Image Storage And Retrieval
A Tool For Museum Collection Management

Richard Llewellyn
Registrar
Australian War Memorial
GPO Box 345
Canberra ACT 2601
AUSTRALIA
Fax: 06 243 4325

Abstract

The Australian War Memorial is a large and complex museum and research institution. It has recently introduced a multi-million dollar computer system to control and assist in every facet of the management of its collections. The use of on-line image retrieval of collection material is a major integral part of this system, not only for client access to collection material but also for basic documentation purposes. The author examines the role of images for collection management in some depth.

Australia is a young nation. Until Federation in 1901, we were a collection of British colony states. Federation was not the culmination of any titanic anti-colonial struggle; indeed it was something of the reverse, for we were prodded and coaxed, almost hesitantly, out from under the mothering wing. Because we became a nation gently and without any massive social upheaval, we had no truly definite sense of national identity, no heroes, no national archetypes - with the possible exception of the bushranger. The first identifiable Australian national persona came with the first world war. That character was the Digger - the tough, laconic, disrespectful of authority but loyal and brave bronzed Aussie. Like most good tales, the story of the Digger was almost certainly 90% fanciful embroidery upon a thin and hard-stretched cloth of truth. The myth of the Digger has passed into folklore but certainly not out of memory, and this is of primary importance in understanding the place that the Australian War Memorial plays in our cultural heritage.

The concept of the Memorial was formed on the battlefields of France in 1916. The idea of commemorating those who served and died in this war was not unique to Australia, but

the institution that has evolved is truly unique in the world. The Australian War Memorial is not merely a ceremonial place, a shrine to honour the dead, but a museum where the living can come to understand, through its displays and its research resources, the full meaning of war.

Australians are great collectors - that is to say, great thieves at the impersonal level. What started out as a collection of the official records of the war rapidly expanded to a collection of anything that was not totally immovable - although we even have a mosiac liberated, tile by tile, from the wall of an ancient building within the confines of the Ottoman Empire, a piece of a middle-ages French cathedral, and a pair of medieval stone lions from Belgium (the lions were, however, a present from the Belgian government). Most people's idea of records does not include tanks - we have seven. Naval projectiles weighing several tons can perhaps be considered paperweights and therefore connected loosely with records, but by no stretch of the imagination would you normally consider field rations to be records, yet last year we found a piece of 1917 cheese amongst our souvenirs. It was well on the way to maturity, and we hope yet to unearth a good port of a similar vintage to match it.

I won't labour the point any further, but I hope that by synthesis you can comprehend our beginnings: a massive collection of disparate objects, without classification or order, and often with little or no description. We also had no staff and no home - indeed, even at the end of the war, no legal entity. Balanced against this was a massive emotional push for the creation of a memorial, for Australia, in proportion to its population, had suffered the greatest number of casualties of any of the combatant nations. There were very few families who had not lost someone, and it is not stretching things too far to say that the first 'crop' of young men of the new nation had been decimated, for there were recorded instances of boys as young as fourteen and fifteen enlisting, and dying - boys who were being born even as the ink was drying on the documents of Federation.

Australia has never forgotten the effects of war on its development as a nation. The Australian War Memorial was the first truly national cultural institution to be created, and in many ways it remains our most important national cultural institution. We have also tended to hold a somewhat favoured status for funding and staff because of the generally high level of community identification and support we receive. However, it must also be appreciated that the community also demands a very high level of service from the Memorial. No other cultural institution in Australia is so strongly regarded as belonging to the people, or has anywhere near the same high public profile. We have to be especially responsive to the demands for service and access placed upon us.

The Memorial is one of the largest museums in the country, with some 8,500 square metres of exhibition space, and over 200 staff. We are currently planning to add about another 8,000 square metres of exhibition space to cater for our in excess of half a million visitors per year. We occupy five individual buildings, spread over a distance of about fifteen kilometers. Two are storehouses, of about 6,000 square metres. These are not normally occupied by staff, and access by the public is only possible through prior arrangement. The

Memorial building houses all of the public exhibition galleries, the Research Centre which combines a library and an archives office, and our educational programs area, as well as our formal commemorative areas. Adjacent to the Memorial building is the Administration building, with the majority of the curatorial staff, Registration, and the normal complement of administrative staff. The casual visitor does not use the Administration building, but there are several specialist reference areas open to serious researchers. This building also houses our computer network central facility. Some eleven kilometers away is our Conservation facility, which also has environmentally-controlled storage areas for paper and textiles, and some of the less stable wood and metal relics.

Our collection includes around 800,000 photographs and about 70,000 accession records of relics - but of course there are many more individual relic items. We have nearly 30,000 works of art, posters and pieces of sculpture, 60,000 plus monographs and serial titles, thousands of personal records - diaries, sketchbooks and the like, thousands of maps - the list goes on. In addition to providing displays in our galleries, we support research into all aspects of Australia's military history - including the other side of the coin, so to speak, of anti-military and anti-war apects such as anti-conscription rallies. We also have a legislative requirement to provide information from our holdings of official government records to establish war veteran's rights and entitlements such as pensions, health care and so on.

The Memorial is quite a large institution, with a broad range of material to manage and an equally broad range of client service demands. I think it is therefore reasonable to say that the functional requirements for our image storage and retrieval system are no less than those of any other large and diverse museum. Having said all that, I'd like to move to the focus of this session - collections management. Before I continue, let me define one point of language. I use the term Image Storage and Retrieval to describe our system, for that is precisely its function. This system does not provide any manipulation, enhancement or adjustment of images beyond mild colour balance correction on both input and output devices. While we also have image manipulation facilities, our use of these is very restricted, primarily to archival reconstruction of images.

