response</b>			
CUT	Goliath	L	<a href="mailto:lgoliath@cut.ac.za">lgoliath@cut.ac.za</a>
DUT	Raju	Roy	<a href="mailto:rajur@dut.ac.za">rajur@dut.ac.za</a>
Mantec	Zingu	EC	<a href="mailto:Zingu@julian.mantec.ac.za">Zingu@julian.mantec.ac.za</a>
MEDUNSA	More	Rachel	<a href="mailto:rmore@medunsa.ac.za">rmore@medunsa.ac.za</a>
NMMU	Schonken	Andre	<a href="mailto:andre.schonken@nmmu.ac.za">andre.schonken@nmmu.ac.za</a>
VUT	Gozo	Joyce	<a href="mailto:joyce@vut.ac.za">joyce@vut.ac.za</a>
WSU	Mpako-Ntusi	Tembeka	<a href="mailto:tmn@wsu.ac.za">tmn@wsu.ac.za</a>
<b>Science Councils/ National Facilities - Responded and/or Submitted</b>			
ARC	Newby	Terry	<a href="mailto:terry@arc.agric.za">terry@arc.agric.za</a>
Council for Geosciences	Price	Roger	<a href="mailto:rprice@geoscience.org.za">rprice@geoscience.org.za</a>
CSIR	V d Merwe	Adele	<a href="mailto:avdmerwe@csir.co.za">avdmerwe@csir.co.za</a>
CSIR	Neumann	Helmut	<a href="mailto:hneumann@csir.co.za">hneumann@csir.co.za</a>
HSRC	Reagon	Faye	<a href="mailto:FReagon@hsrc.ac.za">FReagon@hsrc.ac.za</a>
Mintek	Kaniappen	Manil	<a href="mailto:manilk@mintek.co.za">manilk@mintek.co.za</a>
MRC	Luke	Nomfundo	<a href="mailto:nomfundo.luke@mrc.ac.za">nomfundo.luke@mrc.ac.za</a>
NHLS	Wadee	AA	<a href="mailto:bronwyn.maroun@nhls.ac.za">bronwyn.maroun@nhls.ac.za</a>
NICD	Schoub	BD	<a href="mailto:barrys@nicd.ac.za">barrys@nicd.ac.za</a>
NICD	Steenberg	Lea	<a href="mailto:leas@nicd.ac.za">leas@nicd.ac.za</a>
SABC	Assmann	Ilse	<a href="mailto:assmanni@sabc.co.za">assmanni@sabc.co.za</a>
SABS	Nel	Anne	<a href="mailto:nelar@sabs.co.za">nelar@sabs.co.za</a>
SAEON	Stevens	Nicola	<a href="mailto:nikki@saeon.ac.za">nikki@saeon.ac.za</a>

SAIAB	Skelton	Paul	<a href="mailto:p.skelton@ru.ac.za">p.skelton@ru.ac.za</a>
SANBI	Willoughby	Selwyn	<a href="mailto:willoughby@sanbi.org">willoughby@sanbi.org</a>
SASRI	Huckett	Barbara	<a href="mailto:barbara.huckett@sugar.org.za">barbara.huckett@sugar.org.za</a>
<b>Science Councils/ National Facilities - No Response</b>			
HartRAO	Booth	R	<a href="mailto:roy@hartrao.ac.za">roy@hartrao.ac.za</a>
HMO	Malinga	S	<a href="mailto:smalinga@hmo.ac.za">smalinga@hmo.ac.za</a>
iThembalabs	Lawrie	JJ	<a href="mailto:lawrie@tlabs.ac.za">lawrie@tlabs.ac.za</a>
ORI	van der Elst	RP	<a href="mailto:rudu@ori.org.za">rudu@ori.org.za</a>
SAAO	Charles	PA	<a href="mailto:pac@sao.ac.za">pac@sao.ac.za</a>
SAWS	Terblanche	DE	<a href="mailto:deon.terblanche@weathersa.co.za">deon.terblanche@weathersa.co.za</a>

