5

Shoehorning complex metadata in the Living Archive of Aboriginal Languages

Catherine Bow, Michael Christie and Brian Devlin

The Living Archive of Aboriginal Languages is making endangered literature in Australian Indigenous languages publicly available online (Bow et al. 2014). Like any other project attempting to package a vastly complex body of work into an accessible repository, this project has grappled with a number of complex issues. Wrangling a variety of text types, languages, locations, digitisation processes, metadata and other issues into an accessible online repository requires a great deal of shoehorning. While straightforward decisions can be made simply, a number of decisions require complex solutions, and are readily dumped into an expanding 'too-hard basket'. Accordingly this paper – informed by Christie's work on Aboriginal knowledge traditions and digital technologies (2004, 2005), and picking up on many of the issues identified in Nakata (2007) on Indigenous digital collections – explores some of the 'too-hard basket' issues emerging from the Living Archive project, and various attempts to resolve them. This is not to suggest that the solutions reached within this project should be seen as normative or appropriate for other projects, but rather to explore and describe some

of the decision-making processes and the impact of these solutions and compromises on the overall project.

The initial aim of the Living Archive project was to collect materials created in Australian Aboriginal languages of the Northern Territory for schools with bilingual education programs, and to create digital versions of these materials for both preservation and access (Christie et al. 2014). Copies of the books were sourced, collected (from the schools themselves, or from libraries or private collections), catalogued, and digitised in both preservation and presentation formats. With the permission of the original creators, both the metadata and the digital objects were uploaded to a digital repository at the Charles Darwin University Library. A web-based interface was developed to facilitate open access to the materials, with a restrictive Creative Commons licence.[1] The data was stored in a MySQL database, with the metadata catalogued according to library standards, using both Metadata Object Description Schema (MODS) and Open Language Archives Community (OLAC)[2] schemas. These standardisation practices and recommendations serve as the basis for many of the decisions made in the establishment and structure of the Living Archive, in order to conform to best current practice for the digital archiving of language resources so as to facilitate discoverability and interoperability (Bird and Simons 2003).

The focus on published texts is a key point of difference between the Living Archive and a number of other related archives, such as PARADISEC (Thieberger and Barwick 2012), the collection at AIATSIS[3] and others listed at http://www.language-archives.org/. While the addition of audio, video and other materials would enhance the archive significantly, uncertainty about ongoing funding limits these options. Despite the impression that the focus on published texts should make the creation of metadata simple and unambiguous (cf. Nathan 2013),

---

1   Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Australia: https://creativecommons.org/licenses/by-nc-nd/3.0/au.

2   The Metadata Object Description Schema is an XML-based bibliographic schema developed for library applications. The Open Language Archives Community is an international partnership of institutions and individuals creating a global virtual library of language resources.

3   http://www.aiatsis.gov.au.

this paper demonstrates some of the complexities involved in creating useful metadata. As the archive is designed as both a research tool and a community resource, good metadata is vital to accessibility. Engaging different kinds of users in appropriate and interesting ways (e.g., Woodbury 2014; Trilsbeek and König 2014) requires that the contents of the archive be easily discoverable, searchable and navigable. Mediation (Holton 2014) may be required to assist users to navigate the complexity of the materials, so that users do not feel 'lost in a thicket' (Woodbury 2014, 22). The act of inviting community members to participate in the collection and correction of metadata of existing resources (Garrett 2014) is an important form of engagement that assists in bringing the archive to life.

While the complexity of the collection and its 'standardisation' into a usable archive inevitably leads to disagreement and error, it also allows for – and even requires – engagement with local authorities (Linn 2014). Community leaders and language authorities are well placed to review the information available and make decisions about metadata, categorisations, inclusion or exclusion of items and other key components of the archive. An important feature of this process has been negotiating the conflicting demands of standardisation with the often heterogeneous nature of the materials while taking account of the requirements of the various users. While sharing the load of such activities as forms of 'crowdsourcing' (cf. Birch 2013; Bird 2013) has the potential to capture the utility of digital distribution and the power of social networking, it is important not to neglect the value of sitting with community leaders 'offline' to perform such tasks. Further development of the archive will involve consultation with community members and this is expected to challenge some of the requirements of the standardised system, and to open up a number of interesting empirical and theoretical questions.

Three clusters of issues are raised in this paper: (1) naming or identification conventions (as applied to languages, places, people and book titles); (2) categorisation practices (the use of controlled vocabularies, grouping into communities and linking related items); and (3) determining whether to include or exclude materials (including photo books, and annotated versions).

Naming/identification

Languages

Language identification should ordinarily be a straightforward issue, especially since the majority of books in the collection clearly name the language used. In some cases, the language could be unambiguously assumed given the item's publication site; for example, all the books published on Bathurst Island are in the Tiwi language, while Yipirinya School published books in four different languages, each of which is clearly identified in the metadata. The wider collection raised a few challenges in this area.

The current recommended international standard for language identification (ISO 639-3) allocates a three-letter code to all languages listed in its database,[4] in order to assist with consistent language identification and to facilitate discoverability of resources. While such an international standard is useful, mapping the languages in this archive to this specification can be quite complex, as the codes did not always match the nomenclature used in the communities or in the books themselves. Where there was a direct mapping from language identification to language code, these were used (e.g., the code [mph] identifies the Maung language); however, in many cases the language name is listed as an alternative name or dialect in ISO 639-3 (e.g., Wubuy, discussed below).

A particularly complex case is in the Yolŋu area of north-east Arnhem Land (Christie 1993), where the different levels of clan affiliations and moiety distinctions are not incorporated into the ISO 639-3 codes. For example, Yolŋu people often classify their languages by the word used for 'this', which yields a set of eight related language groups: Dhuwal, Dhuwala, Djaŋu, Djinba, Djinaŋ, Dhaŋu, Dhay'yi, and Nhaŋu (Schebeck, 2001). Each language contains a pair of corresponding dialects (*matha*), which are categorised by moiety and linked to clans (*mala*). So people of the Djambarrpuyŋu *mala* speak a version of Dhuwal which is often referred to by their clan name, Djambarrpuyŋu, while the closely related Gupapuyŋu people speak Dhuwala (although

---

4   http://www-01.sil.org/iso639-3/codes.asp.

the books are listed as being in Gupapuyŋu language). The use of clan names to differentiate 'dialects' is not incorporated in the ISO 639-3 system. A book in the Wangurri language should ideally be identified as Wangurri, Dhaŋu and Yolŋu, but this is currently not permissible using the ISO 639-3 system.

The ISO 639-3 codes are also subject to change, which needs to be monitored; for example, in 2012 a request to change the name of the Dhangu language [dhg] to Djangu (with the same three-letter code) was accepted by the ISO,[5] requiring a change in metadata for the Living Archive. In 2013, a further request resulted in that change being reversed, reflecting the distinction between the Dhangu and Djangu languages. Consequently the ISO 639-3 situation needs to be regularly monitored to identify and process any further changes to the system.

A further example of the complexity of language nomenclature involves the book *Jatdi Na-yahwurt* ('The little frog', Galmur 1994) (Figure 5.1). The cover indicates that the book is in Mayali language; however, ISO 639-3 has no code for Mayali, listing it only as an alternative name for Gunwinggu [gup], along with Kuninjku, Kunwinjku and Gunwinjgu.[6] Garde (2014, personal communication) states that Mayali is a variety that is distinct from Kunwinjku, Kuninjku and Gundjeihmi, and notes that Evans (2003) recommends the term 'Bininj Gun-wok' as a collective name for this dialect chain. Other books in the Living Archive collection are identified as Kuninjku and Kunwinjku, but none use the ISO 639-3 name of Gunwinggu. A compromise was required for the Living Archive in order to conform to ISO 639-3 codes for discoverability, while still respecting community usage and the printed identification of the language by the book's creators. This was achieved by creating an additional metadata field labelled 'Language note' to allow for clarification of otherwise potentially confusing information. In the case of this book, the metadata displays the language as Kuninjku (linked in the background to ISO code [gup]) and the language note specifies 'Mayali language'.[7] While such a solution is less than satisfac-

5   http://www-01.sil.org/iso639-3/chg_detail.asp?id=2012-047&lang=dhg. Here the ISO 639-3 spellings are used without special characters.

6   http://www.ethnologue.com/language/gup.

7   http://laal.cdu.edu.au/record/cdu:31258/info.

tory, it aims to meet both the demands of standardisation and the local preferences of the community.

