

**INQUIRY INTO FOLLOW UP OF AUDITOR-GENERAL'S
2010 FINANCIAL AUDIT REPORTS**

Organisation: Australian Museum
Name: Mr Steven Alderton
Position: Acting Director
Telephone: 9320 6000
Date Received: 4/10/2011

Theme:

Summary

OFFICE OF THE DIRECTOR
Mr Frank Howarth

Phone: (02) 9320 6110; Fax: (02) 9320 6074
Email: frank.howarth@austmus.gov.au



4 October 2011

Mr Jonathan O'Dea MP
Chair
Legislative Assembly
Public Accounts Committee
Parliament of New South Wales
Macquarie Street
Sydney NSW 2000

Dear Mr O'Dea

Re: Auditor-General's Financial Audit Repeat Recommendations

I'm writing in response to your letter to Mr Frank Howarth dated 7 September 2011 in which you requested "a submission outlining the Australian Museum's response to the Auditor-General's recommendation that the Museum complete its program of recording its collections on an electronic database as soon as possible".

Over the last 12 months the Museum has continued to re-deploy existing resources, seek additional funding and developed innovative approaches using volunteers to progress this mammoth task.

Attachment A outlines the history of applications to relevant external funding bodies and internal projects funded by the Museum to accelerate digitization of our collection records over the last 12 years.

Since the Auditor-General's review of our collection management performance we have:

- Continued the allocation of set proportions of time for collections staff to dedicate to databasing collection records;
- Allocated internal capital funds to acquire systems for recording object details within collection areas to accelerate the manual process of data entry;
- Successfully sought external funding and allocated projected staff salary savings for databasing priority collections;

nature culture discover

.../2

- Initiated projects (with Australian Government funding through the Atlas of Living Australia project and the Australian Museum Foundation) using both on-site and on-line volunteers to database biological and archival collections (see progress reports of this 'Rapid Digitisation Project' at Attachment B); and
- Engaged in the process to develop a national research infrastructure strategy being led by the Department of Innovation, Industry, Science and Research, successfully lobbying for the inclusion of collection digitization as a national priority. Availability of funding is unknown at this stage.

In late 2009 the Museum participated in a substantial funding application with other NSW Cultural Institutions who face similar challenges in databasing their collection records. A copy of the business case for this substantial project is at Attachment C. The Museum will lead the development of a refreshed bid later this year.

Under the circumstances of limited resources I consider that we are making good progress in recording our collections and thank you for your continuing interest in this regard.

Yours sincerely



Steven Alderton
Acting Director

Attachments:

- A. Summary of funding applications and projects to support accelerated digitization of collection records at the Australian Museum, 1999 - 2011.
- B. Progress reports: The Rapid Digitization Project.
- C. Business case: Digital New South Wales.

Attachment A.

**Summary of funding applications and projects to support
accelerated digitization of collection records at the Australian
Museum, 1999 – 2011.**

Summary of funding applications and projects to support accelerated digitization of collection records at the Australian Museum, 1999 - 2011

1. **Cultural Institutions Digital Asset Management proposal, 2011** : The AM is leading a project to develop a new bid to Treasury to digitise cultural collections in NSW Cultural Institutions. The first meeting of the group will be held at the Australian Museum on 21 October 2011.
2. **SEWPaC project to database the Australian Museum ant collection, 2011-12 (\$29,000):** Contract from SEWPaC to database and supply ant data for the ANHAT database.
3. **AM salary savings Digitisation Projects, 2010-11** : In 2010-11 the AM generated salary savings due to unfilled research positions and directed \$20K to databasing in each of the priority collections for databasing – Mineralogy, Palaeontology, Anthropology and Malacology.
4. **Rapid Digitisation Project, March 2011 - October 2012, (\$297,000)** : Project funded by the Atlas of Living Australia to explore innovative data and image capture using volunteers. The funding is for salary for volunteer coordinators and imaging equipment. More than 9000 collection records have been digitised as at 17 Sep 2011.
5. **Australian Museum Digitisation Project (\$90,000), 2010-11** : Capital funds were allocated to a digitisation project to complete the digitisation of Entomology types (top priority digitisation project for Natural Science Collections) and high priority cultural collection objects.
6. **Communities NSW 2009-10, Digital Asset Management (DAM) proposal** : We have participated in a bid with other cultural agencies in Communities NSW for support through the TAM plan to prepare a business case for digitisation of collections. Treasury allocated funds to prepare a business case which was completed at the end of October 2010 and subject to Gateway review in November/December 2010. We expected to hear in February 2011 if the business case would be funded. We are now assuming that this proposal will not be funded by the new government.
7. **AM salary savings 2009-10** : In 2009-10 the AM generated salary savings due to unfilled research positions and directed \$48K to a databasing position in each of the priority collections for databasing – Mineralogy, Palaeontology, Anthropology, Entomology and Malacology.
8. **Virtual volunteer Project, 2010 – 2012:** We are exploring innovative solutions for data capture in collections with the Atlas of Living Australia using virtual volunteers to transcribe data from images of data labels on the web and then upload the data to museum databases. Commenced 2010, underway.
9. **National Collaborative Research Infrastructure Strategy (NCRIS)** : We participated in the Council of Australasian Museum Directors (CAMD) proposal to the Commonwealth government to support digitisation as a part of the Atlas of Living Australia (ALA)

project. The ALA was funded but databasing was explicitly not supported through the ALA.

10. **DEWHA funding, 2005-2009** : The AM has successfully sought support from DEWHA for databasing collections of particular taxonomic groups in order to supply data to the DEWHA database (ANHAT).
11. **CHAFC, 2004** : The AM prepared a bid to the Commonwealth for funding on behalf of major museums in Australia (through Council of Heads of Australian Fauna Collections, CHAFC) to database collections. The bid was unsuccessful.
12. **Global Biodiversity Information Facility (GBIF) DIGIT Project** : The Council of Heads of Australian Faunal Collections (CHAFC) developed a project which was funded by GBIF for Australian museums to digitise records and images of vertebrate and mollusc holotypes. The Australian Museum received \$40,000 and vertebrate and mollusc holotypes were digitised.
13. **Maintenance of Effort Capital Proposal, 2003-4** : A proposal was put to Treasury following the ICAC recommendations. The proposal was to digitise collections over a five year period and to install an inventory/monitoring system. The proposal was not funded.
14. **KPMG, 2002** : We again commissioned KPMG to develop a business case for \$3M for databasing collections but this was unsuccessful.
15. **KPMG, 2000** : Integrated Collection Management System . We commissioned KPMG to develop a business case for an Integrated Collection Management System and databasing of collections. We were successful in receiving \$3M for the ICMS (EMu) but the databasing case was not funded.
16. **NSW Biodiversity strategy, 1999-2002**: The AM received funding for databasing NSW collections from the state government. We successfully argued that collection data was critical for achieving the biodiversity strategy and received funds for new records and improving the quality of existing records.

Dr Penny Berents, 21 Sep 2011
Head of Natural Science Collections

Attachment B.

Progress reports: The Rapid Digitization Project.

The Rapid Digitisation Project

Progress Report Summary

Paul Flemons, Rhiannon Stephens, Leonie Prater, Michael Elliott

The project has been proceeding for almost 6 months. The following represents a brief dot point summary of inputs and outputs of the project thus far. Parts of this summary will be included in the final report executive summary for the project to be submitted to the ALA.

1. Project Timelines

- Project commenced March 2011
- First volunteers began transcribing 10th May 2011
- These statistics for a period of 42 volunteer days from 10th May (from when volunteers started digitizing) to July 30th 2011
- Current funding ceases on October 30th 2011

2. Inputs

Inputs – setup (one offs)

- **Stage 1 (2 workstations – funded through ALA Rapid Digitisation Project)**
 - Digitising Lab
 - Cleaning and removal of furniture – included in Museums existing cleaning
 - Establish power points - ? (funded by Facilities)
 - Establish Network cabling – Approx \$5000 (capital funded)
 - Equipment
 - Cameras, Computers, Furniture, Miscellaneous lab consumables
 - 2 workstations - \$16000 (one specimen, one archives)
 - Permanent Staff Inputs
 - Training digitizing officers
 - Who then trained volunteers
 - Setting up lab
 - Purchasing equipment
 - Collection management
 - Setting up database and entry interface
- **Stage 2 (4 more workstations plus Leica microscope – funded through ALA Imaging Infrastructure Project))**
 - Equipment
 - Cameras, Computers, Copy Stands, Leica Digital Microscope and camera
 - 4 workstations - \$55000 (3 specimen labels, one high resolution large specimens)
 - 1 Leica digital microscope and computer - \$40000

Inputs – ongoing (estimated)

- Digitising Lab
 - Cleaning (covered by existing Museum arrangements)
 - Electricity and network maintenance (covered in existing Museum budgets)
 - Equipment Maintenance (included in existing Museum arrangements)
- Staff Hours
 - Digitising Officers (Leonie and Rhiannon)
 - 21 hours each/week – 42/week total

- 2184 hours a year
- Dave and Jacqui
 - 1 hrs week (fortnightly meeting plus fixing specimens)
 - 52 hours a year
- Paul – 2 hours a week, 104 a year
- Michael – 2 hours a week , 104 a year
- **Total Staff hours = 2444 per year**

3. Outputs

Fully Equipped Digitising Laboratory

- The ALA funding has enabled the establishment of a very well equipped digitizing laboratory which by mid August will consist of:-
 - 6 digitising workstations - 4 specimen/label , 1 high resolution macro specimen, and 1 archive print material - made up of:
 - Two desks
 - A powerful desktop computer
 - Camera, lens and light source
 - Copystand - for vertical photography
 - 1 Leica Microscope workstation for high quality imaging of small specimens
 - This system has only just been acquired
 - Suitably skilled volunteers will be identified and trained in its use
 - Digitising activities to be carried out on this system will be identified and prioritized in conjunction with scientific and collection staff

Well Trained Volunteer Workforce

- By mid August some 50 volunteers will be trained in the use of the specimen/label and registers digitizing workstations. This will provide the digitizing lab with a robust , versatile and flexible volunteer workforce which will enable the lab to function continuously 4 days a week, and provide reserves to cover any volunteers who are away as a result of sickness or holidays or wish to have a break. Volunteers will continue to be trained as needed to ensure the availability of enough volunteers to keep the lab functioning at full capacity 4 days a week.
- Suitably skilled volunteers will be identified and trained in the use of both the digital microscope and high resolution large specimen workstation

Specimen Digitising

- **Outcome**
 - Results in a record in Emu with taxon, registration number and an image of the specimen and associated labels in the multimedia module
 - Images of specimen and labels are made available through the Atlas of Living Australia Volunteer Portal where they are transcribed
 - Currently 800 of the 1151 labels made available have been transcribed in the space of 4 weeks – this is expected to increase as improvements are made to the site)
 - These records will be validated in the volunteer portal before being imported into Emu.

- Alternative mechanisms for capturing the full record have been identified and will be enacted should the Volunteer Portal prove not sufficiently productive for our needs (see 6. Recommendation).

- **Outputs (One Workstation)**

- Number of specimens digitized – 4814
- Number of drawers digitized – 48
- Number of workstation hours – 204.4 hrs – the number of hours the workstation has being operated by volunteers (either 1 or 2).
- Number of breakages for digitized specimens – 71 (1.65%)
- Number of specimens digitized per workstation – 24 specimens/hour
- Quality – 9 wrong out of 287 images checked – 3% (expected to reduce as system becomes more used)

Archive Digitising

- **Outputs (One Workstation)**

- Number of Register pages digitized – 7098
- Number of Register volumes digitized – 36
- Number of volunteer hours – 163 – the number of hours the workstation has being operated by volunteers (either 1 or 2).
- Number of breakages for archival material – 0
- % of breakages of archival material – 0%
- Number of Register pages digitized per volunteer hour – 32 pages/hour

4. Projected Outputs Per Year

Specimen Digitising (Four Workstations)

- Number of specimens digitized using 2444 staff hours (see ongoing Inputs above) (actual digitizing - 8 volunteers 4days/week, 5hours/day)
 - $24 \times 4 = 96$ per hour
 - At 20 hours a week - 99840 per year
 - 40.9 specimens per staff hour

Register Digitising (One Workstation)

- Number of Register pages – 66,560 (1 staff, 2 volunteers 4day/week, 7hour/day)
- This would well and truly complete the digitizing of Registers
- There is ample Archive material that could be captured to keep this workstation operational full time.

5. Table of Comparison with specimen digitization figures from KPMG report *Digitisation of Collections Records, Business Case, Australian Museum 2002*

Method of digitisation	hourly rate ¹	quantity per day ²	records entered in a year ³	Records entered per staff hour ⁴	Notes
Measured					
Direct data entry using AM technical staff ⁵	11	55	11,440 (5 hrs)	7.9	Rate per staff hour depends on how many hours are spent by a staff member in actual digitising – it is unlikely to be 7 hours when other activities are taken into account such as retrieving draws, fixing broken insects etc.
Image Capture using volunteers (1 workstation)	24	120	24,960	10.2	Assuming 2444 staff hours per year for 24960 records
Creation of Image based EMu record	24	120	24960	10.2	Process of importing images and text into Emu is about 1 minute per day of staff time and 1 hour per day of computer time. This therefore has little effect on capture rates
Creation of Full record in EMu	24	120	24960	9.9	Full record captured using ALA Volunteer Portal. Assumes VP site validates records that can then be imported into EMu. Images and metadata are exported in bulk to VP and data imported in bulk. Estimate 1 hour/month in export and 7 hours/month in import. This adds 88 hours/year to staff time making it 2532 staff hours . This will only affect the rate per staff hour not the overall capture rate
Projected (4 workstations)					
Image Capture using Volunteers ⁶	96	480	99,840	40.9	
Creation of Image based EMu record	96	480	99,840	40.9	Process of importing images and text into Emu is about 1 minute per day of staff time and 1 hour per day of computer time. This therefore has little effect on capture rates.
Creation of Full record in EMu	96	480	99,840	39.4	Full record captured using the ALA Volunteer Portal. Assumes VP site validates records that can then be imported into EMu. Images and metadata are exported in bulk to VP and data imported in bulk. Estimate 1 hour/month in export and 7 hours/month in import. This adds 88 hours/year to staff time making it 2532 staff hours . This will only affect the rate per staff hour not the overall capture rate and will not be greatly affected by a greater number of workstations operating

¹KPMG Appendix A assumed 7 hours per day, whereas I assume five hours on-the-job. The hourly rate is shown to correlate with Table 3.2: entering data for 192,857 specimens in 15 man-years.

²KPMG Table 3.2 and Appendix A suggest 3 staff x 5 years x 240 days/year = 192,857 specimens, ie 55 specimens per day.

³ Assuming working 4 days/week, 5 hours/day, 52 weeks a year

⁴ As per footnote 1 - assuming a maximum of 5 hrs spent digitising a day. The other 2 hours spent organising or attending to other matters.

⁵ Derived from KPMG report *Digitisation of Collections Records, Business Case, Australian Museum 2002*

⁶ Based on 4 workstations requiring volunteers supervised by 1 staff member

6. Recommendation

Over the past 10 years the Museum has explored a number of options for digitizing and registering specimens and these have been documented through the following reports:

- KPMG report *Digitisation of Collections Records, Business Case*, Australian Museum 2002
- *Data capture of specimen labels using volunteers*, John Tann and Paul Flemons Collection Informatics Unit Australian Museum 2008
- Report on Second Pilot Project to Digitise Hardcopy Registers and trial workflows and procedures for imaging specimens and their associated label(s). John Gollan and Paul Flemons , Collection Informatics Unit Australian Museum December 2010
- This report – Rapid Digitisation Project , Paul Flemons, Rhiannon Stephens, Leonie Prater, Michael Elliott, Collection Informatics Unit Australian Museum 2011

Based on these documents, the infrastructure developed through the Rapid Digitisation and Imaging Infrastructure projects and the impressive results achieved in these projects we believe the best approach for the Museum to pursue for digitizing its entomology collections (with potential for other collections) is as follows:

- Use a group of volunteers, trained and supervised by a paid digitizing officer employed by the Museum, to register specimens and image specimens and their labels for the purpose of digitization (see Tann and Flemons 2008 for more on pros and cons)
- Import images, and basic metadata captured with the images, into Emu to create a valid “short” database record .
- Create the “full “ database record by transcribing e full label details and georeferencing the records using one or more of a range of potential approaches including:
 - Offsite transcription through the ALA Volunter Portal then import the resultant data into Emu
 - Onsite transcription into excel spreadsheets (using the volunteer group established through the RDP) for import into Emu
 - Direct onsite entering of label data and georeferences into Emu via desktop or MVWise.

The Rapid Digitisation Project
Progress Report Summary 17/09/2011

Paul Flemons, Rhiannon Stephens, Leonie Prater, Michael Elliott

1. Outputs

Specimen Digitising

- **Outputs from all active workstations (recently increased to 4) (at 17/09/2011)**
 - Number of specimens digitized – 9154
 - Number of drawers digitized – 125
 - Number of workstation hours – 480.9 hrs – the number of hours the workstations has being operated by volunteers.
 - Number of breakages for digitized specimens – 521 (5.7%)
 - Number of specimens digitized per workstation hour – 19 specimens/hour
 - Quality – 9 wrong out of 287 images checked – 3% (expected to reduce as system becomes more used)

Archive Digitising

- **Outputs (One Workstation) (at 17/09/2011)**
 - Number of Register pages digitized – 12767
 - Number of Register volumes digitized – 63
 - Number of volunteer hours – 274.5 – the number of hours the workstation has being operated by volunteers (either 1 or 2).
 - Number of breakages for archival material – 0
 - % of breakages of archival material – 0%
 - Number of Register pages digitized per volunteer hour – 46.5 pages/hour