<b>Museums - Responded and/or Submitted</b>			
eThekweni Loc Hist Mus,	Ngubane	Mlungisi	<a href="mailto:Ngubanem@durban.gov.za">Ngubanem@durban.gov.za</a>
Iziko	Robertson	HG	<a href="mailto:hrobertson@iziko.org.za">hrobertson@iziko.org.za</a>
McGregor M	Jacobson	L	<a href="mailto:jake@museumsnc.co.za">jake@museumsnc.co.za</a>
Transvaal M	Mananela	Klaas	<a href="mailto:manamela@nfi.museum">manamela@nfi.museum</a>
<b>Museums - No Response</b>			
Albany M	Webley	LE	<a href="mailto:l.webley@ru.ac.za">l.webley@ru.ac.za</a>
Bayworld	Van Zyl	Sylvia	<a href="mailto:svanzyl@bayworld.co.za">svanzyl@bayworld.co.za</a>
Cape Med Museum	Kaplan	Berkah	<a href="mailto:'bkaplan@pawc.wcape.gov.za">'bkaplan@pawc.wcape.gov.za'</a>
Durban Nat Science M	Redman	GT	<a href="mailto:redmang@durban.gov.za">redmang@durban.gov.za</a>
East London M	Cole	KC	<a href="mailto:kcole@elmuseum.za.org">kcole@elmuseum.za.org</a>
Kruger Museum	Haw	Nikki	<a href="mailto:nikki@nfi.museum">nikki@nfi.museum</a>
Natal M	Taylor	Peter	<a href="mailto:ptaylor@nmsa.org.za">ptaylor@nmsa.org.za</a>
National Cultural History Museum	Balkwill	Glyn	<a href="mailto:glyn@nfi.museum">glyn@nfi.museum</a>
National M	Nuttall	RJ	<a href="mailto:Direk@nasmus.co.za">Direk@nasmus.co.za</a>
National Zoo	Nxomani	C	<a href="mailto:clifford@zoo.ac.za">clifford@zoo.ac.za</a>
NFI	Makgolo	M	<a href="mailto:mmakgolo@nfi.museum">mmakgolo@nfi.museum</a>
SAMilitary Museum	Keene	John (Maj)	<a href="mailto:Milmus@nfi.museum">Milmus@nfi.museum</a>
Sammy Marks Museum	Walters	Nerina	<a href="mailto:marks@nfi.museum">marks@nfi.museum</a>
Tswaing Crater Museum	Jikelo	C (Ms)	<a href="mailto:27127902302@worldfax.co.za">27127902302@worldfax.co.za</a>
Willem Prinsloo Agric Museum	Zdara	Mariana	<a href="mailto:zdara@nfi.museum">zdara@nfi.museum</a>

<b>NGOs - Responded and/or Submitted</b>			
Africa South Art Initiative (ASAI)	Pissarra	Mario	<a href="mailto:info@aisi.co.za">info@aisi.co.za</a>
Alan Paton Centre	Liebenberg-Barkhuizen	Estelle	<a href="mailto:liebbe@ukzn.ac.za">liebbe@ukzn.ac.za</a>
DISA	Liebetau	Pat	<a href="mailto:liebetau@ukzn.ac.za">liebetau@ukzn.ac.za</a>
Helen Suzman Foundation	Taljaard	Raenette	<a href="mailto:raenette@hsf.org.za">raenette@hsf.org.za</a>
IfaLethu	Ramdhani	Narissa	<a href="mailto:nramdhani@ifalethu.org.za">nramdhani@ifalethu.org.za</a>
NMF	Sheldon - Harris	Verne	<a href="mailto:nmf@nelsonmandela.org">nmf@nelsonmandela.org</a>
SA History Archive	Pigou	Piers	<a href="mailto:director@saha.org.za">director@saha.org.za</a>
SABINET	Hattingh	Ros	<a href="mailto:ros@sabinet.co.za">ros@sabinet.co.za</a>
SAHRA	Gous	Herma	<a href="mailto:hgous@sahra.org.za">hgous@sahra.org.za</a>
<b>NGO's - No Response</b>			
Africa Centre			<a href="mailto:info@africacentre.net">info@africacentre.net</a>
African Art Centre			<a href="mailto:anthea@afri-art.co.za">anthea@afri-art.co.za</a>
African Assn of Zoos and Aquaria			<a href="mailto:paazab@paazab.com">paazab@paazab.com</a>
African Conservation Trust			<a href="mailto:info@projectafrica.com">info@projectafrica.com</a>
Afrikaanse Taal en Kultuurvereniging			<a href="mailto:atkv@atkv.org.za">atkv@atkv.org.za</a>
Apartheid Museum	Masiza	Jacqui	<a href="mailto:research@apartheidmuseum.org">research@apartheidmuseum.org</a>
Arts & Media Access Centre			<a href="mailto:graham.f@amac.org.za">graham.f@amac.org.za</a>
Arts and Teaching Initiative	Moodley	Segi	<a href="mailto:atinitiative@agnet.co.za">atinitiative@agnet.co.za</a>
Bag factory Artists Studio			<a href="mailto:info@bagfactoryart.org.za">info@bagfactoryart.org.za</a>