Another means of managing these discrepancies was the establishment of a synonym list which linked the ISO 639-3 codes (and AUSTLANG codes,[8] another system commonly used for identification of Australian languages, though not compliant with OLAC standards) with the language names as used in the Living Archive, and allowing discoverability via both the official ISO names or common alternative names. For example, the Numbulwar community calls their language Wubuy, while ISO 639-3 and many other sources retain the label Nunggubuyu. Simply listing Nunggubuyu in accordance with ISO 639-3 standards may prevent people finding Wubuy materials, or alienate users more familiar with the name Wubuy. The synonym list allows searches for both Nunggubuyu and Wubuy to return the same results, and while Wubuy is retained as the preferred language name in the archive, it is directly connected to Nunggubuyu as the ISO standard name. In some cases selecting the language names was uncontroversial (e.g., Warlpiri, Tiwi) but in many it was necessary to shoehorn complex information into a simplified structure.

Places

While the majority of items were clearly identified as coming from a specific Literature Production Centre based within a school, the naming of place was not always straightforward. Over the four decades of literature production, the official names or spelling of some locations changed. Some places had changed name (e.g., Oenpelli is now known as Gunbalanya), and some are known by both a Western and an Indigenous name (e.g., Docker River is also known as Kaltukatjara). The use of the synonym list mentioned above for language identification included alternative names (or spellings) of place names to return appropriate search results.

In addition, OLAC conventions allow for the inclusion of location information beyond place of publication, such as 'geographical origin'

---

8   http://austlang.aiatsis.gov.au/main.php.

(used for the setting of a story – for example, the story *Kapirdi-langu-patu* (Martin 2011), published at Yuendumu in Warlpiri language, is based on a Dreaming belonging to the Gundungurra people from the Blue Mountains in NSW) and 'origin of story' (used if a story told in the book originated from a location different from the one indicated by the language or location of the publication). An example of this is the Maung story of a mother turtle (Kurrunama and Margalgala no date), which lists Warruwi as the origin of the story, was translated into Burarra language in Maningrida and into Djambarrpuyŋu language in Galiwin'ku. These details are particularly useful for an audience which is consciously aligned to place; however, for non-Indigenous staff entering data in the Living Archive database, it was sometimes difficult to clearly identify and categorise these additional locations as either geographical origin or story origin. This gives further opportunity for local users to validate and enhance the metadata of these objects, especially where some of the metadata fields are optional or not relevant, while more detail allows for greater enrichment and complex searchability.

People

Identification of people can become complex given the use of different Aboriginal naming practices (Christie 1993), such as changing names due to marriage or death, different spelling conventions, the use of both Indigenous and non-Indigenous names, each person having a number of different personal names, and the emergence of surnames in some places. The challenge is to link records to a single person, if the name appears differently in different books. For example, one author has books attributed to her by a number of different names and alternative spellings, including Ampi, Margaret Ampi, Margaret Umpi, Margaret Ampi Poulson, and Margaret Poulson. The infrastructure on which the Living Archive is built allows for authors (and other contributors) to be allocated a unique four-digit code. Using this, each record displays the name of the contributor as it is listed in the book metadata, and is also linked to any records of the same contributor identified by any alternative name. All books by the above author can be viewed together by clicking on any variant of her name, while the integrity of the metadata for each individual record is maintained.

Nonetheless, community knowledge is required to identify these connections and correct any errors. For example, in the Gunbalanya collection, two names were given separate codes until a local contact identified the two names as referring to the same person. Such errors are often identified, and there are undoubtedly additional errors like this in the existing archive, exposing another area where local input is needed to identify problems and correct errors. A strategy is being developed to allow and enable community members and appropriate stakeholders to identify and rectify these situations.

Different community practices raise different issues with regard to attribution of authorship. A series of Maningrida readers includes no mention of authors, yet local people know exactly who wrote them. This information can then be supplemented in the metadata in the archive using square brackets. Also in Numbulwar, a conscious decision was made at the time many books were produced not to include the names of authors, so most items are listed as written or illustrated by 'Numbulwar Community Education Centre'. Since the project team is asking creators for permission to make the books public (Bow et al. 2014), this creates a problem, and identification of authors risks negating the original reason for their exclusion. Such questions can best be answered at the local level, yet involve wider legal and ethical implications.

Titles

While the identification of a book's title is generally unproblematic, some of the materials in the Living Archive do not conform to standard practice in this area. There are a small number of books which have no clear 'title' on either the cover or a title page, and even some books with no words, for which titles had to be devised. In other cases the distinction between a title, a subtitle or a series title was unclear. In a few cases there are discrepancies between the title as it appears on the title page and on the cover, such as 'Kukaku Anu' which on the cover is listed as 'Kukaku Yanu' (Raggett 1979). Such problems are not unique to this collection, although they may be more difficult in this case as some of the materials are in languages not understood by the people developing the archive.

Figure 5.1 'The little frog' by Judy Galmur. http://laal.cdu.edu.au/record/cdu:31258/info/.

Categorisation

Library cataloguing is a carefully managed process, with strict protocols addressing a range of possible scenarios. The books produced in Literature Production Centres do not consistently conform to these standards, making it sometimes difficult to shoehorn them into the required categories, leaving many in a nebulous 'too-hard basket'. In the Living Archive this has led to additional metadata fields which may only be relevant for a small number of books each. While this enhances the granularity of the archive and allows for complex searching and sorting, it also requires careful decisions for data entry. Questions such as 'Should this be a description note or an abstract?', 'Should this be considered a series name or a subtitle?', 'How do we catalogue this book which has a different title on the cover and on the title page?' were regularly discussed within the project team and with a librarian with expertise in metadata. Where some questions were outside the standard metadata practice, new solutions had to be identified.
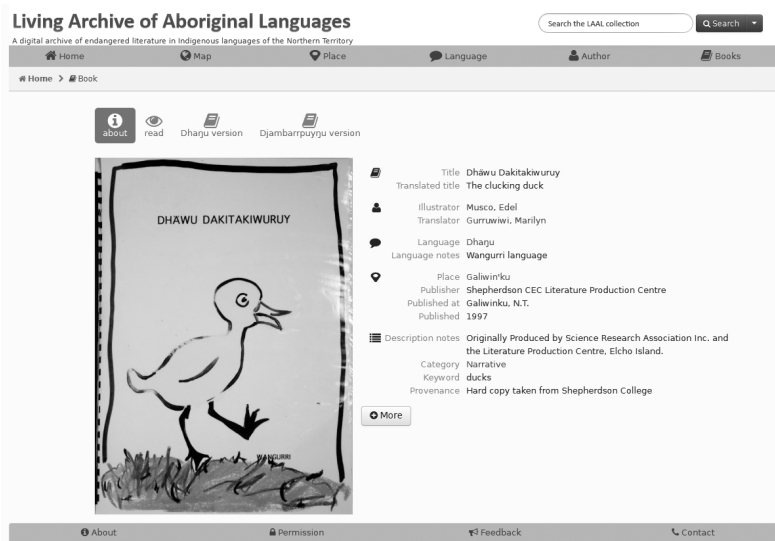
Figure 5.2 'The clucking duck'. http://laal.cdu.edu.au/record/cdu:31310/info/.

Controlled vocabularies

Use of controlled vocabularies is a standard feature of many database programs, forcing the use of a set of pre-defined terms, rather than free text. Limiting this range of options enforces a consistency which aids manageability – for example, avoiding spelling errors, and grouping like with like. In the Living Archive, controlled vocabularies are used in a number of metadata fields (category, type of resource, scan source, etc.), each with its own set of compromises, such as those involved in using controlled vocabularies for language names outlined above.

The use of a 'category' field assists end users to navigate through a huge range of materials. However, categorising Aboriginal literature into Western-style literary genres can create more problems than it solves, because it undermines traditional classifications of modalities (Christie 2005). Currently the archive lists seven genres or categories: traditional, instruction, language instruction, map, memoir, narrative and song. While it may be helpful to distinguish between a text of

language instruction and a memoir, or a song and a narrative, these distinctions are not always transparent, and often depend on the perspective of the person inputting the data. Putting things into such categories is useful for navigation, yet grossly oversimplifies the issue. Further discussion and community engagement is desirable to further explore appropriate categories, and these will most likely be quite different in different communities. Already community feedback has prompted a change, where what was once categorised as 'folktale' according to library standards was changed to 'traditional' (i.e., traditional story) in response to a request from community members who thought the word folktale devalued the significance of traditional stories.

Most of the materials in the Living Archive are books; however, a limited selection of additional related materials, such as audio files, videos, and other multimedia files (e.g., ebooks) is included in the archive. Using the controlled vocabulary of 'Type of resource' to classify all records using MODS categories Text, Sound, MovingImage and InteractiveResource allows for a search that simply retrieves the appropriate type. This expands the infrastructure of the archive to incorporate more materials as they may emerge.

Digitisation of materials in the Living Archive has been distributed across a number of different sites. Using a controlled vocabulary for the metadata field 'Scan source' allows easy access to data sets from these different locations, whether digitised at Charles Darwin University Library, by staff at Australian National University (which is a project partner), on scanners or multi-function printers at schools or Literature Production Centres as part of previous digitisation projects (such as at Barunga School), by Department of Education staff in Alice Springs (for materials from a number of desert communities), or those created digitally at Literature Production Centres, which required no scanning. This is an example of refining metadata fields to assist in project management as distinct from using metadata to facilitate access or discoverability.

Using these controlled vocabularies is an attempt to simplify the database from both a data-entry and an end-user perspective (i.e., allowing faceted browsing); however, this disguises much of the complexity behind each of these issues.

Grouping into communities

The initial grouping of materials was complex due to a lack of one-to-one mapping between a language and a place. Some languages were linked to several places (e.g., Warlpiri language materials came from Yuendumu, Lajamanu and Willowra; Kunwinjku materials from Gunbalanya, Maningrida and Barunga), while some places produced materials in several different languages (e.g., Maningrida produced a large number of Burarra and Ndjébbana materials, plus smaller numbers of Kunwinjku, Gurrogoni, Djinaŋ, etc.; Yipirinya School produced materials in Warlpiri, Luritja, Central Arrernte and Western Arrernte). In fact, one-to-one mapping was rare (only Nguiu for Tiwi language materials and Wadeye for Murrinh-Patha[9]). Even apparently simple cases had some degree of complexity (e.g., Warruwi Literacy Centre only produced Maung language materials, but these were mostly published in Maningrida, as Warruwi had no printing facilities). Even if one location produced materials in only one language, another location producing anything in the same language removed the possibility of any 1:1 mapping (e.g., Santa Teresa only produced Eastern Arrernte materials, but Yipirinya School also produced items in this language among others). From a project management perspective it was helpful to create functional groupings that allowed for materials to be located in a single 'collection', even if they overlapped with another group. As a result, 16 collections were formed, based on either a language or a location. This, however, is not transparent in the resulting online archive, as the browse options are currently limited to either language or place. This is yet another example of the careful thinking needed about how the archive would be used, and how this would be affected by decisions made at the project management level.

---

9   Materials from the Wadeye bilingual program are only in Murrinh-Patha language; however, materials in other languages are likely to emerge as the search expands beyond these school sources.

Related items

The Living Archive includes materials that connect with one another in various ways. There are examples of books which have translations into different languages; others appear in various versions (whether with varying layout, such as 'big books' and 'instant readers', or in different editions, such as updated or revised versions), and some which have related multimedia objects. In other cases the relationship is more along the lines of membership of a series. In order to facilitate discovery and access to these related items, additional fields were included in the metadata to enable links to be made visible. Two types of related items were explicitly addressed: those linking to multimedia formats and those linking to other versions. The first category allows for audio or video files to be uploaded and linked directly to the book to which they refer – for example, an audio file and an epub version of *Ŋalindiy bumara yolŋunha mala* ('The moon killed people', Djäwa 1975). The second category allows for the different 'versions' noted above, with translations distinguished by including a language field, while alternative versions are in the same language. An example is *Dhäwu Dakitaki-wuruy* (Figure 5.2) ('The clucking duck', literally 'A story about ducks', Unknown 1977) which has an original version in Djambarrpuyŋu language from 1977, linking to an updated A4 sized version from 1997, as well as translations in two varieties of what ISO 639-3 refers to as Dhaŋu language (Wangurri and the combined Gälpu/Golumala/Ŋaymil version). Other relationships are accessible through hyperlinking specific fields, such as contributor names and categories, as well as place and language.

Inclusion and exclusion

The wide range of materials in the archive meant that the project team needed to decide which items should be prioritised for inclusion, which should be excluded, and which left in the 'too-hard basket' – a receptacle noted for its ever-growing capacity. For example, at Yirrkala a series of workshops with traditional elders and school children exploring facets of culture and teaching led to the production of books. These 'Galtha Rom' books remain a valuable repository of traditional

knowledge and experience. The community had already been through a process of deciding which ones could be shared – made accessible in the community through the school library, for example – and which should be protected from view. However, these decisions needed to be reconsidered in light of internet technology, which has the potential to take the materials beyond the community to a much wider public, and so gives questions about accessibility broader significance. These books stand in a separate category as their authorship is complex, so it is almost impossible to find all those involved in their creation and request permission from everyone concerned or their families for the materials to go online. Decisions about whether such materials belong in an open access repository such as the Living Archive must be made by the community. A scanner was provided to the school to enable them to digitise materials themselves as required, ensuring preservation, but leaving open the question of access. In a number of situations, items may escape digitisation because appropriate permissions for such complex (and often quite politically charged) material may be simply too hard to get given the funding, time and personnel available.

Photographic books are another potential hotspot, and caveats about Indigenous people not displaying images of deceased people are well known. However in some communities these rules are loosening, and people enjoy seeing photos of relatives. Also photos of those who are still alive may be contentious, especially with certain sensitivities surrounding images of children. While the creation of the books incorporated parental permission for including photos of the children in various settings (often on school excursions or involved in school or community activities), those permissions did not explicitly include (or preclude) any publication of the images on the internet. Many of those children are now adults and may want a say in whether or not their images may be reproduced for the world to see. Similarly, many books in the collection are made up of stories and pictures produced by school children as part of a class activity. The effort to track down each child and ask permission for the material to go online is beyond the capacity of the current project. Economies of scale make it difficult to address each individual item in the collection with the appropriate authorities, and so it is the easy cases which make their way to the front of the line. Further development will enable community-level enrichment and en-

largement of the archive to provide an opportunity for some of these issues to be resolved authoritatively and decisively. An online feedback form is already available[10] and offline feedback is regularly sought, particularly with Indigenous community members.

The books in the archive were created for classroom and community use, and consequently some hard copies bear signs of wear and tear. While attempts were made to select the 'best' copy for digitisation, there are many examples of marked, defaced, and damaged artefacts in the collection, as well as some with careful annotations. In some cases it was possible to digitally enhance the original documents; however, there is still some merit in retaining the evidence of use, showing that the books have been handled and engaged with by various users. Annotated copies present a different challenge, particularly when it is not known who made the annotations (often spelling modifications, changes in diacritic use or punctuation) and therefore if they are to be accepted. At the current stage of development, the archive only allows a single plain text version to accompany the scanned PDF files. However, a solution whereby different textual variants can be included is desirable.

Conclusion

The materials created for bilingual education programs in remote Indigenous community schools over a number of years represent a wealth of knowledge, experience and skills. They belong to a specific context of time, place, people and situation, which cannot be retained when bringing digital versions to life in an online archive. The vastly different array of possibilities for dissemination and access to materials were unimaginable when the programs began, and so the resources take on a new identity in this environment. While Western knowledge systems force certain requirements on today's archiving standards, traditional knowledge systems should still be respected. In the Living Archive project, the development of the archive has seen compromises and solutions drawing on both traditions, yet sometimes satisfying nei-

---

10   http://laal.cdu.edu.au/feedback.

ther. Such is the nature of a project such as this one, in managing a wide range of material in various categories, with a number of cases not fitting neatly into established protocols. This 'too-hard basket' is where some of the most interesting and complex issues reside, and much is lost in the oversimplification (or shoehorning) of rich and complex data.

Works cited

Birch, Bruce (2013). 'The Ma! Project: Crowdsourcing software for language documentation.' Presented at Research, Records and Responsibility (RRR): Ten Years of the Pacific and Regional Archive for Digital Sources in Endangered Cultures, The University of Melbourne, 2–3 December. http://hdl.handle.net/2123/9858.

Bird, Steven and Gary Simons (2003). 'Seven dimensions of portability for language documentation and description.' *Language* 79(3): 557–82. http://tiny.cc/2003-7dim-birdsimons.

Bird, Steven, Florian Hanke and Haejoong Lee (2013). 'Collaborative language documentation with networked smartphones.' Presented at Research, Records and Responsibility (RRR): Ten Years of the Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC), The University of Melbourne, 2–3 December. http://hdl.handle.net/2123/9857.

Bow, Catherine, Michael Christie and Brian Devlin (2014). 'Developing a living archive of Aboriginal languages.' *Language Documentation and Conservation* 8: 345–60.

Christie, Michael (1993). 'Yolngu linguistics.' *Ngoonjook: A Journal of Australian Indigenous Issues* 8: 56–77.

Christie, Michael (2004). 'Computer databases and Aboriginal knowledge.' *Learning Communities: International Journal of Learning in Social Contexts*. 1 (February), https://www.cdu.edu.au/northern-institute/lcj/learning-communities-journal-archive.

Christie, Michael (2005). 'Aboriginal knowledge traditions in digital environments.' *The Australian Journal of Indigenous Education* 34: 61–66.

Christie, Michael, Catherine Bow, Brian Devlin and Jane Simpson (2014). *The Living Archive of Aboriginal Languages*. Darwin: The Northern Institute, Charles Darwin University.

'Dhäwu Dakitakiwuruy' (1997). Translated by Marilyn Ganyinurr. Galiwin'ku: Shepherdson CEC Literature Production Centre. http://laal.cdu.edu.au/record/cdu:31310/info.

Djäwa (1975). *Ŋalindiy Bumara Yolŋunha Mala*. Milingimbi: Milingimbi Bilingual Centre. http://laal.cdu.edu.au/record/cdu:31730/info.

Evans, Nicholas (2003). *Bininj Gun-wok: A Pan-dialectal Grammar of Mayali, Kunwinjku and Kune*. Canberra: Pacific Linguistics.

Galmur, Judy (1994). *Jatdi Na-yahwurt*. Barunga: Barunga Press. http://laal.cdu.edu.au/record/cdu:31258/info.

Garrett, Edward (2014). 'Participant-driven language archiving.' *Language Documentation and Description* 12: 68–84.

Holton, Gary (2014). 'Mediating language documentation.' In *Language Documentation and Description* 12: 37–52. London: SOAS.

Kurrunama, Rosemary and Maggie Margalgala (n.d.). *Inyarlgen*. Warruwi: Warruwi Literacy Centre. http://laal.cdu.edu.au/record/cdu:31684/info.

Linn, Mary S. (2014). 'Living archives: A community-based language archive model.' *Language Documentation and Description* 12: 53–67.

Martin, Margaret (2011). *Kapirdi-langu-jarra*. Translated by Shirley Martin Napanangka, Tess Ross Napaljarri and Serena Shannon Nakamarra. Yuendumu: Bilingual Resources Development Unit.

Nakata, Martin (2007). 'Indigenous digital collections.' *Australian Academic and Research Libraries Journal* 38(2): 99–110.

Nathan, David (2013). 'Access and accessibility at ELAR, a social networking archive for endangered languages documentation.' In *Oral Literature in the Digital Age: Archiving Orality and Connecting with Communities*, edited by Mark Turin, Claire Wheeler and Eleanor Wilkinson, 21–40. Cambridge: Open Book Publishers.

Raggett, Obed (1979). *Kukaku Anu*. Papunya: Papunya Literature Production Centre. http://laal.cdu.edu.au/record/cdu:31914/info.

Schebeck, Bernhard (2001). *Dialect and Social Groupings in Northeast Arnhem Land, Australia*. Munchen: Lincom Europa.

Thieberger, Nick and Linda Barwick (2012). 'Keeping records of language diversity in Melanesia: The Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC).' In *Language Documentation and Conservation, Special Publication No. 5: Melanesian Languages on the Edge of Asia: Challenges for the 21st century*, edited by Nicholas Evans and Marian Klamer, 239–53.

Trilsbeek, Paul and Alexander König (2014). 'Increasing the future usage of endangered language archives.' *Language Documentation and Description* 12: 151–63.

Woodbury, Anthony C. (2014). 'Archives and audiences: Towards making endangered language documentations people can read, use, understand, and admire.' *Language Documentation and Description* 12: 19–36.