

## **PARADISEC**

Australia lies within a region of great linguistic and cultural diversity. Over 2000 of the world's 6000 different languages are spoken in Australia, the South Pacific Islands (including around 900 languages in New Guinea alone) and Southeast Asia. Within the next century this number is likely to drop to a few hundred. The majority of these 2000 languages and their associated cultural expressions such as music are very poorly documented. Even in those languages that have begun to be documented many of the most developed cultural expressions (such as the dense and highly allusive language used in song) have never been studied.

PARADISEC (Pacific And Regional Archive for Digital Sources in Endangered Cultures) offers a facility for digital conservation and access for endangered materials from the Pacific region, defined broadly to include Oceania and East and Southeast Asia. At present no such facility exists, and our research group is developing models to ensure that the archive can provide access to interested communities, and conforms with emerging international standards for digital archiving.

In 2003 we aim to establish a framework for accessioning, cataloguing, digitising audio, text and visual material, and preserving digital copies. The primary focus of this initial stage is safe preservation of material that would otherwise be lost, especially field tapes from the 1950s and 1960s.

## **Data Preservation**

Recorded material needs to be preserved in a way that allows it to be read into the future. Reel to reel and cassette tapes are becoming obsolete and the ability to play them is becoming more difficult. If there is only one copy of the recorded material it is even more vulnerable to being lost, in cyclones, fires or simply as a result of poor storage conditions.

The current option for preserving this data is to digitise it at the best quality available and to store several copies in separate locations, as PARADISEC is doing.

## **Information discovery**

Cataloguing information (or *metadata*) will accompany all items in the collection. The collection is catalogued using descriptors based on Dublin Core and the Open Languages Archives Community (OLAC) recommendations. By using these metadata standards we are able to share information about what is located in the collection. The goal is that any resource from the region be discoverable regardless of where it is located, and regardless of where the researcher is located. Access to the data itself requires permission which is specified for each item in the collection.

## **Intellectual property issues**

Normal copyright restrictions apply, and each item in the collection has its own access conditions, as specified by the depositor and performer. If an item is distributed, the moral rights of speakers and performers are asserted.

## **Cultural renewal**

A founding principle for PARADISEC is that small and endangered cultures need support for locating and reintroducing material that was recorded in the past. Ensuring that the material is well cared for means that it can be made available into the future. Digital outputs from PARADISEC will be available in various formats depending on the needs of the users. While audio files will be stored at high resolution, they can be made available as MP3 or other formats for delivery on CD or over the web.

## **Technicalities**

The main focus of the current project is the digitisation of audio files. Our Quadriga system uses the AudioCube workstation to digitise audio material at 24-bit, 96 kHz Broadcast Wave Format (BWF). A backup version of all data is held offsite at the Australian Partnership for Advanced Computing (APAC) facility in Canberra, using the GrangeNet network to deliver the data from Sydney.

PARADISEC will establish a standard cataloguing method using metadata that conforms to the Open Languages Archives Community (OLAC).

PARADISEC is managed by a steering committee representing a consortium of three Universities, the University of Sydney, the University of Melbourne and Australian National University.

### Services offered

Digitisation of audio material.

Training in ethnographic documentation techniques of recording, data management, and data linkage.

Topics covered include:

- use and comparative advantages of types of recording equipment;
- transcription and interlinearising of texts;
- linkage of digitised audio and transcripts;
- managing information as it is collected to facilitate later access.
- assessment and cataloguing of data collections.

### Safe backup of digital data.

We offer a service of safely housing data which is not accessioned into the collection, especially work in progress or data that is not for general distribution.

### Future directions

While there is a great need for the repository offered by PARADISEC it is currently operating on limited project funding. We are investigating funding options to ensure the archival function can be put on a stronger footing.

### Contact

For further information, contact the project manager.

[nickt@paradisec.org.au](mailto:nickt@paradisec.org.au)

### Email for general enquiries

[admin@paradisec.org.au](mailto:admin@paradisec.org.au)

<http://www.paradisec.org.au>