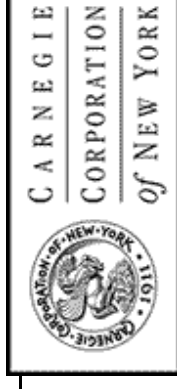
Cape Craft and Design Institute			<a href="mailto:butlerd@cput.ac.za">butlerd@cput.ac.za</a>
Cape Tercentenary Foundation	Barret	Gloria	<a href="mailto:info@cape300foundation.org.za">info@cape300foundation.org.za</a>
Cape Windjammers Education Trust			<a href="mailto:antje@capewindjammers.com">antje@capewindjammers.com</a>
Champions development trust	Kruger	Jono	<a href="mailto:jono7@webmail.co.za">jono7@webmail.co.za</a>
Childrens' Resource Centre	Fisher	Eugene	<a href="mailto:crcchild@telkomsa.net">crcchild@telkomsa.net</a>
City at Peace – Cape Town	Kamwendo	Solomon	<a href="mailto:capetown@cpnational.org">capetown@cpnational.org</a>
Craft Council SA	Moshoai Cook	Nomvula	<a href="mailto:info@craftcouncil.co.za">info@craftcouncil.co.za</a>
Cultural Development Trust			<a href="mailto:info@cuhede.org.za">info@cuhede.org.za</a>
DALRO	Davis	Yasmin	<a href="mailto:dalro@dalro.co.za">dalro@dalro.co.za</a>
Dance for All	Boyd	Philip	<a href="mailto:admin@danceforall.co.za">admin@danceforall.co.za</a>
Dance for Life	Govender	Mandy	<a href="mailto:mgoven@sa-dance4Life.org.za">mgoven@sa-dance4Life.org.za</a>
District Six Museum Foundation	Layne	Valmont	<a href="mailto:info@districtsix.co.za">info@districtsix.co.za</a>
DramAidE	Gumede	Nkhonzeni	<a href="mailto:gumedem3@ukzn.ac.za">gumedem3@ukzn.ac.za</a>
Ekhaya Multi Arts cNtre			<a href="mailto:info@ekhayaartcentre.co.za">info@ekhayaartcentre.co.za</a>
Endangered Peoples Trust	Seabell	Allan	<a href="mailto:allan@endangeredpeople.com">allan@endangeredpeople.com</a>
Extra-Mural Education Project	Gevisser	Jonathan	<a href="mailto:jg@emep.org.za">jg@emep.org.za</a>
Federasie van Afrikaanse Kultuurverenigings			<a href="mailto:fak@fak.org.za">fak@fak.org.za</a>
Field Band Foundation	Cilliers	Retha	<a href="mailto:rethacil@iafrica.com">rethacil@iafrica.com</a>
French Institute of SA	Person	Xavier	<a href="mailto:ifas@ifas.org.za">ifas@ifas.org.za</a>
From the Hip: Khulumakahle	Adams	Floss	<a href="mailto:clowns@fthk.co.za">clowns@fthk.co.za</a>
Girl Guides Association of SA	Eddy	MJ	<a href="mailto:info@ggasa.org.za">info@ggasa.org.za</a>
Grahamstown Foundation	Clayton	Louisa	<a href="mailto:admin@foundation.org.za">admin@foundation.org.za</a>
Gumboots Foundation			<a href="mailto:sa-info@gumboots.org.za">sa-info@gumboots.org.za</a>
Hope for the children			<a href="mailto:info@hopeforthechildren.co.za">info@hopeforthechildren.co.za</a>
Indigenous Peoples of Africa Coordinating Committee			<a href="mailto:ipacc@iafrica.com">ipacc@iafrica.com</a>
New Readers Publishers	Keyser	Sonya	<a href="mailto:keyser@ukzn.ac.za">keyser@ukzn.ac.za</a>
Nzumbululo Heritage Solutions	Khalushi	Rudzani	<a href="mailto:hessa5@telkomsa.net">hessa5@telkomsa.net</a>
Project Literacy			<a href="mailto:info@projectliteracy.org.za">info@projectliteracy.org.za</a>
SA Arts Exchange	Moonsamy	Roshnnie	<a href="mailto:office@artsexchange.co.za">office@artsexchange.co.za</a>
SA Catholic Bishops Conference	Brennan	Fr Vincent	<a href="mailto:vbrennan@sacbc.org.za">vbrennan@sacbc.org.za</a>
SA Choral Society	Van Der Walt	Annemarie	<a href="mailto:sachoral@xsinet.co.za">sachoral@xsinet.co.za</a>
SA History Online	Light	Joni	<a href="mailto:joni@sahistory.org.za">joni@sahistory.org.za</a>
SA San Institute	Schippers	Meryl-Joy	<a href="mailto:sasi@iafrica.com">sasi@iafrica.com</a>
SA Tennis association			<a href="mailto:satennis@mweb.co.za">satennis@mweb.co.za</a>
Sibikwa Community Theatre	Klots	Phyllis	<a href="mailto:phyllis@sibikwa.co.za">phyllis@sibikwa.co.za</a>
The BAT Centre			<a href="mailto:malangeb@mweb.co.za">malangeb@mweb.co.za</a>
Two Oceans Aquarium	Farquhar	Michael	<a href="mailto:aquarium@aquarium.co.za">aquarium@aquarium.co.za</a>

<b>Major City Public Libraries - Responded and/or Submitted</b>			
eThekwini Municipal Library	Rabillall	Renu	<a href="mailto:Rabilall@durban.gov.za">Rabilall@durban.gov.za</a>
Johannesburg Public Library	Antrobus	Ann	<a href="mailto:anna@joburg.org.za">anna@joburg.org.za</a>
<b>Major City Public Libraries - No Response</b>			
BA-Phalaborwa Lib	Mokgaboki	Steve	<a href="mailto:mokgabokis@ba-phalaborwa.org">mokgabokis@ba-phalaborwa.org</a>
Cape Town	Davids	Fadeelah	<a href="mailto:Hangberg.Library@capetown.gov.za">Hangberg.Library@capetown.gov.za</a>
Cape Town	Steyn	Ninnie	<a href="mailto:ninnie.steyn@capetown.gov.za">ninnie.steyn@capetown.gov.za</a>
Ekurhuleni Metro	Wickens	Stella	<a href="mailto:stellaw@ekurhuleni.com">stellaw@ekurhuleni.com</a>
Free State	Schimper	Jacomien	<a href="mailto:jacomien@sac.fs.gov.za">jacomien@sac.fs.gov.za</a>
Gauteng	Meyer	Koekie	<a href="mailto:Koekie.Meyer@gauteng.gov.za">Koekie.Meyer@gauteng.gov.za</a>
Jeffreys Bay Lib	Jack	Linda	<a href="mailto:ljack@ec108.org.za">ljack@ec108.org.za</a>
Mangaung	Francis	Elizabeth	<a href="mailto:bib15@civic.fs172.co.za">bib15@civic.fs172.co.za</a>
N Cape	van Dyk	Fritz	<a href="mailto:fvandyk@solplaatje.org.za">fvandyk@solplaatje.org.za</a>
Nelson Mandela Metro	Chigumbu	Bongiwe	<a href="mailto:bchigumbu@mandelametro.gov.za">bchigumbu@mandelametro.gov.za</a>



Trevor Barlow Lib Mangaung	Jacobs	Ronel	<a href="mailto:trevorb@civic.mangaung.co.za">trevorb@civic.mangaung.co.za</a>
Tshwane Community LIS	Magoro	Johannes	<a href="mailto:johannesmag@tshwane.gov.za">johannesmag@tshwane.gov.za</a>
Western Cape	Swiegelaar	Johan	<a href="mailto:jswiegel@pawc.gov.za">jswiegel@pawc.gov.za</a>

<b>National and Provincial Departments - Responded and/or Submitted</b>			
NLSA	Mpholefole	Narios	<a href="mailto:Narios.Mpholefole@nlsa.ac.za">Narios.Mpholefole@nlsa.ac.za</a>
DAC	Oberholzer	Andries	<a href="mailto:andries.oberholzer@dac.gov.za">andries.oberholzer@dac.gov.za</a>
National Archives	Motsi	Alexsio	<a href="mailto:alexsio.motsi@dac.gov.za">alexsio.motsi@dac.gov.za</a>
<b>National and Provincial Departments - No Response</b>			
DEAT	Augustyn	CJ	<a href="mailto:augustyn@deat.gov.za">augustyn@deat.gov.za</a>
Dept of Arts & Culture, KZN	Ngcoya	Sibusiso	<a href="mailto:ngcoyas@dact.kzntl.gov.za">ngcoyas@dact.kzntl.gov.za</a>
Dept of Arts and Culture, KZN	Nzimande	Sbongile	<a href="mailto:nzimandes@dact.kzntl.gov.za">nzimandes@dact.kzntl.gov.za</a>
Dept of Arts and Culture, NW	Mayet	Gulam	<a href="mailto:gmayet@nwpq.gov.za">gmayet@nwpq.gov.za</a>
Dept of Arts & Culture, N Cape	Hall	Andrew	<a href="mailto:ahall@ncpg.gov.za">ahall@ncpg.gov.za</a>
Dept of Arts & Culture, W Cape	Du Preez	Hannetjie	<a href="mailto:Hdupreez@pawc.gov.za">Hdupreez@pawc.gov.za</a>
DME	Swart	Elize	<a href="mailto:Elize.swart@dme.gov.za">Elize.swart@dme.gov.za</a>
DoE	Dlamini	Busi	<a href="mailto:busi.dlamini@doe.gov.za">busi.dlamini@doe.gov.za</a>
DoH	Neethling	Susan	<a href="mailto:susan.neethling@housing.gov.za">susan.neethling@housing.gov.za</a>
DST	Seleti	Yonah	<a href="mailto:Yonah.seleti@dst.gov.za">Yonah.seleti@dst.gov.za</a>
NHC	Mancotywa	Sanwabile	<a href="mailto:s.mancotywa@nhc.org.za">s.mancotywa@nhc.org.za</a>
SADF	Prinsloo	Gerald	<a href="mailto:sandfdoc@mweb.co.za">sandfdoc@mweb.co.za</a>



## **NRF Carnegie Audit of Digitisation Initiatives in South Africa Request for Information Introduction**

### **Output**

The Audit will culminate in an open access database of digitisation projects accompanied by a set of proposed prioritisation principles that will have been workshopped among stakeholders. The database will become a registry of collections that have been digitised and that warrant digitisation. The latter will be a handy focus for funding agencies.

### **How to Register your Ongoing and Prospective Digitisation Projects**

You are invited to contribute to the database and it is suggested that you cast the net wide to minimise the risk of valuable digitisation opportunities being missed. If you wish to contribute by completing the questionnaire in this Excel format, please feel free to do so. Note that you should duplicate the worksheet Part 2 for each collection.

On completion, please Save As 'nrf-carnegie digitisation audit – name 1- name 2.xls' where name 1 is a brief name for the collection and name 2 the initials of the person completing the questionnaire, e.g.: nrf-carnegie audit –UP archaeology – UvdS.xls and emailed to [roy@pageshipp.co.za](mailto:roy@pageshipp.co.za)

Any queries should be addressed to the Audit Contractor, Roy Page-Shipp, at the same e-mail address or **082 447 6289**.

## Structure of the Questionnaire

### **Part 1: Management and Governance**

This Section deals with the management and governance of the organisation which houses the digitisation projects. It needs to be completed only once for each organisation.

### **Part 2: Description of the collection(s) digitised or to be digitised**

This section needs to be completed for every collection, whether digitised, being digitised, planned for digitisation or warranting digitisation but with no plan formulated.

#### **Part 2.1: Contact Details of Collection Champion(s)**

This section is for contact details of the person who is directly managing the digitisation process and can answer queries re technical aspects of the project as well as on short-term and sustainability funding.

#### **Part 2.2: Collection description**

This section contains descriptive details of the items in the collection. It needs to be completed for every collection.

#### **Part 2.3: Technical aspects**

This section deals with technical aspects of the digitisation and preservation processes in place or envisaged. Please provide this information for every collection, by copying and pasting if the information is common to all collections.

#### **Part 2.4: Evaluation of the Collection. .**

This section deals with the quality of the collection (uniqueness, completeness, links to other collections etc) and the rationale for digitising it (greater research access, deterioration of specimens, relevance to heritage or research priorities). It needs to be completed for every collection

### **Part 3: Resources - Skills, Technology and Funding**

This section details the available or required resources. If the same resources are applied to several projects, please copy this section for each collection.

### **Part 4: Sustainability of Preservation**

This section describes measures in place, planned or needed for ongoing preservation of the digitised collection(s)

### **Part 5: Lessons learnt in digitising this (these) collection(s).**

This section shares special insights obtained in hands-on digitisation.

## Part 1: Management and Governance

<b>Name of Organisation</b>						
<b>Name of Unit/Department</b>						
<b>Organisation type</b>		Higher Education	Government	Science Council/ National facility	NGO/ Civil Society	Business/Commercial Research house
Mark with X						
<b>Physical address</b>						
City						
Postal Code						
<b>Postal Address</b>						
City						
Postal Code						
<b>Name of CEO/Responsible Executive</b>						
Tel. no.						
Fax No.						
Email:						

**NB THIS PART NEEDS TO BE COMPLETED IN FULL ONLY ONCE FOR EACH ORGANISATION. IF YOU ARE SUBMITTING DETAILS OF SEVERAL COLLECTIONS IN YOUR ORGANISATION, COMPLETE ONLY 'Name of Organisation' WITH THE SECOND AND SUBSEQUENT COLLECTIONS.**

**Part 2: Description of the Collection(s) digitised or to be digitised**

**Name of Collection:** \_\_\_\_\_

**Name of Department Organisation:** \_\_\_\_\_

Please complete this part for every Collection

**Part 2.1: Contact Details of Collection Champion**

**Collection Champion:** The person who manages the digitisation and preservation of this collection and can answer technical queries as well as immediate and sustainability funding queries. If several 'Champions' are involved, please provide their details

<b>Name of Collection Champion</b>			
<b>Telephone No</b>			
<b>Cell phone No</b>			
<b>Fax No</b>			
<b>e-mail address</b>			

**Part 2.2: Collection Description**

<b>Status of Collection (mark with X)</b>	Digitised	Digitising under way:	Planned	Not yet Planned
		Date of Completion		

**URL of Digitised Collection** \_\_\_\_\_

<b>Type of Material (Mark with X)</b>	<b>Text</b>			
	Printed pp	Manuscripts	Letters	Diaries
				Other (specify)

Language							
<b>Graphic</b>							
Cartoons	Photos	Slides	Posters	Artworks	Drawings: Design/Architect's	Other (specify)	
<b>Audio-Visual (Film, sound recording, DVD etc)</b>							
Specify:							
<b>Specimens (plants, seeds, microscope slides etc)</b>							
Specify:							
<b>Other</b>							
Specify:							

<b>Size of Collection (approx No of Items)</b> Mark with X	<100	100-1000	1000 - 5 000	5000 - 10 000	10 000 - 50 000	>50 000	Don't know

**Additional Information** - use free text to add relevant detail such as geographic coverage, discipline, origin of collection etc

**Copyright / Ownership** - who owns the collection, is it under copyright protection, how is copyright managed?

**Written digitisation / preservation policy** available? If so, give URL or other ref  
Example of such a policy

Yes No (Delete inapplicable)  
<http://www.nla.gov.au/policy/digitisation.html>

<b>Accessibility of Collection</b> Mark with X	Open Access	Partial Open Access	Restricted members	to
Access via Archivist				
Other: Specify				

### Part 2.3: Technical Aspects

If this is different for each collection, please complete a separate copy for each collection.

<b>Classification/Indexing Standards</b> Mark with X	Dewey Decimal	Library of Congress Subject Headings	Sears	Free indexing	text
	Other: Specify				

<b>Metadata Schema</b> Mark with X	Dublin Core	Dublin Core Qualified	METS	EAD	ISAD
	TEI (Sound) Other: Specify				

<b>Metadata Encoded as</b> xml	Other: Specify				
-----------------------------------	----------------	--	--	--	--

<b>File Formats</b> (pdf, tiff, jpeg etc) If only selected items are digitised, how are they selected?	Specify and give reasons for choice:				
---	--------------------------------------	--	--	--	--

### Part 2.4: Collection Evaluation

Please complete this part for every Collection

<b>What was, or would be, the justification for digitising and preserving this collection?</b>	Consider uniqueness, condition, urgency, deterioration of specimens, importance for research or heritage etc.				
<b>Who will benefit from the digitisation and preservation of this collection?</b>					
<b>Links</b> (if any) to related collections					

### Part 3: Resources - Skills, Technology and Funding

Digitisation is	conducted <b>In-House</b>	See below
	<b>Outsourced</b>	Name of contractor:

<b>f In-House</b>						
<b>Staff Expertise</b>	IT App Devt	Metadata	Conservators	Equipment operators	Other: Specify-	
<b>Staff Numbers</b>						

<b>Hardware</b>	<b>Type of Equipment</b>	<b>Equipment Specification</b>
	Cameras	
	Scanners	
	Camera Back	
	Other	Specify:

<b>Software</b>	<b>Capture</b>	<b>Content Management</b>	<b>Presentation</b>
Indicate whether Open Source (OS) or Proprietary (P)	Specify	Specify	Specify

<b>Quality Assurance</b>	Give details of QA and other monitoring and evaluation processes or plans		
<b>Funding</b>	<b>Full</b>	<b>Partial</b>	<b>Duration</b>
<b>Digitisation</b>			
<b>Amount Source</b>			

<b>Preservation</b>			
<b>Amount Source</b>			

<b>Ongoing Access</b>			
<b>Amount</b>			



Source

### Part 4: Sustainability

Do you have a plan for sustainability of :-

Staff?	Y?N	Digitisation	Preservation	Other: Specify
Equipment?	Y?N	Digitisation	Preservation	Media Upgrades
Access?	Y?N	Migration	Media Upgrades	
Other resources?		Specify:		

Other comments on sustainability

### Part 5: Lessons Learnt

**Thank you.** Please check that you have a Part 2 for every collection that you are submitting. Remember you can reuse the questionnaire again if you wish to submit more later.  
 Please Save As 'nrf-carnegie digitisation audit – name 1- name 2.xls' where name 1 is a brief name for the collection and name 2 the initials of the person completing the questionnaire.  
 Please Save As 'nrf-carnegie digitisation audit – name 1- name 2.xls' where name 1 is a brief name for the collection and name 2 the initials of the person completing the questionnaire eg: nrf-carnegie audit –UP archaeology – UvdS.xls and emailed to [roy@pageshipp.co.za](mailto:roy@pageshipp.co.za)

## 24 Appendix F: Staff Levels by Organisation

Name of Organisation	Collection Name	Staff Expertise	Numbers of Staff Full-Time and Part-Time					Total
			IT Apps	Meta data	Conser- vation	Equipment Operation	Other	
Agricultural Research Council	Soil Map Archive	ITAppDev, Metadata, EquipmentOperators		1				1
Cape Peninsula Univ of Tech	Institutional history	Library expertise, Project management, Metadata, Eqpt Ops		2		2		4
Council for Geosciences	Unpublished Maps	Conservators, EquipmentOperators			2	2		4
CSIR	National Digitisation and Preservation	ITAppDev, Metadata, EquipmentOperators, Preservation, Digital data management, Knowledge Management						0
Centre for Popular Memory		IT AppDev, EquipOp,	1			1	2	
Dept of Arts and Culture	Permanent Art Collection	Conservators, Equipment Operators						0
eThekweni Municipality	Ulwazi Indigenous Knowledge	Conservators, Other, operators		1		1		2
Human Sciences Res Coun	Collection of micro-data	Librarians, data managers	1	1				2
Iziko Museums of Cape Town	Natural History slides	Domain Specialists and Collection managers					10	10
Medical Research Council		ITAppDev, Metadata, EquipmentOperators	2	2	0	2		6
National Library of S A	SA Official Publications	ITAppDev, Metadata, Conservators, EquipmentOperators	3	5	2	5		15
Northern Flagship Institute	Coleoptera	Metadata, Conservators	1	1	4			6
North-West University	Potchefstroom Campus Institutional Repository	ITAppDev, Metadata, EquipmentOperators, Database development. Copyright and IP, Digital OCR	3			2		5
Rhodes University	Rhodes Resources	Metadata, Conservators, EquipmentOperators, Other, Sound engineer and sound assistants	None	2	2	4	3	11
Robben Island Museum	Mayibuye	ITAppDev, Metadata, Conservators	2	2	2			
S A Inst for Aquatic Biodiversity	SAIAB Art Collection	ITAppDev, Metadata, Conservators, EquipmentOperators	2	1	1	1		5
S A Bureau of Standards	SANS standards collection	Management, IT and outsourced	4	2	2	2	None	10
SA History Archive		Metadata		3				
Sabinet Gateway	African Online Journal Archive	Project management , Publisher Liaison Staff, Quality Controller		5		2		7
Sabinet Online	SA Gazettes		5					5
SAEON	Kruger Nat'l Park Rangers Diaries	Metadata, Conservators		1	1			2

Name of Organisation	Collection Name	Staff Expertise	Numbers of Staff Full-Time and Part-Time						Total
			IT Apps	Meta data	Conser- vation	Equipment Operation	Other		
Stellenbosch University	Documentation Centre for Music (DOMUS)	ITAppDev, Metadata, Conservators, EquipmentOperators	2	3	1	3		9	
Tshwane Univ of Tech	TUT Archives	ITAppDev, Metadata	3	6				9	
University of Johannesburg	J. Boyazoglu Artifact collection	ITAppDev, Metadata	1	3				4	
University of Johannesburg	Afrikaner on the Witwatersrand c	ITAppDev, Metadata, Conservators, EquipmentOperators	1	1	1	1		4	
University of UKZN - DISA	Struggles for Freedom, 1950 to 1994	ITAppDev, Metadata, Equipment Operators, Web site development. Digital image QC	2	1		5	4	12	
University of Pretoria	UP Collection(s)	Metadata, Digitisation expertise, basic presevation, scanning, photo editing, OCR		1		2	1	4	
University of the Free State	Archive for Contemporary Affairs (ARCA)	ITAppDev, Metadata	1	1	2	2		6	
University of Zululand	Uzulu Collection			3				3	
Witwatersrand University	Wits Art Collection	ITAppDev, EquipmentOperators	4			2		6	
Witwatersrand University	Wits University Rock Art Collection	IT Outsourced		2	1	2		5	
Witwatersrand University	Adler Museum of Medicine	curator, professional officer and collection manager			1		2	3	
Witwatersrand University	Wits Historical Papers	ITAppDev, Metadata, Conservators, Equipment Operators, Content experts						0	
<b>TOTALS</b>			<b>39</b>	<b>51</b>	<b>22</b>	<b>43</b>	<b>24</b>	<b>179</b>	

---

## 25 Appendix H: List of Contacts

---

In any exploration of this kind, the contractor encounters role-players who are especially useful sources of information. The following are recorded for future use.

Organisation	Name	e-mail
Rhodes University	Thomas, Gwenda	<a href="mailto:g.thomas@ru.ac.za">g.thomas@ru.ac.za</a>
Wits	Pickover, Michele	<a href="mailto:Michele.Pickover@wits.ac.za">Michele.Pickover@wits.ac.za</a>
Council for Geosciences	Price, Roger	<a href="mailto:rprice@geoscience.org.za">rprice@geoscience.org.za</a>
HSRC	Reagon, Faye	<a href="mailto:FReagon@hsrc.ac.za">FReagon@hsrc.ac.za</a>
SABC	Assmann, Ilse	<a href="mailto:assmanni@sabc.co.za">assmanni@sabc.co.za</a>
Iziko	Robertson, Hamish	<a href="mailto:hrobertson@iziko.org.za">hrobertson@iziko.org.za</a>
Inst for Human Evolution	Thackeray, Francis	<a href="mailto:mrsples@global.co.za">mrsples@global.co.za</a>
DISA	Liebetrau, Pat	<a href="mailto:liebetrau@ukzn.ac.za">liebetrau@ukzn.ac.za</a>
IfaLethu	Ramdhani, Narissa	<a href="mailto:nramdhani@ifalethu.org.za">nramdhani@ifalethu.org.za</a>
SA History Archive	Pigou, Piers	<a href="mailto:director@saha.org.za">director@saha.org.za</a>
RARI, Wits	Smith, Benjamin	<a href="mailto:benjamin.smith@wits.ac.za">benjamin.smith@wits.ac.za</a>
TUT, Centre for Advanced Manufacturing Technology	Wynne, Tony	<a href="mailto:wynnead@tut.ac.za">wynnead@tut.ac.za</a>