Avoiding extinction: the case for a national research collection.

Dr. Craig Anderson.
University Librarian,
Swanston Library, RMIT University, Box 2476, Melbourne 3000
Craig.anderson@rmit.edu.au

Abstract
Rapidly changing technological and publishing environments coupled with shifting influences in tertiary education will require dramatic solutions for the storage of low-use legacy print material, especially monographs. This paper discusses recent developments in the academic library environment which may support the need for shared last copy/single copy collections. The paper suggests a way forward which can provide cost effective, reliable long term access to print material which may otherwise disappear. This paper focuses on the preservation of monographs within Australian academic libraries and offers some preliminary costs of such an exercise using an existing last copy facility as a case study. Paper developed from presentation at ALIA Biennial 2012 Conference: Discovery. (Anderson, 2012)

Keywords: Last copy, Print repository, Shared collections, CAVAL, CARM

Background:

Decreasing use of physical materials
Over the past ten years there has been a precipitous drop in the number of loans of print on paper books (p-books) from Australian academic libraries, both in terms of both absolute numbers and loans per student. The median number of loans per library rose between the years 2000 and 2002, reaching a peak of 392,000 in the latter year, and has fallen almost every year since, with the 2010 figure reported at 333,000.(Figure 1).
The drop has been more dramatic than Figure 1 indicates. There was strong growth in student numbers from 2000 to 2010, with the median number of students per institution growing from 14,000 to 24,000. When these two sources of data are combined as median loans per student the trend is clear, with a fall from 22.07 to 12.44 loans per student (Figure 3).

There is a clear trend towards fewer loans, and while the loans per student figure may have stabilized between 2009 and 2010 the overall trend is for a decreased use of physical
collections. This decrease in loans has not been matched to this point by a decrease in the production and sales of p-books.

Sales of p-books have shown a trend which runs counter to academic library loans. In spite of conflicting information provided by Amazon at various times, p-book sales have been variable, and in some cases have seen dramatic increases. An example of this can be seen in the Publisher’s Weekly report for February 2012, where an increase of 24.4% in adult hardcover books is noted when comparing with the same month in 2011 (AAP StatShot report February 2012), while e-book sales had only risen 9.9%. These loans and sales figures illustrate that there is still a great deal of publishing of print on paper, but that academic library users are borrowing fewer books.

As well as sales of p-books being variable, there are still a large number of books which are available as print on paper only. According to Spiro and Henry (2010, 11) “much content is not yet available electronically, business models are unsettled and multifarious and universally satisfactory solutions for reading long-term scholarly works on screen have not yet emerged”. The Spiro and Henry report goes into considerable detail on the challenges of creating a hypothetical new research library which is digital only and discusses the lack of digital materials.

Increasing need for space

In spite of the decreased number of loans, and consequent reduced immediate need for p-books, academic libraries would appear to need more space for their other activities rather than less. Much of this need for space is being driven by the use of libraries as student study space. Universities have actively encouraged this use, along with increased space for private study, group work and the availability of computer workstations. While gate count figures are collected by only a selection of libraries, a steady increase is shown in the Council of Australian University Librarians (CAUL) (2012) “Statistics” with a median of 300,000 entries in 2004, growing to one million entries in 2010.

In addition to growth in the use of libraries, further growth is expected in the overall number of students. The Australian government has set a target of 40% of 25 – 34 year olds having a bachelor’s degree by the year 2025 (Commonwealth of Australia 2009, 12). The current post-secondary education rate for this age group is 32%. An attainment of the goal of 40% would increase student numbers by 25% from the current 857,384 equivalent fulltime students (Commonwealth of Australia 2011) to 1,071,730 equivalent fulltime students. With a conservative assumption that there is no total population growth, this is an increase of 214,000 equivalent fulltime students. These additional 214,000 students will require additional classrooms, as well as space for teaching and support staff. More space and other resources will also be required for other facilities including libraries. The Tertiary Education Facilities Management Association (TEFMA,) recommends allowing 14.5 square metres of built campus space for each equivalent full time student (Bradley 2009, 11). According to those guidelines more than 3 million additional square metres of built campus space will be needed to accommodate the proposed additional students.

Common perception that “print is dead”

Amongst academic administrators as well as much of the general public there appears to have been something of a perception that “print is dead”. This is driven to some extent by
enthusiastic press releases from organisations such as Amazon who have invested heavily in an electronic infrastructure and devices and are keen to sell these devices (Galante, J. 2010). In addition, the general press often seems to have reported on the migration to e-books in an apparently uncritical manner. When actual sales figures are examined, the e-book vs. p-book picture appears more complicated than a simple report on e-book sales might indicate. While p-book sales are trending downwards in some areas, they have been fairly resilient in others. The Bowker “New titles and editions report” (Bowker 2011), which is based on ISBN assignment provides details of the growth of traditional print on paper for the period 2002 – 2011. This report shows a 20 % growth in education titles, and 14% growth in both philosophy and psychology when comparing 2010 with 2011

In addition to the replacement of p-books with e-books through traditional sales channels such as Amazon, the advent of Google books has presented another alternative to print on paper. While the reality is different, there appears to be a perception that “it is all available on the internet”. A report titled “Redefining the Academic Library” (University Leadership Council 2011) is one of the more recent examples of this. This report states:

In only the past 10 to 15 years, a number of non- and for-profit organizations have amassed collections that dwarf academic libraries in size and scale. While libraries from community colleges to massive research universities typically house between 50,000 and 1 million volumes, companies like Amazon and Google offer 7.5 and 15 million volumes, respectively.

While the library’s unique role in offering access to their collections free of charge distinguishes it from the corporations, inexpensive and instant digital access to books has already quickly displaced the value proposition associated with access to physical print collections, and the availability of free, quality resources outside of the library is expanding rapidly (University Leadership Council, 2011, p. 8) (emphasis added).

The report provides minimal coverage of the very real constraints around copyright. These constraints are only recognized on page 20 with “Most importantly, large scale efforts at providing access to millions of scanned books are currently limited by copyright protections”. Page 22 goes on to say:

OCLC’s Constance Malpas recently studied HathiTrust’s collection in depth and found that (assuming a workable licensing model for accessing in-copyright works is established) its archive already duplicates a significant portion of even the largest research libraries, and could therefore allow considerable space and storage cost savings for those who elect to withdraw volumes held electronically in HathiTrust (University Leadership Council 2011, 22) (emphasis added).

While the report does mention copyright as a restriction to moving to an e-book based library, it is expressed in the manner of a minor, and soon to be resolved, problem. No mention or prediction is made around the Judge Chin decision or the future of Google books (Hamblett 2011). This may provide academic administrators, who are the target for the report with the impression that a predominantly e-book environment is imminent when in reality these very real barriers continue to exist.

Nor are the problems around non-U.S. based access to Google books mentioned. If the Google books settlement had been found acceptable it would only provide access to North American users as the settlement did not cover any areas outside of North America.
Does it matter?

There is considerable informal debate within Australian academic librarianship about whether there is a need to retain lesser used items within collections. While some libraries have a policy of retaining nearly all unique items, others have a much more dynamic collection with the aim of steady state collection, where an item is discarded for each item which is ingested.

Much of the dialogue within Australian academic librarianship around print monograph retention seems to follow one or more of three themes. These are:

- There will always be a copy somewhere else, and this will be available on inter-library loan
- Low circulation items take valuable space and should be discarded
- In the near future there will be an electronic copy, or the electronic copy already on HathiTrust or Google will be accessible

An alternative viewpoint might consider the possibility of loss of research material as libraries simultaneously discard the same item, including the last copy or at some future time, universities become much more competitive and refuse to inter-loan material, or only do so at very high prices. In 2011, 336,521 items were withdrawn from Australian academic libraries (CAUL 2013). While these could have all been additional copies of titles already held and of which a single copy was retained, the quantum of this figure gives rise to concerns that some final copies may have been discarded. The implicit risk in all of the above arrangements is that the last copy of a unique title in Australia or in the world may be discarded. This loss of intellectual diversity might be compared to the extinction of a species or the loss of history as we cannot predict what future generations may find useful and what they may need to reinvent.

Some answers:

Australian libraries have variously considered or implicitly rely one or more strategies for sourcing print based research material in the future. These include:

Reliance on “informal libraries of last resort”

A number of libraries, particularly those in the “Group of Eight”¹ have well understood policies on last copy retention, whereby they intend to keep their final copy of any title purchased for an indefinite period. Some libraries with a more dynamic collection policy which supports discarding lesser used material rely on these “informal libraries of last resort” to have the required research material in the future.

¹ The “Group of Eight” is a group of the oldest Australian universities which have generally built strong research collections over a long period of time. These universities include the Australian National University, Monash University, The University of Adelaide, The University of Melbourne, The University of New South Wales, The University of Queensland, The University of Sydney, and The University of Western Australia. “The Group of Eight (Go8) is a coalition of leading Australian universities, intensive in research and comprehensive in general and professional education. The Go8 exists to:

- enhance the contribution of its member universities to the nation’s social, economic, cultural and environmental well-being and prosperity; extend the contribution of its member universities to the generation and preservation of the world’s stock of knowledge;…” (Group of Eight Australia, 2013)
Reliance on commercial suppliers

Implicitly many libraries rely on commercial suppliers such as Springer, Proquest, Elsevier, Thomson-Reuters or others to provide them with the necessary electronic content. While this has worked well with journal content which is continually refreshed, monographs have been a smaller part of their offerings, and there appears to have been very little done with out of print monographs, especially “orphan works” for which no rights owner can be easily found. This lack of interest by commercial vendors should not be surprising as there is little commercial market for most of these items, although their value to future scholarship cannot be estimated.

Reliance on Google/ HathiTrust

Either through a lack of understanding of the copyright implications, or an optimism that access will be provided at some future time, a number of libraries appear to implicitly believe that it is safe to discard low use items on the basis that they are, or soon will be, available in digital form. This philosophy can be seen in the “Redefining the Academic Library” (University Leadership Council 2011).

Single copy / last copy collection

A group of CAVAL members created a single copy/last copy collection in the mid 1990’s (commonly known as the CARM – CAVAL Archival and Research Materials - Collection) which is still growing. This collection has few collecting guidelines beyond specifying that the items must not already be held in the shared collection and must not represent a threat to the physical preservation of other items held. This collection has been built over the past 15 years with low use items which have been removed from member libraries collections and ceded to CAVAL. Currently the collection holds over 700,000 items. A single copy / last copy collection allows libraries to confidently remove items from their own collection, and in particular from open shelves in the knowledge that should an item be required for research in the future it will still be available. The remainder of this article discusses the rationale for, and cost of a single copy / last copy shared collection.

Single copy / last copy shared collection

What a single copy / last copy shared collection might look like

A shared collection which holds the last copy of a work on behalf of all participants, such as the CAVAL shared collection appears to be both cost effective and safely retains research material in perpetuity. Although there are many advantages to a shared collection, particularly when held in a high density facility, funding remains a perceived problem. The following section outlines the possible costs. While the costs provided below are indicative, they are all based on actual experience. As the cost of building and operating such a facility is shared across the entire sector, costs would be far lower than having multiple institutions holding the same item. The cost differential is dramatic.

Cost per volume

According to Courant and Nielsen (2010) the cost to hold an item on open shelves is typically $US 4.26 (approximately $AUS 4.09) per volume per annum. This cost is calculated

---

2 CAVAL is a “CAVAL is an Australian not-for-profit company established to provide library services to libraries in Australia, New Zealand and Asia”. (CAVAL 2013). The author has been a member of the Board of Directors since 2005
on a conventional net present value methodology and includes construction of a new facility every 40 years as well as maintenance and operation. Using the same calculation methodology their cost for high density storage is $US 0.86 ($ AUS 0.83) per volume per annum. Sharing high density storage obviously lowers this cost again – if 39 Australian universities\(^3\) shared the cost this would bring the median cost per library to $AUS 0.022 per volume per annum for each university (this, and all subsequent costs are provided in Australian dollars unless otherwise noted).

The above figures all relate to U.S. building costs as at 2009 and do not include a depreciation allowance, but do include rebuilding costs on a 40 year cycle and all operating costs, as well as the cost of circulation. By way of triangulation, CAVAL, which recently built and currently operate a high density facility calculate the annual storage cost as $AUS 0.84 per volume per annum over a period of 10 years, dropping to $AUS 0.75 per volume per annum for a 50 year term. The CAVAL figures do not include the cost of rebuilding of the facility, nor circulation costs but do include depreciation. It would appear that Courant and Nielsen's (2010) costs are comparable with the Australian experience.

The following calculation is based upon a 20 year storage period on the assumption that by the mid 2030's either scholarship will have changed dramatically and that older materials will no longer be required, or that mass digitization costs as well as copyright barriers will be resolved and the stored material no longer needed. It also assumes that the Courant and Nielsen (2010) figures are close enough to Australian costs to provide a valid comparison.

Courant and Nielsen (2010) also assign a cost for open storage. In the absence of any comparable Australian data, these are used to compare the cost to store an item for 20 years in open and high density access. At a cost of $US4.26 ($AUS 4.09) per volume per annum, it will cost each university $AUS 81.80 to keep an item on open shelves for 20 years. If the same item was in high density storage, at a cost of $AUS 0.84 per volume per annum the 20 year cost is $AUS16.80. If this cost were shared equally across the 39 Australian universities listed on the Universities Australia website (Universities Australia 2013) the cost per university would be $AUS 0.43 over the entire 20 year period.

How big?
One of the key questions in regard to a joint storage scheme is the number of titles which need to be held in such a facility. This critical question of quantum needs to be based upon a number of factors including a definition of what “unique copy” means, the type of material held (materials such as textbooks, popular reading and titles which are held commonly held by all universities are all candidates for exclusion from such a collection) and the ownership model of materials in the facility. At one end of the scale, such a collection might be as large as ten million volumes. This would be more than triple the size of the collection held by any university library in Australia and is based on the figures provided by Missingham and Walls (2003). This figure corresponds to the number of unique titles held by all Australian universities which are on the NBD (National Bibliographic Database) in 2003. This would, in essence, provide a national university library print collection of lesser used research items.

\(^3\) There are 39 Universities in Australia. For more information please see the Universities Australia web site (Universities Australia 2013).
The following section estimates the cost for such a facility on two cases. The first case is for a facility which holds 10 million volumes and provides a large majority of the research items Australian university researchers might need. The second example is more modest and suggests a collection of just three million volumes. These examples are polar ends of the scale of such a facility and are provided by way of illustration of the possible cost savings.

Cost of a shared “Academic Library of Australia”

Using the Missingham and Walls paper (2003) regarding collection overlap as a starting point, 9.5 million unique titles were identified as being held by Australian university libraries in the NBD (National Bibliographic Database) in 2003. In order to cast this idea as widely as possible, the following calculations assume that:

- It is desirable to keep one copy of every monograph even if not highly used which is currently held by university libraries.
- That the State and the National Library collections are ignored as these are beyond university control.
- That special and public library collections similarly are beyond control and are disregarded.

There are a series of one-time building costs, ingest costs, and operating costs. The figures below are based on the building and operation of the CARM 2 facility, which was opened in December 2010. While the Courant and Nielsen (2010) figures (above) were used for cost comparison between high density and open shelving, the CAVAL CARM 2 figures are used in the following section as they are more granular, more recent, and they are Australian. These costs are provided as an illustration of the possible sector-wide savings and would have to be refined with any more formal proposal. The savings are dramatic enough that any refinement of the costs will approximate the same degree of savings.

The first scenario suggested in the following section posits that a facility capable of holding 10 million monograph volumes is built and is filled over a 10 year period, with one million items added each year by universities around Australia.

Building costs:
The CAVAL CARM 2 facility, with a capacity of 2 million volumes in high density storage cost $AUS 14 million (exclusive of land costs). Holding 2010 cost as a constant, this produces an initial building cost per volume of $AUS 7.00. Extrapolating this to 10 million volumes provides an initial cost of $AUS 70 million. This estimate does not include any economies scale of a larger building nor does it include the land cost.

Ingest costs:
In order to maximize use of space, items in high density storage facilities are stored according to size, rather than classification order. To allow for the storage and retrieval of items it is necessary to use a warehousing system and/or modified library management system, and to process each item individually. CAVAL experience has shown this cost to be in the order of $AUS 1.15 per volume, (Jilovsky 2012) although this cost does vary according to a number of factors. For the purposes of this exercise, on the assumption of the
possible ingest of one million volumes per annum, the annual ingest cost for a million volumes would be in the order of $AUS 1.150 million.

Building related operating costs:
Operating costs include expenditures such as building insurance, maintenance, retrieval of requested items and depreciation. Other operating costs include maintenance of systems such as fire control and warning systems, climate control and governance. A recent annual operating cost for the CARM 2 facility is provided as $AUS 950,000 (Jilovsky 2012). As some, but not all of the operating costs are independent of facility size an annual cost has been extrapolated by calculating the cost of operating a facility five times that of CARM 2’s capacity of two million volumes, (to allow for 10 million volumes) on the basis that the larger facility will cost more to run and then halving that figure, on the basis that there would be some savings due to economies of scale as well. This provides an estimated annual operating cost (exclusive of item ingest) of about $AUS 2.5 million annum.

Annual operating cost:
On the basis of a 30 year accounting life for the building and straight line depreciation at 3.3%, the annual cost to operate such a facility would be $AUS 3.65 million. This estimate allows for depreciation, ingest costs and operating costs but not initial build cost, and consists of the following components:

Ingest for 1 million volumes per annum:
(to prepare for high density storage) $AUS 1.15 mill.

Building related operating costs: $AUS 2.5 mill.

Depreciation allowance: $AUS 2.3 mill.

Total annual operating cost: $AUS 5.95 mill.

Shared across 39 universities this provides an average cost of $AUS153,000 per annum per university.

The projected costs in the previous paragraph only refer to the annual operating running cost exclusive of capital building costs. The capital building costs could possibly be funded by the Commonwealth government, or if this funding could not be found, and universities needed to pay this initial cost themselves, the one-time capital contribution would be an average of $AUS 1.8 million per university.

To summarize the above an initial contribution of $AUS 1.8 million per university, and an ongoing contribution of $AUS 153,000 per university would provide a research collection more than triple the size of that of the largest university library in Australia. This collection would be jointly owned and controlled by the university sector. All libraries across the sector could safely discard items in the knowledge that at least one copy of each item would still be preserved and accessible. It is reported that the average number of holdings per bibliographic record in Australian university libraries is 1.32 items (Genoni and Wright 2010). On that basis, a last copy/single copy collection of 10 million volumes would relieve universities of 13.92 million volumes being stored on open shelves, which, at a cost of $AUS
4.08 per annum provides an annual cost saving of just over $AUS 56 million per annum across the sector.

A second scenario would be to aim for 30% of total unique titles (approximately three million) to be held in a shared collection. If we hold the above cost figures constant but reduce costs by 70%, this provides an initial build figure of $21 million, and an ongoing operational cost (including depreciation, ingest and operating but not initial build cost) of $1.09 million. If this annual cost was shared equally across 39 universities, the cost would be $28,000 per annum per university.

The sector wide savings even on the smaller example of three million volumes would be significant. Using the Genoni (2012) figures again, a last copy/single copy collection of three million volumes would relieve universities of 3.96 million volumes being stored on open shelves. At an on-shelf storage cost of $AUS 4.09 per annum this would represent a savings of $AUS 16.2 million per annum across the sector.

**Other advantages of a shared store**

In addition to significant cost savings, the creation of such a storage facility would present several other advantages, including efficiencies of scale, joint control and the strength of collective wisdom.

**Efficiencies of scale:**

A facility of either 3 million or 10 million volumes would introduce an operational scale which would allow a greater degree of efficiency than would be possible in numerous smaller facilities. This greater scale would allow for better support of the governance and operating costs, as well as providing the opportunity for the acquisition of equipment such as automated digitising equipment which might be too expensive for a single university.

**Joint control:**

Policy control of such a collection would need to be vested at a high level such as CAUL or Universities Australia. This would reduce the risk of short-term local decisions being made which are driven by local priorities. This joint control would allow for the greater good of the sector and the wider community to be taken into account.

**Collective wisdom:**

Building and management of such a collection could enable the exercise of a collective (or “crowd sourced”) wisdom guiding the philosophy, policy and practice of a joint collection. This collective wisdom could support a greater degree of independence from local pressures and might lead to increased community benefit which may not occur with local collections.

**Barriers**

There are many potential barriers to such a collection. Some of these barriers are summarized by Genoni (2011, 41) in a qualitative survey of 12 Australian academic librarians. Genoni concluded that there is “some support for collaboration as a principle, this alone is not likely to be sufficient for Australian university librarians to support the adoption of a system-wide approach to storage issues”.
Some of the barriers mentioned by Genoni’s participants, and echoed by Payne (2007) include cost, concern about university rankings (and in the U.S. concerns about accreditation), resistance from academic staff, and the financial impact on the University’s asset base. These are discussed below.

Cost
The cost of building any shared storage facility is prohibitive for several reasons. The first is that the asset is unlikely to be owned by the funding authority, and while universities may sometimes be willing to fund facilities for direct use of their staff and students, they are unlikely to fund a facility which is not of direct and visible benefit to their own academic community. In addition the cost of such a facility is in absolute terms substantial, even if that cost was shared amongst a number of universities.

Rankings
University rankings are important in the increasingly competitive environment in which Australian universities operate, and fear that a reduced monograph collection may reduce rankings has been informally stated by some Australian University Librarians, however neither the Time Higher Education rankings (Times Higher Education 2012) nor QS rankings (QS World University Rankings – 2012) refer to collection size as a component in their ranking systems.

Resistance from academic staff
In many cases academic librarians respond to real or feared pressure from academic staff to retain items on open access. Academic staff often state that they require monographs and journals on close-by open access to promote serendipity and to provide ease and speed of searching.

Conclusion
There is clear evidence that substantial cost savings and advantages are available both to the university sector as a whole and to individual institutions though shared storage of legacy collections. This paper proposes that it is in the long term interest of university libraries, the university sector and the community as a whole to investigate these possible savings and the advantages (and disadvantages) of a national research collection in the interests of Providing cost effective, reliable long term access to print material which may otherwise disappear.
References:


