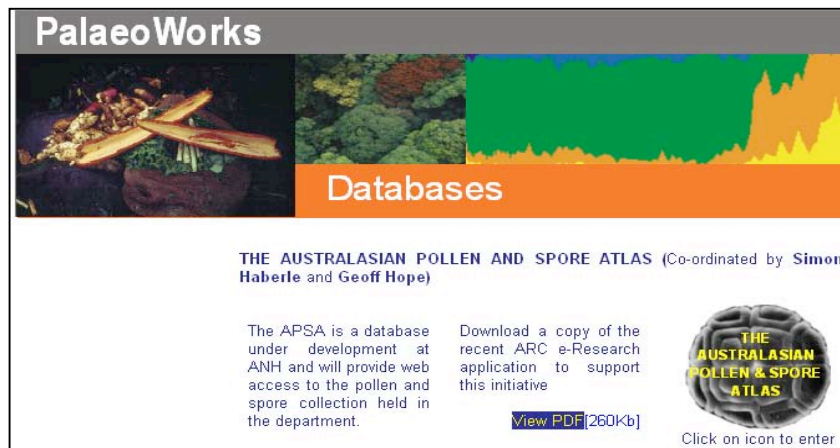


I've been asked to provide some snapshots of projects which I have been investigating in association with Kevin Bradley from the National Library as part of our work on risk analysis and metadata. We commenced by examining a sample of the kinds of projects and data collections which have been or are being done in the partner universities. At this stage we have spoken to nearly thirty people, but time will allow me to talk about only five of them. I'm not going to talk about metadata or risk analysis – that will come later – but I would like to demonstrate the variety of projects we have found and some of the features which they demonstrate.



The screenshot shows a website titled "PalaeoWorks" with a header image of pollen grains and a map. Below the header is a section titled "Databases" with the following text:

THE AUSTRALASIAN POLLEN AND SPORE ATLAS (Co-ordinated by **Simon Haberle** and **Geoff Hope**)

The APSA is a database under development at ANH and will provide web access to the pollen and spore collection held in the department.

Download a copy of the recent ARC e-Research application to support this initiative

[View PDF\[260Kb\]](#)

Click on icon to enter

The icon is a globe with the text "THE AUSTRALASIAN POLLEN & SPORE ATLAS" on it.

I'd like to start off with the **Australasian Pollen and Spore Atlas** at the Australian National University. The project is being co-ordinated by Dr Simon Haberle and Professor Geoff Hope of the Research School of Pacific and Asian Studies.

There are about 15,000 samples of pollens and spores already collected and in a physical archive. The samples come from Australia, the Pacific and Asia and have been taken from living plants – flowering plants. These constitute modern reference material that is used by archaeologists here at ANU to identify material in fossilised environments. There is a much wider range of interest in the pollen than that, however, so they are planning to create high-resolution digital images of the pollen samples which can be made available via the internet. This will entail taking the microscope slides which contain the pollen grains held in liquid and digitising them using a high-powered Zeiss microscope with axio-cam.

So who might be interested in this collection? The answer is that many disciplines have an interest in the largest pollen collection in the Southern Hemisphere. The ANU archaeologists are currently collaborating with medical researchers with interests in respiratory disease and asthma at the Menzies School of Health Research in Darwin and at the University of Sydney. There are palaeobotanists and palaeoecologists located in institutions such as the Botanic Gardens here in Canberra, Monash University, Southern Cross University and the University of Sydney Key Centre of Microscopy and Microanalysis. The Australian National Nuclear Research and Development Organisation, better known to us all as ANSTO, has an interest, although we don't know why, and then there is the honey industry. So the digitisation of a collection such as this has a potentially high impact and usage factor.

I'll move now from a proposed database to an existing and complete suite of databases held at the University of Sydney. The **Australian Beach Safety and Management Program** has been led by Professor Andrew Short who is Director of the Marine Studies Centre there, working in co-operation with Surf Life Saving Australia and with the financial assistance of the ARC. Professor Short has six collections of digital data associated with this Program and allied

analogue collections. The major database lists all 10,685 beaches in Australia and a further 850 beaches located on inhabited surrounding islands, together with measurements of 42 different variables covering such details as name, geographic location, length, gradient, presence of car parks and accommodation, zoning, beach hazards and so on. A second database provides details of the barriers behind each of these beaches, the width, height and volume of sand-dunes, etc. A third database records drainage – the existence of streams, estuaries, creeks, mangroves, lagoons, and so on. A fourth database provides references to known maps and aerial photographs, all of which are held in hard copy. A fifth digital collection contains digitised photographs of all the beaches, all taken by Professor Short himself, so there are no copyright issues. A sixth collection contains short documentary descriptions of each of the beaches, and these are used as the basis for his published books. At the moment all of the databases are held in Excel files on two computers, one at home and one at work. Images are on CD. Non-digital collections of sand, maps, aerial photographs and other photographs are all held in the Centre and are integral to understanding the digital collections. All of this has been put together since 1987.

There is significant interest in all of this information. Federal, state and local government agencies of all kinds have interests ranging through the environment, fishing, tourism, land use, town planning and maritime safety. University researchers also have interests in select aspects of the data. Individuals are interested too for details of beaches and facilities. Over time, this information will provide the start of a time series for comparative purposes. Ongoing responsibility for the digital collections is currently under negotiation. How these will continue to relate to the non-digital collections in future remains unresolved.

I'd like to move now from the beaches of Australia to **land use in Papua New Guinea** and some of the specialised collections to be found in the Research School of Pacific and Asian Studies here at the Australian National University. Dr Bryant Allen is one of a number of academics who have focussed on Papua New Guinea, looking particularly at rural development, agriculture, population and land use. Many of the academics here are approaching retirement age, and their collections of field notes, photographs, maps, and other ephemera contain invaluable information, often deteriorating and lacking any comprehensive curatorial oversight. Ideally, the School would like to see all of these collections digitised so that their contents are readily available to a wider audience for further analysis and comparisons.

Dr Allen has worked in the field of economic geography for over thirty years and is currently engaged in trying to get some of these collections digitised. One in particular is of interest. It is made up of about 2,000 photographs, primarily aerial photographs taken during intensive surveys going back to the 1950s. These were taken by the CSIRO Land Use Division and handed on to Dr Allen when the Division was closed down. The photographs came with a paper-based index and this has been converted recently into digital form. The photographs have been scanned and will be linked to the index before moving to DSpace. Ideally, they would like to add geo-spacial location and AgSystem information so that someone can select an area and see the photographic image and related agricultural details. This would, however, be costly. All of this data was created with funding from the Australian Government, through various agencies. Funding has been found to convert the images, and basic metadata in this instance was available from the index provided. For the many thousands of other photographs held by Dr Allen, however, there is no funding available for conversion. While scanning is reasonably straightforward, the creation of metadata, those descriptions which provide access to the collections, is time consuming. It can only be done with reference to the notebooks of the field workers. To have this information available digitally would enable longitudinal analysis of land use trends and practices.

Dr Bruce Millar is the Deputy Director of the Research School of Information Science and Engineering at ANU. Over the past 25 years, he has collected a significant number of recordings of spoken language, originally using analogue recording techniques and more recently digital. The early recordings have been lost as technology has been superseded. These have been collected with a view to understanding information processing models, but have been used for various purposes by researchers from different disciplinary backgrounds. Together, they occupy some 40 gigabytes of computer memory. The collection includes the **Australian National Database Of Spoken Language (ANDOSL)** prepared at the School of Electrical Engineering at the University of Sydney, the National Acoustic Laboratories, the Speech Hearing and Language Research Centre at Macquarie University and the Computer Sciences Laboratory of the Australian National University. Major funding for the project was provided by the Australian Research Council in the form of research infrastructure grants (1991-1995).

Australian National Database of Spoken Language (ANDOSL)



Dr Millar is currently seeking a permanent home for this collection as he is due to retire in the next couple of months. It is a valuable collection, containing a comprehensive record of the speech habits of Australians, both native-born and overseas-born. The speakers were rigorously selected within phonologically defined speaker groups and each group balanced for age ranges and gender. The recordings are of high quality and each recording is comprehensively described. Dr Millar is interested in the possibility of using DSpace to house the collection.

I'd now like to move from a big collection to a tiny collection, just one individual work. In 1911, King George V visited India to be presented as the crowned Emperor of India. Following the ceremony he indulged his passion for hunting in a tiger shoot using the elephant "hunting ring" technique. In the course of the hunt, 600 elephants were employed and the kill numbered 39 tigers, 18 rhinos and 4 bears.



Someone, and we don't know who, photographed the shoot. Fifty of the photographs were assembled into an album, beautifully bound and inscribed. This album was discovered many decades later being used as a tea tray in a rural Indian home in Madhya Pradesh by Dr U.N. Bhati, later a Visiting Fellow in Economics and Marketing in the School of Resources, Environment and Society at The Australian National University. Dr Bhati acquired the album and in time donated it to the ANU Library. Library staff have scanned the

individual photographs which have been made available on the web together with further details of the hunt.



In 2004, a member of the Library staff received an email from a researcher in Britain seeking permission to use some of the photographs in a forthcoming journal article and suggesting that two of the Library staff should be named as co-authors, given the depth and quality of their research on the hunt. They agreed and the article is to be published in the next issue of the online journal, *Pachyderm*. The researcher concerned works for the Rhino Resource Centre in England, an offshoot of the World Conservation Union Species Survival Programme.

We have learned a number of lessons from this survey. Firstly, we have learned that research is indeed multidisciplinary and that information collected for one purpose may have multiple other uses, sometimes in ways that no one could have foreseen. We have learned that there are various models of ownership and production of digital data. The collection owned by the individual researcher or group of researchers is one model. But also important are collections maintained and developed for other purposes by entities such as libraries, archives, publishers and digital repositories. We have also learned that collections of digital data may relate closely to collections in physical formats, and that responsibility for the care of physical objects is not always defined. We have learned that there is a significant cost to the conversion of digital objects and that a high proportion of that cost is in the description of that data, the creation of metadata.