Implementation of the Data Seal of Approval

The Data Seal of Approval board hereby confirms that the Trusted Digital repository Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC) complies with the guidelines version 2010 set by the Data Seal of Approval Board.
The afore-mentioned repository has therefore acquired the Data Seal of Approval of 2010 on September 18, 2012.

The Trusted Digital repository is allowed to place an image of the Data Seal of Approval logo corresponding to the guidelines version date on their website. This image must link to this file which is hosted on the Data Seal of Approval website.

Yours sincerely,

The Data Seal of Approval Board

Assessment Information

Guidelines Version: 2010 | June 1, 2010
Guidelines Information Booklet: DSA-booklet_2010.pdf
All Guidelines Documentation: Documentation

Repository: Pacific and Regional Archive for Digital Sources in Endangered Cultures
                                                (PARADISEC)
Seal Acquiry Date: Sep. 18, 2012

For the latest version of the awarded DSA   http://assessment.datasealofapproval.org/seals/
for this repository please visit our website:

Previously Acquired Seals: None

This repository is owned by: **Pacific and Regional Archive for Digital Sources in Endangered
                                                Cultures (PARADISEC)**
                                                c/- School of Languages and Linguistics
                                                University of Melbourne

                                                Vic 3010 Parkville
                                                Australia

                                                T +61383448952
                                                E thien@unimelb.edu.au
                                                W http://paradisec.org.au/

Assessment

1. The data producer deposits the research data in a data repository with sufficient information for others to assess the scientific and scholarly quality of the research data and compliance with disciplinary and ethical norms.

Minimum Required Statement of Compliance:

> 3. In progress: We are in the implementation phase.

This guideline cannot be outsourced.

Applicant Entry

Statement of Compliance:

> 4. Implemented: This guideline has been fully implemented for the needs of our repository.

Evidence:

> Outline of deposit requirements are here:
> http://www.paradisec.org.au/deposit.html

> They include filenaming conventions and a deposit form to be filled out.

> The catalog is here:
> http://catalog.paradisec.org.au/

> The catalog distinguishes a collection and item level, with contextual information provided at the collection level.

> A deposit form is filled out be depositors which provides PARADISEC with a licence to use the deposited data, and confirms that it was collected according to appropriate standards and with the approval of those recorded.

> We actively promote data sharing and reuse and encourage depositors to apply our 'standard access conditions' http://www.paradisec.org.au/PDSCaccess.rtf

Reviewer Entry

Accept or send back to applicant for modification:

> Accept

Comments:

2. The data producer provides the research data in formats recommended by the data repository.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

This guideline cannot be outsourced.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Evidence:

We provide information to depositors about preferred formats for deposit and run regular training workshops to ensure practitioners understand what the formats should be, and that they use them from the creation of their research data. See http://www.paradisec.org.au/downloads.html for a range of information available to depositors.

We are able to digitise analog audio (and have done this for open reels, cassettes and minidisk) and have outsourced digitisation of video. We have also scanned or imaged manuscripts for accession into the collection.

We provide depositors with access (logon) to the catalog to enter their own metadata.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

3. The data producer provides the research data together with the metadata requested by the data repository.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

This guideline cannot be outsourced.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Evidence:

Are deposit forms which hold resource discovery metadata used?
Yes http://www.paradisec.org.au/PDSCdeposit.rtf

Are there other user-friendly ways for users to provide metadata?
Yes, via a spreadsheet (to start a collection with basic metadata items)
http://www.paradisec.org.au/PDSC_minimal_metadata.xls

This is then followed up by entry into the online catalog.

Does the repository have quality control checks to ensure that the data producer adheres to the request for metadata?
We require at least five basic metadata terms to create an item (PID, Title, Date created, country, language)

Are there tools to create metadata at the level of files?
We create metadata at the level of items, which may contain multiple files, but typically are just one file.

Are metadata elements derived from established metadata standards, registries or conventions? If so list them, and show the level of adherence to those standards.
DC, OAI, Open Language Archives Community (we are a five-star OLAC archive: http://www.language-archives.org/metrics/paradisec.org.au)

Are these metadata items relevant for the data consumers?
They are relevant inasmuch as they allow the metadata to be structured and searched via aggregators, e.g. http://www.language-archives.org/search?query=thesis&archive=paradisec.org.au&page=1

What is the repository's approach if the metadata provided are insufficient for long term archiving?
We have deceased estates where the data may not have sufficient metadata and in those cases we put as much contextual information as we can around the data and request users to assist in enriching the metadata. We also store an XML version of each item's metadata with the data in the repository.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

4. The data repository has an explicit mission in the area of digital archiving and promulgates it.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

This guideline can be outsourced.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Evidence:

Our webite lists our goals

We have an operations manual that we update regularly: http://www.paradisec.org.au/OperationsManual2011.pdf

We have promotional leaflets (in English, French, Indonesian and Chinese http://www.paradisec.org.au/downloads.html) and present at relevant conferences.

We have become part of the Australian National Data Service and are recognised a collection of national significance. We have a steering committee that will manage the collection should one of the Chief Investigators (Project Managers) stop working on it.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

5. The data repository uses due diligence to ensure compliance with legal regulations and contracts including, when applicable, regulations governing the protection of human subjects.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

This guideline cannot be outsourced.

Applicant Entry

Statement of Compliance:

3. In progress: We are in the implementation phase.

Evidence:

Is the repository a legal entity? If so, please describe its legal/organizational status.
PARADISEC is an incorporated body as described on our sponsorship page:
http://www.paradisec.org.au/sponsorship.htm

Does the repository use model contract(s) with data producers?
Yes, deposit forms as described earlier.

Does the repository use model contract(s) with data consumers?
Yes, access forms as described earlier.

Are the repository's conditions of use published?
Yes, each item has its own conditions published.

Are there measures in place if the conditions are not complied with?
This is difficult. We have had an example where a user in Europe broke the agreement he had signed with us and we could do little more than write to him and ask him to desist. Fortunately, he complied.

Does the repository ensure knowledge of and compliance with national and international laws? How?
We follow what is happening in this area but are very short -staffed and it is not a priority.

Are special procedures applied to manage data with disclosure risk?
We have levels of access to the data and strong security on the repository itself. In general we do not store material that has disclosure risk.

Are data with disclosure risk stored appropriately to limit access?
Up until now all data have been held in a closed repository with access mediated by our staff.

Are data with disclosure risk distributed under appropriate conditions?

Are procedures in place to review disclosure risk in data and to take the necessary steps to either anonymize files or to provide access in a secure way?
We have not had to do this to date but are confident we could if required.

Are staff trained in management of data with disclosure risk?
Yes

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

6. The data repository applies documented processes and procedures for managing data storage.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

This guideline can be outsourced.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Evidence:

Does the repository have a preservation policy?
Yes, this is documented in our online Operations Manual (cited above) and on our website:
http://www.paradisec.org.au/services.html

Does the repository have a strategy for backup / multiple copies? If so, please describe.
On ingest files are sent to a storage system that backs them up to tape. They are then sent to the primary repository in another city (Melbourne) where a secondary copy is also made. A copy of the whole collection is then synchronised in another city (Sydney)

Are data recovery provisions in place? What are they?
We keep a list of all archived files and the archive tapes on which they are to be found. We have regular disaster recovery exercises wherein we locate and restore a file from the backup to ensure its integrity.

Are risk management techniques used to inform the strategy?
Not as much as could be.

Are there checks on the consistency of the archive?
Yes, we run weekly checks of various kinds, and use checksums whenever files are copied.

What levels of security are acceptable for the repository?
We have high-level password access at the item-level in out repository. We allow metadata to be hidden from public view if required, and have password access to editing the catalog.

How is deterioration of storage media handled and monitored?
We constantly migrate the collection (we have been doing this for ten years now)

Or if applicable to which TDR have you outsourced?
NA

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

7. The data repository has a plan for long-term preservation of its digital assets.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

This guideline can be outsourced.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Evidence:

We adhere to international standards for data formats and trust that this will allow files to be migrated and read into the future. See the 'technicalities' section on this page: http://www.paradisec.org.au/services.html

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

8. Archiving takes place according to explicit work flows across the data life cycle.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

This guideline can be outsourced.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Evidence:

The workflow is published online as cited earlier (Operations manual)

Adherence to standards means that we consider data transformations on ingestion and have done so for some time. Deposited material in non-standard formats is converted.

We have employed trained audio conservation officers and outsource video conversion to a professional provider. Metadata systems have been developed by our team in association with specialists at the National Library of Australia, the Open Language Archives Community and the National Film and Sound Archive.

We store audio, video, images, and text, selected on the basis of relevance to support for endangered or small languages and cultures. We do not accession material related to large languages which are adequately recorded elsewhere.

Data producers can read our policies on handling of the data on our website:
http://www.paradisec.org.au/services.html

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

9. The data repository assumes responsibility from the data producers for access and availability of the digital objects.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

This guideline cannot be outsourced.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Evidence:

Does the repository have licences / contractual agreements with data producers? Please describe.
As noted earlier we have deposit and access agreements on our website
http://www.paradisec.org.au/downloads.html

Does the repository enforce licences with the data producer? How?
We apply whatever conditions the data producer (=depositor) requires for their collection.

Does the repository have a crisis management plan? Please describe.
We have recovery procedures in case of file corruption. We store XML versions of catalog entries each time an entry is saved in the catalog, thus reducing reliance on the catalog database. In case of fire/flood or the like we have offsite backups of the entire collection and the catalog.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

10. The data repository enables the users to utilize the research data and refer to them.

Minimum Required Statement of Compliance:

2. Theoretical: We have a theoretical concept.

This guideline cannot be outsourced.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Evidence:

We provide data in simple file formats (wav, mp3, mov, mp4, text, xml etc). We also accept files in formats specific to linguistic and musicological research traditions (Toolbox, Elan, Childes formats – all of which are plain text)

The catalog is searchable at http://catalog.paradisec.org.au.

OAI searching to item level is possible. RIF-CS searching is at the collection-level (http://services.ands.org.au/home/orca/rda/view/?key=paradisec.org.au%2Fcatalog)

We have PID to the item-level and are now investigating the use of doi to the file-level

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

11. The data repository ensures the integrity of the digital objects and the metadata.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

This guideline cannot be outsourced.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Evidence:

Does the repository utilise checksums? What type? How are they monitored?
We use MD5 and use checksums whenever files are moved to ensure integrity.

How is the availability of data monitored?
We have an active base of users and depositors who report on any issues.

Does the repository deal with multiple versions of the data? If so, how? Please describe the versioning strategy.
We do allow users to deposit work in progress and then recommend that they date the next version's filename (e.g., 2010831) to allow for citation of various versions.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

12. The data repository ensures the authenticity of the digital objects and the metadata.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

This guideline cannot be outsourced.

Applicant Entry

Statement of Compliance:

3. In progress: We are in the implementation phase.

Evidence:

Does the repository have a strategy for data changes? Are data producers made aware of this strategy?
We do not change data, except to add metadata if required, as in the case of BWF headers added to Wav files. Data producers sign over the material to us and so far trust that we are looking after their material properly.

Does the repository maintain provenance data and related audit trails?
We store collection-level data which provides a context for the items within the collection.

Does the repository maintain links to metadata and to other datasets, and if so, how?
We reference any external links in the catalog. In some instances we only store links to relevant items in order to bring them into the OAI search mechanism, e.g.,
http://oaister.worldcat.org/title/mota-language-bible-text/oclc/679651945&referer=brief_results

Does the repository compare the essential properties of different versions of the same file? How?
No, we leave this to depositors

Does the repository check the identities of depositors?
Yes. As we are based in a community of linguists and musicologists we are able to work out if a collection proposed for deposit is valuable or not.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

13. The technical infrastructure explicitly supports the tasks and functions described in internationally accepted archival standards like OAIS.

Minimum Required Statement of Compliance:

3. In progress: We are in the implementation phase.

This guideline can be outsourced.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Evidence:

We take seriously the need to adhere to standards. In general, we follow the International Association of Sound and Audiovisual Archives (IASA) guidelines (http://www.iasa-web.org/tc04/audio-preservation). We produce BWF audio files as recommended by IASA (http://www.iasa-web.org/technical-guidelines). We are cited as an exemplar using Digital Mass Storage Systems by the IASA www.iasa-web.org, in their Guidelines on the Production and Preservation of Digital Audio Objects (IASA-TC04). Aarhus, Denmark: International Association of Sound and Audiovisual Archives (IASA), 2004, p. 51.
"The Sub Committee on Technology of the Memory of the World Programme of UNESCO recommends these guidelines as best practice for Audio-Visual Archives."

All files conform to standard file formats, and metadata is exported to OAI-PMH (http://oaister.worldcat.org/search?q=paradisec&qt=owc_search&dblist=239&scope=0&oldscope=0) and RIF-CS (http://researchdata.ands.org.au/paradisec).

During 2012 we are redeveloping our catalog to allow for HTML5 streaming access to the collection.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

14. The data consumer complies with access regulations set by the data repository.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

This guideline cannot be outsourced.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Evidence:

Does the repository use End User Licence(s) with data consumers?
Until now we have asked users to sign a paper Access agreement and send it to us. As from the end of 2012 we will have an online clickable agreement for users.

Are there any particular special requirements which the repository's holdings require?
No

Are contracts provided to grant access to restricted-use (confidential) data?
They can be, but this is on a case-by-case basis.

Does the repository make use of special licences, e.g., Creative Commons?
Yes, a great deal of manuscript material we have produced is made available subject to CC-licences, e.g., http://paradisec.org.au/fieldnotes/AC2.htm

Are there measures in place if the conditions are not complied with?
As noted earlier, we have had an instance where the conditions were breached and we wrote to the user asking him to desist and to stop distributing material from our collection. He complied, but we do not have the resources to pursue wrongdoers.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

Provide link to access agreement

15. The data consumer conforms to and agrees with any codes of conduct that are generally accepted in higher education and scientific research for the exchange and proper use of knowledge and information.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

This guideline cannot be outsourced.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Evidence:

Does the repository need to deal with any relevant codes of conduct?
No. PARADISEC does not conduct research, it holds the products of other people's research and they are the ones responsible for ensuring their work conforms to relevant ethical standards.

Does the repository need to deal with codes of conduct specifically pertaining to protection of human subjects?
No. PARADISEC does not conduct research, it holds the products of other people's research and they are the ones responsible for ensuring their work conforms to relevant ethical standards.

What are the terms of use to which data consumers agree?
As set out in the access agreement here: http://www.paradisec.org.au/PDSCaccess.rtf

Are institutional bodies involved?
Yes, we are a multi-institutional consortium.

Are there measures in place if these codes are not complied with?
We do not have the resources to pursue wrongdoers, but can apply sanctions, such as blocking further access to the collection.

Does the repository provide guidance in the responsible use of confidential data?
We ask depositors to ensure that they have the right to deposit and that they are able to assign appropriate levels of access to their collections.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments:

16. The data consumer respects the applicable licenses of the data repository regarding the use of the research data.

Minimum Required Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

This guideline cannot be outsourced.

Applicant Entry

Statement of Compliance:

4. Implemented: This guideline has been fully implemented for the needs of our repository.

Evidence:

Are there relevant licences in place?
We have deposit (http://www.paradisec.org.au/PDSCdeposit.rtf) and access (http://www.paradisec.org.au/PDSCaccess.rtf) forms online

Are there measures in place if these codes are not complied with?
Yes, but we are not resourced sufficiently to take legal action if the situation arose that required it.

Reviewer Entry

Accept or send back to applicant for modification:

Accept

Comments: