Cross-institutional research output
the PARADISEC experience
PARADISEC’s mission

To preserve and make accessible Australian researcher’s field recordings of endangered languages and cultures of the Asia-Pacific region
Background

• Over 2000 of the world’s 6000 languages in the Asia-Pacific region

• Number likely to fall to a few hundred by 2100 (UNESCO)

• Australian researchers active in region since 1950s - making unique recordings of unrepeatable events

• Recordings now themselves endangered (format obsolescence, media deterioration, loss of metadata)
Pacific and Regional Archive for Digital Sources in Endangered Cultures

Cross-institutional

Cls: Cliff Goddard
Hugh de Ferranti

Cls: William Foley
Allan Marett
Jane Simpson

Cls: Andrew Pawley
John Bowden
Malcolm Ross
Alan Rumsey

Cls: Steve Bird
Nick Evans
Cathy Falk
Janet Fletcher
John Hajek

Project Manager
(Metadata guru)
Nick Thieberger

Store account - web interface
Stuart Hungerford

Audio Archiving Unit
Director: Linda Barwick
Audio: Frank Davey
Project Liaison: Amanda Harris

Barwick - APSR National Perspectives on Sustainability Colloquium - 10 June 2005
Cross-disciplinary

ETHNOGRAPHIC MEDIA
(audio, video, text, image)

Qualitative analytical methodologies

Field linguistics
Ethnomusicology
Anthropology

Barwick - APSR National Perspectives on Sustainability Colloquium - 10 June 2005
Ethnography

- “The study of people in their natural settings; a descriptive account of social life and culture in a defined social system, based on qualitative methods (e.g., detailed observations, unstructured interviews, analysis of documents)” (Bowling 1997)
Qualitative research data

“Qualitative research produces data that are rich and voluminous, shedding light on the lived experience of the ‘being-in-the-world’ and the interactions inherent in complex social phenomena. Analysis of such data, however, is complex and time consuming” (Roberts & Wilson, 2002)
Cross-media

- **Audio** (interviews, music sessions, language elicitation, conversations; variety of historical formats)
- **Video** (movement studies, social interaction, sign language ... variety of historical formats)
- **Images** (photos, diagrams, graphs, maps, paintings, sketches)
- **Databases** (lexical, genealogical, taxonomic, GIS etc)
- **Texts** (fieldnotes, transcripts, analyses, dictionaries, articles) sometimes using specialist fonts (phonetic, music notation)
### Data example

**Wurm collection, Solomon Islands, 1979**

Digitised cassette tape with page image of transcript, and Wurm’s language map.

<table>
<thead>
<tr>
<th>Name</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identifier</td>
<td>SAW2-009</td>
</tr>
<tr>
<td>Title</td>
<td>Nákenaa à Nuu Nubulaa</td>
</tr>
<tr>
<td>Date</td>
<td>1979-07-19</td>
</tr>
<tr>
<td>Media</td>
<td>NSW Gov C-60 cassette</td>
</tr>
<tr>
<td>Country</td>
<td>SOLOMON ISLANDS</td>
</tr>
<tr>
<td>Type</td>
<td>lexicon</td>
</tr>
</tbody>
</table>

**Description**

Words and Explanations. Image files of accompanying documents.

**Contributor**

Wurm, S.A. (researcher)

**Contributor**

Moia, Martin (recorder)

**Metadata notes**

The "Name" column in the table above shows the name of the metadata element in the catalog and the "Value" column the value of that element. Note that not all elements have values for all catalog metadata elements.
International content

- 2100 records in PARADISEC catalogue with data on 390 languages from 50 countries including: American Samoa, Australia, Bangladesh, Botswana, Cambodia, Chile, China, Cook Islands, Fiji, French Polynesia, Greenland, Hong Kong, Iceland, India, Indonesia, Israel, Italy, Japan, Kiribati, Republic Of Korea, Lao People’s Democratic Republic, Madagascar, Malaysia, Malta, Marshall Islands, Mexico, Micronesia, Federated States Of, Myanmar, Nauru, Nepal, New Caledonia, New Zealand, Nigeria, Niue, Palau, Papua New Guinea, Philippines, Reunion, Samoa, Singapore, Solomon Islands, South Africa, Taiwan, Province of China, Thailand, Tonga, Uganda, United States of America, Vanuatu, Viet Nam, Wallis And Futuna (data as of June 2005)
Global research community

Pacific and Regional Archive for Digital Sources in Endangered Cultures

DELAMAN archives

LACITO (Paris)
ANLC (Alaska)
EMELD (Michigan)
AILLA (Texas)
PARADISEC
DOBES (Netherlands)
ELAR (London)
AIATSIS (Canberra)
AMPM (Auckland)
Global standards & formats

OPEN LANGUAGE ARCHIVES COMMUNITY

- Sub-community of Open Archives Initiative
- Worldwide virtual library of language resources
- PARADISEC one of 32 participating archives
- OAI-PMH harvesting

AIMS
- develop consensus on best current practice for digital archiving of language resources
- develop network of interoperating repositories & services for housing & accessing such resources
Re-use of research data

• Historical studies (cultural, linguistic or musical change over time)
• Cross-cultural typologies (e.g. cultural, linguistic or musical “universals”)
• Research training (hons/postgraduate projects)
• Community use (family history, repatriation to local archives, etc)
• Educational materials
• Publishing (CDs, documentaries, etc)
Baseline sustainability

If …

• Robust technical environment

• Viable backup and migration strategies for digital objects

• Longterm funding base for collection curation and maintenance

• Policy changes to recognise and encourage researcher data deposits as valued research output

• Viable governance structures for institutions and researchers to participate in voluntary federated systems

…then
Specific sustainability issues

• Managing standards compliance, version control and authentication in distributed system

• Maintaining relationships between different representations of the same event or cultural domain

• Finegrained annotation and analysis of objects - how to support and reference diversity of disciplinary approaches and annotation standards

• Rights management - ethical commitments; access for community users and international researchers; differences in international IP laws; difficulties of tracking and enforcement in networked environment
Contact us

http://www.paradisec.org.au

Linda.Barwick@paradisec.org.au (Director)

Nicholas.Thieberger@paradisec.org.au (Project Manager)

Telephone:

+61 2 9036 9557 (Sydney unit)

+61 3 9344 5185 (Melbourne unit)
Generous support from:

- Partner Universities (Sydney, Melbourne, ANU, UNE)
- Australian Research Council LIEF scheme
- Grangenet
- Australian Partnership for Sustainable Repositories (SORRT testbed)
- Australian Partnership for Advanced Computing
- ANU Internet Futures