

The ANU Archives' implementation of ICA-AtoM software

This paper documents a case study presented at a workshop for the 15th Pacific Regional Branch of the International Council on Archives (PARBICA) conference, Honiara, Solomon Islands held in August 2013. An abbreviated version was published in the PARBICA Panorama Annual 2014.

In 2010 the International Council on Archives released open-source web-based archives management software called ICA-AtoM (Access to Memory). It had been developed by Artefactual Systems Inc. for the ICA to make it easy for archival institutions worldwide to put their holdings online using the ICA's descriptive standards. The software can be downloaded at www.ica-atom.org.

The Australian National University Archives is one of the larger archives in Australia with over 20 kilometres of records in its custody. It collects business records and archives of trade unions and industry organisations in the Noel Butlin Archives Centre, and also holds the University Archives and the Pacific Research Archives.

Existing finding aids

With a diverse collection built up over sixty years, there had been a variety of approaches taken in documenting the Archives' holdings. Documentation about the records and their creators and detailed lists of items had been typed up and placed in binders to be consulted by researchers in the reading room. Subject index cards were also prepared to assist retrieval. A summary list of holdings was printed as a booklet and regularly revised, and was then accessible on the Archives' website at www.archives.anu.edu.au.



Existing finding aids: red binders and subject index

A stand-alone computer system was developed in the early 1990s using 4th Dimension software. It operated for a number of years and was used to generate paper documentation to add to the binders. Unfortunately the database was not supported by the Information Technology area and at the time it crashed staff who were familiar with its operation had either retired or moved on. Since then collection documentation had been produced as Word documents and Excel spreadsheet lists and added to the reading room binders.

The Archives' development of an image database of digitised photographs was a catalyst to develop a database for the whole collection. With many other archives digitising collections and making them accessible online, it was evident that the ANU Archives would need to make our finding aids accessible to researchers online, to provide a framework for further digitisation and to go some way to meeting our users' expectations of online research.

Searching for new software

Archives staff evaluated a number of software products, concentrating on those that would be supported by the IT area in the light of previous experience. The ICA-AtoM software which was available for beta testing in 2008 was attractive to the Archives because:

- the database would be online, accessible from our website and searchable by search engines such as Google
- it was based on the ICA descriptive standards and provided the opportunity to bring our documentation up to that standard
- as open source software it did not tie us to investment in proprietary software (or the cost of licences) and also would allow an exit strategy in the future as data could be exported easily to another system if needed
- it facilitated multiple repositories: our two major collections (the University Archives and the Noel Butlin Archives Centre) could be presented in one database and there would be the potential to join future collaborative databases
- it had a CSV (comma separated value) upload facility which would allow automatic loading of electronic lists
- support was offered in the form of documentation: a user manual and an administrator manual, and online user group discussion
- our IT area was prepared to support it because other systems at the University used the underlying open-source operating codes.

Getting ready

Some preparatory work had to be done to work out how we would transfer the information in our old finding aids to a database. An initial decision was that if item lists were part of the database, then the information on subject index cards (which derived from the item titles) would not need to be duplicated. The information in the binders which would need to be captured consisted of:

- **typed descriptions and lists:** options included converting these to electronic documents using OCR (optical character recognition) software then loading using the CSV facility or manually entering into the database
- **printouts from the old database:** a rescue mission was launched on the old Macintosh computer to retrieve the data that had been entered into the original database, producing a 40,000 line Excel spreadsheet which represented the 40,000 items which had been listed in the database. This was then divided up by deposit so that the items could be loaded using the CSV facility against the archival description for each deposit.
- **printouts of electronic documents:** electronic documents such as Excel spreadsheets could be massaged into the CSV template and automatically loaded.

Assessment of current documentation

The documentation of series in the University Archives was most consistent with the ICA standards as the accessioning of university material had only begun in earnest in 2001. There were only six binders of series descriptions and lists, and these had been created as Excel spreadsheets and could be loaded using the CSV facility. This was a good place to start to test out the software and become familiar with it.

The Noel Butlin Archives Centre collections, on the other hand, had been gathered over the past sixty years and the documentation varied from detailed page-by-page descriptions of large company collections to rough box lists of material that had flooded into the Archives when trade unions had amalgamated in the 1990s. The data that was recorded, the order in which it was recorded, and the level of detail had varied over the years. No standard format had been followed.

A 'record group' approach had been taken so that often records were documented as records of the company or union transferring them, when they were actually created by a predecessor organisation. Some information required by the ICA descriptive standards did not appear in the finding aids. Information about, for example, access conditions and archival history had been recorded in accession registers.

Early 'deposit descriptions' included information about the creator of the records and about the deposit in the same document. The ICA descriptive standard sets out that the 'creator' information should be separate from information about the records, so if these documents were successfully OCR'd we would have to cut-and-paste information to two separate entries on the database.

These issues meant that we decided that it was more cost-effective to enter the data about creators and deposits manually as staff would have to make judgements about where data would be recorded and to what level of detail, and also check sources such as the accession registers.

Developing local procedures and forms

But before we could start entering data, we needed to develop some local procedures and tools. The ICA descriptive standards cover many data elements so we decided which elements would be mandatory for our archives based on the information that had been routinely collected. We developed a procedure and simple entry forms which included only the elements we would be entering for creators (authority records) and deposits. The forms reminded staff about which elements they needed data for and they could record it without worrying about being 'timed out' on the database. Once they completed the form from the various sources they could enter the complete record.

The procedures also gave formats for the completion of some fields so that the documentation would be consistent even if entered by a number of different people. For example, 'Extent and medium' would record the quantity in metres to one decimal point, eg 0.3 m rather than 30 centimetres or 1 foot (early documentation was prepared before metric conversion).

We decided that we needed to highlight our Pacific-related collections in both the University Archives and the Noel Butlin Archives Centre. We added to our procedures that wherever Pacific material was located the subject tag 'Pacific

Research Archives' should be added to the archival description for the deposit or series.

We also developed a thesaurus of terms for the completion of the data element 'Functions, occupations and activities' in authority records. We had always recorded on the summary list of holdings the industries that companies were involved in and the occupations of people whose personal records we held but there was no consistency. A company might be described as 'pastoral' or 'agricultural' and a person might be described as a 'writer' or an 'author' or a 'trade unionist', a 'trade union activist', or a 'union official'. We based the industry classifications on the Australian New Zealand Standard Industrial Classification published by the Australian Bureau of Statistics and the occupation classifications on a National Library of Australia thesaurus. Our lists of terms are available on the 'database help' page: <http://archives.anu.edu.au/collections/finding-aids/database>.

We also devised 'data entry' spreadsheets to record when authority records, archival descriptions and item lists had been entered. For researchers using the binders in the reading room we added a front sheet to each binder to indicate what has already been entered in the database.

Staged data entry

With twenty kilometres of records in our custody and a limited budget, our strategy for entering data was always to break up large 'impossible' tasks into smaller projects and to allocate parts of the project to all staff. We had 4 to 6 staff working on the projects (while doing their other work) over a period of four years. We also had one part-time casual staff member for just over a year.

We would scope each project by choosing a few examples and entering the data. This often raised issues which needed to be covered in the procedures which were then added to and amended in the light of this experience.

The procedure for how to enter archival descriptions for items underwent many amendments as different examples arose. For instance, what to do if an item was loaned back to a depositor and never returned, and how to enter different formats such as microfilms, maps, photographs and audiovisual items (we produced a separate procedure for entering audiovisual items).

We started with the six binders of University Archives series, authority records and item lists. This project was shared by two staff members and substantially completed in about 12 months, with some large lists requiring further manipulation for automatic CSV loading.

We then tackled company records: to share the work equally, each staff member chose one large company (with 8 or more deposits) and a number of smaller companies (with fewer deposits) and prepared the authority records and archival descriptions for the deposits. Item lists were not entered at this stage unless there were only a few items to enter. This project involved about 80 binders and lasted almost a year. At the same time, a volunteer drafted short biographies for the 200+ people whose collections of personal papers we held. This was an easier task than preparing histories of sometimes obscure companies or organisations. The casual

staff member and a work-experience student entered these as authority records and then completed the archival descriptions for their deposits.

The records of trade unions and organisations were next: groups of trade unions were identified based largely on trade unions which exist today. For example, the Construction Forestry Mining and Energy Union (CFMEU) group included predecessor unions for coal miners, the furnishing trade, timber workers, builders' labourers, engine drivers and firemen, plasterers, carpenters and joiners, painters and decorators, stonemasons and slaters, tilers and shinglers. Again each staff member chose one large group and a few smaller groups and entered authority records and archival descriptions for deposits. As staff completed their allocated trade unions they were 'rewarded' with other organisations which needed to be done. The casual staff member completed all the organisations with only one deposit. This project involved more than 120 binders and took about seven months.

We then concentrated on item entry. While we could use the CSV facility to upload electronic lists for University Archives and more recent company and organisation deposits, including those which had been entered in the old database, there were still many typed lists to go into the database. By May 2013 we estimated that only 10% of items had been entered. To increase this percentage we identified smaller deposits (up to 2 metres) from their archival descriptions. Again each staff member was allocated a number of item lists to enter and a year later we are able to say that over 40% of item lists have been entered.

We have tested OCR software on typed lists but there is a high character error rate, due no doubt to the old typeface used. For the shorter lists OCR was not time-efficient; that is, it was quicker to enter short lists manually than to correct the OCR errors and then upload using the CSV facility. Optical character recognition may be the solution for the longer lists that we haven't yet entered. This project continues!

Once items have been entered it is possible to link those item descriptions to digitised content. Our digitised images are held in the University's digital repository, (digitalcollections.anu.edu.au) and we are now linking item descriptions in the database, where the item lists have been entered, to that content.

EXPLORE ANU » A-Z INDEX » Search ANU... WEB CONTACTS MAP

Australian National University Archives

ARCHIVES

- Home
- Noel Butlin Archives Centre »
- University Archives »
- Pacific Research Archives »
- National AIDS Archive Collection
- Treasures of the Archives
- Using the Archives »
- News & events »
- Publications
- Friends of the Noel Butlin Archives Centre
- Frequently asked questions
- Links
- Contact us

Certain Women at ANU exhibition

'Certain Women at ANU' celebrates ANU's first female PhD students, Professors, Research Fellows and senior officers. It also covers the ANU Women's Studies Program and the ANU Club for Women. Now on in the Archives reading room, R.G. Menzies Library.

» read more

Australian National University Archives

The ANU Archives collects and preserves the University's own archives, the Pacific Research Archives, the National AIDS Archive Collection and the Noel Butlin Archives Centre which is the largest collection of business and labour archives in Australia. These records are available for all researchers not just ANU staff and students.

The ANU Archives digitises our most highly used records and makes them available through the Noel Butlin Archives Centre or the University Archives communities of the ANU digital repository.

Search the Archives collection

Search results will appear in a new browser window. The [database](#) includes information about all deposits and series, the organisations and people who created them, and 40% of lists of items as at May 2014. The [help](#) page gives more information about how to use the database.

The ANU Archives database at www.archives.anu.edu.au as at May 2014

Next steps

To date we have been concentrating on making our finding aids accessible to researchers online through the database so that they can search our holdings quickly and easily and discover any digitised content. The ICA-AtoM software can also be used to manage a number of other archival functions which we hope to implement in coming years. These include:

- accessioning of incoming transfers
- retrieval of material from the repository using location data
- recording 'current controlling organisation' for all deposits for access and copyright permissions.

Already the existence of the database has changed the way that we can respond to reference requests. In the past we would photocopy or scan lists to send to a researcher based on our assessment of which deposits might be relevant to the request. Researchers can now undertake their own searches to discover deposits and items of interest to them. Our staff have also increased their own knowledge of the collections we hold by undertaking the various data entry projects, and we now have the framework in place to undertake further digitisation projects in the future.

Maggie Shapley
University Archivist
Australian National University