



Respect, trust and engagement: creating an Australian indigenous data archive

Gabrielle Gardiner
Jemima McDonald
Alex Byrne and
Kirsten Thorpe
University of Technology, Sydney
Gabrielle.Gardiner@uts.edu.au

Meeting: 86. Social Science Libraries with Information Literacy

WORLD LIBRARY AND INFORMATION CONGRESS: 76TH IFLA GENERAL CONFERENCE AND ASSEMBLY
10-15 August 2010, Gothenburg, Sweden
<http://www.ifla.org/en/ifla76>

Abstract

The Aboriginal and Torres Strait Islander Data Archive (ATSIDA) provides a unique link between research data and community development. With data, statistics and datasets that comprise everything from SPSS numerical data, through to bark paintings and song recordings, the project provides both essential infrastructure for current and emerging social science researchers, and community access to datasets which relate to them.

The ATSIDA team works closely with:

- *experienced researchers submitting their 'life's work' to the archive*
- *emerging researchers developing data management plans*
- *Indigenous communities seeking digital repatriation of data relating to them.*

ATSIDA provides a trusted digital archive for social sciences data relating to the Indigenous peoples of Australia through the secure preservation of data in the Australian National University Supercomputer Facility (ANUSF) as a thematic archive within the Australian Social Sciences Data Archive.

ATSIDA develops data literacy skills in early career researchers, through the development of reasonable and comprehensive data management planning strategies.

And importantly, ATSIDA ensures that the Indigenous communities informing the research gain access to knowledge provided by their ancestors and peers.

This paper demonstrates how the project provides sustainability across all levels of the researcher and the researched, as well as developing information literacy skills in both early career researchers and the Indigenous communities informing the research.

Introducing ATSIDA

UTS was invited to establish a national trusted archive for Indigenous data in 2008. This 'node' of the Australian Data Archive (ADA) manages the collection, storage and reuse of Aboriginal and Torres Strait Islander research data. The research data comprises everything from SPSS numerical data through to bark paintings, song recordings and all the contextual information that gives the data meaning. The ATSIDA project draws on the expertise jointly developed by Jumbunna Indigenous House of Learning and the UTS Library.

As part of the development of the archive and in the ongoing dialogue around trust, security, protocols and cooperation, ATSIDA has entered into an agreement with the Australian Institute for Aboriginal and Torres Strait Islander Studies (AIATSIS), renowned as the world's premier institution for information and research about the cultures and lifestyles of aboriginal and Torres Strait Islander peoples, past and present. ATSIDA and AIATSIS staff will collaborate in a variety of areas, including digitization, data capture, repatriation and influencing regulatory and funding bodies.

ATSIDA provides a trusted, secure digital repository for the long term preservation of and access to research about Indigenous Australians. Only a small fragment of data produced by researchers has been available to other researchers, students, policymakers and the communities who have actively participated in research projects. Much of the material has tended to remain in researchers' offices, often eventually lost, rendered unusable due to format obsolescence or destroyed. There are obvious benefits to be gained from preserving and making more widely available the data assembled during research projects that are usually expensive in both cost and the time of researchers and Indigenous Australian participants. In addition, Indigenous Australian participants – as with many in other fields of research – now expect to play a significant role in directing the course of research projects and also to be provided with copies of research products. Digital repatriation of research data is an important component of ATSIDA and the team is committed to working with Aboriginal and Torres Strait Islander communities to create meaningful relationships built on trust and respect.

Aboriginal and Torres Strait Islander Peoples

ATSIDA deals specifically with research data concerning Aboriginal and Torres Strait Islander peoples, Australia's Indigenous peoples. The Australian Indigenous population form about 2.3% of the total national population, some 460,000 people [1]. Although this is quite a small population spread across the continent, it is very varied comprising urban, rural and remote communities with many cultural and linguistic differences.

In common with the experience of other Indigenous peoples, Aboriginal and Torres Strait Islander Peoples have suffered in many ways through the processes of colonialisation including dispossession of land, suppression of language and culture, economic disadvantage and concomitant health and social consequences. To take one example: approximately 250 distinct Aboriginal languages were spoken in Australia at the time of

colonialisation; two-thirds are now extinct or spoken by just a handful of elderly speakers and only about 20 are actively transmitted to children and widely used [2].

Aboriginal and Torres Strait Islander people have endured a history of surveillance and control through forced removals from lands, government control and intervention. In the past, derogatory terms such as “full-blood”, “half-caste” and “octoroon” [3] were used to categorise and define what it meant to be an Aboriginal person. Today the *Aboriginal Land Rights Act, 1983* provides a more contemporary definition of Aboriginality [4]

The definition is however rather rigid compared to Aboriginal people’s connection to their ‘country’ or area. A recent introduction provided by Professor Stephen Kemmis, Charles Sturt University, Wiradjuri Nations Language and Cultural Heritage Recovery project discussed what it means to be an Aboriginal person:

...As I understand it, to identify as an Aboriginal person means something more than the colour of your skin or how you label yourself. I know that in Australian law, being Aboriginal means (1) identifying as an Aboriginal person, (2) being recognised by other Aboriginal people as an Aboriginal person, and (3) having some Aboriginal forbears. That is the legal definition, and it is a strong beginning. But I understand that to have an Aboriginal identity also means something more. It also means being part of one of the many different Aboriginal cultures found in different parts of Australia today. I also understand that being Aboriginal means having some kind of living connection to country, whether in the form of a timeless connection in the place where your people have always lived, or a current connection with the country in the place you live right now. And I also understand it means having deep respect for country, for elders, for language and culture, for community and family, and for people (especially other Aboriginal people).” [5]

Research data deposited into ATSIDA will provide valuable resources for Aboriginal and Torres Strait Islander people in connecting with cultural heritage information that relates to themselves, their families and communities.

The Challenges

This project is ambitious for many reasons. It is tightly defined in that it deals with research data concerning Aboriginal and Torres Strait Islander peoples. The historical background and nature of the research make the work of ATSIDA a good test for many of the issues raised in relation to archiving social sciences research data. The ATSIDA team are working through these challenges to provide a rich service to researchers and to Aboriginal and Torres Strait Islander communities.

Datasets in various locations and formats

One challenge lies in the complexity that results from the dispersion of resources across many collections, agencies and formats. A lot of relevant data is located in datasets which are not focused solely on Indigenous peoples and issues. Examples include census, health, ethnographic, criminological, family and housing datasets. These data are joined with a host

of linguistic, ethnographical, musicological, community and other studies to form a rich but highly scattered resource of value not only to researchers but also to Indigenous Australian peoples themselves. It is challenging to bring such datasets in many formats together in a way that demonstrates the integrated and interrelated relationships.

Privileged nature of the data

The second major challenge lies in the privileged nature of much of the data. Both qualitative and quantitative datasets often concern small communities and sensitive issues. Their premature or injudicious release could breach privacy and confidentiality commitments, expose vulnerable people, or possibly invite misuse. The history of dispossession of Aboriginal and Torres Strait Islander communities underlines the need to handle research data extremely sensitively.

Reluctant depositors

The third major challenge is to engage researchers in this project as depositors to the data archive. Many researchers are reluctant to deposit datasets for a variety of reasons. In the course of their research, they develop strong relationships with the communities and individuals who are involved with and inform the research. Researchers are concerned to honour commitments regarding confidentiality and the uses to which the data might be put. They also want to ensure that the original project design and research ethics are maintained. Ethics and informed consent forms can often prohibit further use of data, making it challenging to deposit into a data archive. Conversely, they may be worried about possible exposure to methodological criticism or potential misuse of the research data. Misuse can involve inappropriate, acontextual use or the application of the data to attacks on authenticity and sometimes hostile purposes.

An example of this is in the 'History Wars' which scarred Indigenous studies in Australia through questioning the data employed in a reconsideration of Aboriginal-settler relationships and hostilities during the colonial period [6].

Besides these specific professional concerns, potential depositors may have doubts about long term sustainability and security of the archive and can be suspicious of the effort that might be required to deposit their datasets.

Western Science vs Indigenous Traditional Knowledge

Framing these challenges are far reaching questions about ontology and the ownership of knowledge. Research data archives such as ADA operate within the Western scientific paradigm and the Western conception of intellectual property. Both of these assumptions underlying data archives are problematic when dealing with Indigenous Australian datasets.

The western scientific paradigm posits neutral, objective and testable data, data which should be able to be reused to replicate research or feed new investigations. Data compiled in the course of investigations are owned by the individual researcher even though they

recognise some obligations to informants as is normal in the social sciences. The data should be able to be replicated and should be available for reuse by the same or other researchers, ethical protocols permitting. The 'openness' to examination, replication and verification implicit in the scientific paradigm can place it in direction opposition to a privileged understanding of knowledge in which some aspects can only be shared under specific rules. These rules are often expressed in terms of secret and sacred knowledge but extend to the privacy of individuals and their families and communities.

The Berne Convention and the work of the World Intellectual Property Organisation have developed regulations which aim to protect "literary and artistic works ... fixed in some material form ... for the benefit of the author and his successors" [7]. They do not protect ideas, only their expression 'in some material form', and they presuppose ownership by an author or creator who may sell, assign or give the rights to others during the period of protection. They do not recognise the community ownership of traditional motifs, iconography or ideas, which is characteristic of Indigenous and traditional knowledge systems. Nor do those regulations accept enduring ownership that never expires and is administered through community mechanisms such as the authority of elders to prevent misuse.

The needs of the western scientists and the Indigenous peoples poses serious questions to us when we are trying to store, preserve and make available research data relating to Indigenous peoples. ATSIDA archivists face questions like: Who has authority to deposit the datasets? How should they be described? Who can authorise access and reuse? How should the perspectives and rights of data compiler (researcher) and knowledge owner (informant) be resolved? Should the subjects of the research data be informed of the existence of the record, and consulted about its use?

These challenges are not solely relevant to Indigenous research data. Similar issues arise elsewhere and are symptomatic of qualitative research methods. Sensitivity, for example, is essential when handling research data concerning sexual abuse, HIV positive communities or many other delicate topics. Likewise, most researchers feel an obligation to serve their informants correctly and few would lightly transgress cultural constraints. But the challenges acquire particular immediacy in the Indigenous domain because of epistemic differences and the history of oppression.

Finding Solutions

Finding ways to resolve the challenges is central to the ATSIDA project. It starts by accepting complexity and recognising the crucial importance of context.

Complexity arises from the need to accept that there are conflicting imperatives that can only be negotiated on the basis of principles. The principles underlying the development of ATSIDA are aligned to those expressed by a World Intellectual Property Organisation Committee, in its paper on general guiding principles which aim "to ensure that the specific substantive provisions concerning protection are equitable, balanced, effective and consistent". [8] Especially relevant to ATSIDA are the:

- principle of responsiveness to the needs and expectations of traditional knowledge holders;
- principle of recognition of rights;
- principle of respect for customary use and transmission of traditional knowledge;
- principle of providing assistance to address the needs of traditional knowledge holders.

In taking a similar approach to these principles, ATSIDA insists on consultation in an appropriate manner with relevant Indigenous communities to ensure practices recognise and respect Indigenous practices and laws. This consultation, and engagement with communities, is an essential part of ATSIDA developing trust in the operation of the archive.

It is necessary for ATSIDA to adopt protocols and standards that will translate this avowal of respect and responsiveness into practice.

But the complexity is not simply related to western versus indigenous practices. There is also great diversity among the Indigenous peoples of Australia. This is manifested in the variety of traditional beliefs and practices which have been maintained. It is evident also in the lives and aspirations of individuals and communities in remote, rural and urban settings.

Consequently, it is important to ensure that the datasets may be brought together, logically, by clan or linguistic group as well as geographically and thematically. Such collocation of datasets facilitates consultation with relevant Indigenous Australian communities and sensitive management as far as possible in accordance with community preferences and traditions.

Besides the dimensions of traditional knowledge and community wishes, context also means depositors must be confident that the datasets will be properly safeguarded, preserved and made available only in accordance with the conditions imposed at the time of deposit. It is also important to ensure there is enough contextual information available to enable a secondary user to clarify the nature of the datasets and the ways in which they might be accessed and used. The researchers who access and reuse data must be clear about the expectations placed upon them and able to honour those expectations. All of the dimensions which might condition access and reuse need to be captured to ensure that the data archive can operate effectively and faithfully translate ethical expectations into practice.

Creating the Data Archive

The ATSIDA project has many requirements in common with other initiatives to archive and curate research data, especially those addressing the needs of researchers in the disciplines of the social sciences and humanities. The particular force the expectations and requirements have in the context of Indigenous peoples and knowledge, means the project to create ATSIDA offers a proving ground for methods of providing a robust eScience infrastructure in the social sciences and humanities.

The primary tasks are to translate the requirements into effective data management system methodologies. For ATSIDA, this is built on the facilities provided by colleagues at the

Australian National University Supercomputer Facility (ANUSF). Provided that it is supported financially by the national research funding agencies in the long term, it will provide the national cyber infrastructure to support the work of the emerging research data archives in many fields.

Locating ATSIDA's data storage in a national supercomputer facility clearly signals that the data archive is a nationally significant eScience initiative.

The standards for metadata and the protocols for preservation, access and use must reflect the principles of respect and reciprocity and must drive the authorisation mechanisms to ensure that the requirements and expectations are respected.

Mechanisms must also be employed to enable consultation with relevant Indigenous communities on a continuing basis. The protocols and their encoding must be sustainable in the long term.

Repatriation, a specific commitment

A particular and unusual commitment we have at ATSIDA is in facilitating the 'repatriation' of the information and knowledge recorded in the research datasets to the peoples to which it relates.

The term 'repatriation', has come to be used for the return of cultural objects, and sometimes human remains, to the nations, communities or peoples from which they were obtained whether legally or not. Repatriation of information and knowledge has great cultural importance but seldom involves the delivery of actual artefacts. It concerns the delivery in appropriate formats of copies of images, recordings, notes, observations and other records of the culture of a people.

The materials can then be managed appropriately by the community and can be used to revitalise particular cultural practices and improve social cohesion. Community members may add to them, correct them or amplify them as a living cultural memory bank.

ATSIDA is developing a website that provides an information sharing portal for the communities that have contributed to ATSIDA, as well as interested people from the general public where open access has been negotiated. While this is centrally administered, the material contained on the site is curated in negotiation with the communities to which it pertains.

An example of this in operation is the work ATSIDA is undertaking in collaboration with the Macleay Museum at the University of Sydney and the Buku-Larrnggay Mulka Centre at Yirrkala in Arnhem Land. This is a significant repatriation project which will be of importance to contemporary Yolŋu people. The project aims to digitise and make available a collection of approximately 100 paintings on bark and related documentation. The bark paintings were created in the late 1940s by Yolŋu artists for Australian ethnologist, Ronald Berndt. They are intrinsically interesting and valuable but gain additional research value through the inclusion of Berndt's field notes and of voice recordings describing the significance of the paintings.

The recordings were made in the 1970s by a Yolŋu elder, Wandjuk Marika. The research dataset will consist of digital photographs of the bark paintings together with the digitised documents and recordings and general contextual material, including correspondence and personal files. In addition to the digitisation and creation of descriptive metadata, the project requires interpretation of Berndt's specifications for access and display and consultation with Yolŋu elders to establish currently acceptable protocols. These will provide the parameters for making available the materials to researchers via ATSIDA, to researchers and the general public at the Macleay Museum and through repatriation to Yolŋu people.

The Aboriginal and Torres Strait Islander communities who engage with the website will have the opportunity to contribute their own knowledge to the data contained on the site. For example, the Yirrkala data may gather additional contemporary meaning through further annotation and interpretation by community members.

Another project is a compilation of photographs, media coverage and interviews with community members involved in the Koori Knockout – a rugby league competition for both men and women, which has been hotly contested throughout the state of New South Wales since 1981 [9]. It is anticipated that making available this research data will encourage others to put forward their own stories and photographs relating to a competition described on the National Indigenous TV site as a carnival that “can be likened to standing in the firing line on behalf of your mob – you are willing to risk it all for the honour and privilege to play for your home-town team, with a chance of being crowned the Champions.” [10]

This of course has implications for the design of the ATSIDA website. It must be accessible to remote communities who have intermittent or non-broadband Internet access. Options for tagging, annotating or otherwise contributing content must be simple and straightforward. At the same time, the website aims to be professional, visually appealing and employ the latest innovations in open source, Web 2.0 technology.

Developing the Data Capability of Early Career Researchers

A key component to the success of the project is in developing the data management skills of early career researchers. This must be done in a way that enhances the capacity of the researcher, enables good practice and does not add to the time taken to fulfil administrative requirements, adding to what are often unreasonable workloads.

It is not necessary for researchers to understand all the principles behind long term preservation of data, server backups, synchronisation or file format migration – these and other technical aspects should be managed by data archivists and programmers. Researchers in Australia do, however, need to understand their obligations to the funding bodies and the public who enabled their research, as well as to their informants and their peers. They must also be cognisant of their responsibilities to comply with the Australian Code for the Responsible Conduct of Research [11].

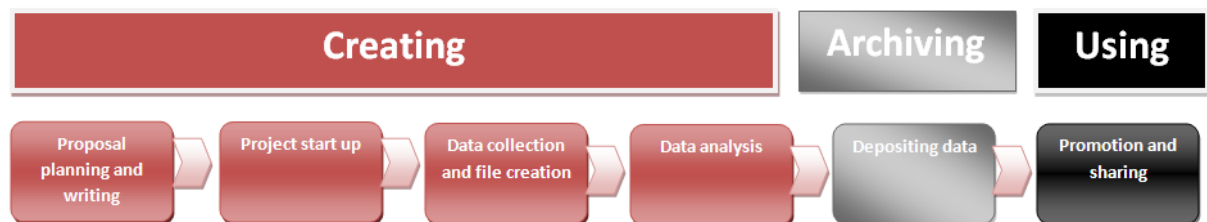
“Policies are required that address the ownership of research materials and data, their storage, their retention beyond the end of the project, and appropriate access to them by the research community.

The responsible conduct of research includes the proper management and retention of the research data. Retaining the research data is important because it may be all that remains of the research work at the end of the project. While it may not be practical to keep all the primary material (such as ore, biological material, questionnaires or recordings), durable records derived from them (such as assays, test results, transcripts, and laboratory and field notes) must be retained and accessible.

The researcher must decide which data and materials should be retained, although in some cases this is determined by law, funding agency, publisher or by convention in the discipline.

The central aim is that sufficient materials and data are retained to justify the outcomes of the research and to defend them if they are challenged. The potential value of the material for further research should also be considered, particularly where the research would be difficult or impossible to repeat.”⁶

Researchers need to understand their obligations for data management throughout the research cycle:



At the *proposal and planning* stage, it is important to guide researchers through tasks such as reviewing existing datasets, identifying any potential challenges for data and factoring in costs associated with data collection and management.

During the *data collection and file creation* stage, researchers may need assistance securing appropriate informed consent, establishing anonymisation techniques and determining appropriate file formats for archiving.

When *depositing data* with an archive such as ATSIDA, there are considerations for determining the suitability of the data for long term preservation and setting appropriate access conditions.

In conjunction with the ADA team, we are developing tools and checklists to guide researchers through these aspects of their research cycle. These tools will enable researchers to quickly and easily evaluate their responsibilities at any given point during the research cycle and facilitate good data management practices without imposing an additional administrative burden.

Conclusion

The project to create a national Australian Indigenous research data archive builds on the trust developed between researchers and Indigenous communities. It is an ambitious project because it must traverse some major challenges with consequences for the technological solutions it adopts as well as its management and operation. In addressing those challenges and taking on the goal of repatriating information where possible, it has the potential to build stronger relationships and to contribute to cultural strengthening. The protocols and standards being developed and tested and the technological means adopted to implement them will inform data archiving in other domains. The tools being developed for both the Aboriginal and Torres Strait Islander communities and the researchers working with these communities, will develop solid data management skills and first hand involvement in the evolution of the Aboriginal and Torres Strait Islander Data Archive.

REFERENCES

- [1] *Population Characteristics, Aboriginal and Torres Strait Islander Peoples 2001*, ABS 4713.0, Canberra, Australian Bureau of Statistics, p. 12.
- [2] *The Encyclopaedia of Aboriginal Australia*, Australian Institute of Aboriginal and Torres Strait Islander Studies, Canberra, 1994, p. 600.
- [3] <http://www1.aiatsis.gov.au/exhibitions/removeprotect/index.html>
- [4] http://www.austlii.edu.au/au/legis/nsw/consol_act/alra1983201/s4.html
- [5] Kemmis, Stephen (2009) Occasional Address – Narranderra TAFE Graduation
- [6] Macintyre, Stuart and Clark, Anna (2003). *The History Wars*, Carlton, MUP.
- [7] *Berne Convention for the Protection of Literary and Artistic Works* (1886-1979), Article 2, Geneva, WIPO, http://www.wipo.int/treaties/en/ip/berne/trtdocs_wo001.html
- [8] IBID
- [9] Norman, H, (1992) 'An unwanted corroboree: The politics of the New South Wales Aboriginal Rugby league knockout', in *Australian Aboriginal Studies*, 2009/2].
- [10] http://nitv.org.au/index.php?option=com_content&view=article&id=507:fierce-footy&catid=18&Itemid=41
- [11] <http://www.nhmrc.gov.au/publications/synopses/r39syn.htm>