We are only now completing the first phase of our image system implementation, so there are experiences to come. However, I can tell you what we have achieved and learned to date, and more particularly, the role we have identified for the system. There are people here far more qualified to talk through the technical intricacies of image systems - I speak as a manager of collections rather than as an image technician. Our image system is in fact one facet of a new computer facility with a total project cost of over A$3 million, to enhance the management of our collection. While the image system is without doubt the show pony, it does not stand outside the rest of the system, as you will see.

The Memorial commenced the use of computers as a tool for collection management in 1983. By 1986, when I started as the Registrar, we had, by and large, our first generation system up and rolling, to control acquisition, location and loan of relics and works of art. This work gave the first really clear picture of what our problems were.

I am sure that most museums will recognise some or all of the following complaints:

- a large collection of badly or inadequately classified and described items;
- a history of lack of control over description vocabulary and format;
- wide variation in the quality, quantity and even availability of control records
- user access demands which are either difficult and costly, or even impossible to meet because of the problem of identifying, locating, or providing access to material;
- physical deterioration of material from repeated handling for user access reasons;
- a need to balance ideal and efficient storage methods with access requirements.

Let me assure you that, if you lack experience of, or exposure to, any or all of the situations on that list, then we can provide a ready educational program featuring live examples. Equally, let me assure you that we see our image storage and retrieval system not as solving, but certainly as helping to address all of these problems.

First, a thumbnail sketch of our overall system. We are now implementing a PC-based network of currently around 70-75 workstations. We use 25-mHz 386 based workstations with 6Mb. of RAM installed. We have two Compaq Systempro 33 mHz 486 file servers, one with 24Mb. of RAM, the other with 8Mb. We are now in the process of instaling out third Systempro to act as a development and back-up machine. The Memorial building is linked to the central network facility via fibre-optic cable, while we have our own microwave link to our Conservation Annex which is some eleven kilometers away. Currently we run on the DOS platform, with Novell Netware networking software.

The image storage and retrieval system allows in-house capture onto WORM laser disks, which are then transcribed onto glass masters. A central bank of players outputs the images to a video-bus network for image transmission to selected network workstations which have an additional image output monitor, controlled from a video frame-grabber card in the workstation.

Logically, the system works like this: users interrogate our computer databases for records which satisfy their enquiry. Not all of our material is yet recorded on our databases, but as part of the overall project we are capturing and even creating records for our total collection. Records which satisfy the search criteria are displayed on the normal workstation screen. When there is an image of the item available, the record indicates this to the user. Users at image retrieval workstations - there are currently seven of these, but the number will grow - can then request display of the image as part of the search routine, by pressing a function key - effectively a "show me that" button. The image location - disk number and frame number - is held on the database record. By software control, the location information is extracted, the enquiry station identification is added, and the resulting request is passed on the network to the image request file server, which commands retrieval of the requested frame or frames from the appropriate player. The image is output onto the video-bus, preceeded by the address of the requesting workstation. The requesting workstation receives the image, stores it via Targa frame-grabber card, and displays it on the attached image monitor screen. Output can also be directed either to an attached low-resolution thermal printer, or to a high-quality video printer attached to the viewing station.

Obviously, this is technically somewhat more challenging than my description implies, but it is all current, off-the-shelf technology.

Let me return now to how the image system fits into our collection management activities, and to how it helps us to solve those problems I mentioned before. These problems might reasonably be divided into two groups: those concerned with documentation of the collection, and those associated with physical access to the collection.

The first three problems I listed form the documentation group; that is, they all revolve around the question of the availability and quality of records of items in the collection. Obviously we cannot manage our collections if we do not have some sort of records. It really doesn't matter a great deal what the format of these records is as long as they provide certain basic information: at the very least, sufficient description of the item to identify it, its current location, and sufficient organisation within the record to ensure that it is retrieved in the first place. The larger our collection, and the more demanding our management requirements, the more informative and complex our records need to be. For a museum of any reasonable size, the old days when the relevant curator knew his or her collection intimately have almost gone. Even if such legendary animals still exist, how are you going to access the collection when they're away? Worse still, how will you manage the collection when they are run over by that apochryphal bus?

I know that I preach to the converted - but in practical terms, if you have a large collection, old and inadequate records, and anything less than a plethora of staff and/or cash resources, you are at best struggling to provide both management of, and access to, your collections at the level you would prefer. That is certainly the experience we have at the Memorial, and the problems of access to the collection both for the client and for the curatorial staff are compounded by distance.

Let me now put into conjunction two hoary old phrases: "like trying to find a needle in a haystack", and "a picture is worth a thousand words."

The haystack to which I refer is made up of the thousands and thousands of words contained in the records of the collection. The needle, or perhaps I should have said "needles," are those few words which identify a specific item, or a small group of candidate items, which you seek. Modern computer searching techniques provide the user with enormous power to seek the needles in our haystack. We can allow quite lengthy descriptive captions onto our database, and yet construct a highly specific search through the use of sophisticated searching techniques: character string matching, adjacency searching, Boolean searching. But, if our records are not good, or not consistent, how do we know if we have in fact located the right item or items? One way is to note their locations, and go and physically look at them. Another way is to have additional descriptive material attached to the record which gives us further information - and that brings us to that loquatious picture.

A picture - or pictures, for three-dimensional objects - of an item, contains an enormous amount of information by which to identify that item. Let us look at an example typical to my institution.

A user comes to us asking if we could identify the old handgun his grandfather used to have years ago. He knows it was a souvenir from France in 1916. He can't remember the make - he was quite young when he last saw it - but it was a pistol, he thinks it had a foreign name, probably French or German. He can remember it had a funny shape to the trigger guard and a sort of a hook on the butt. We have over 2,000 handguns in our collection. By constructing the search on the computer carefully, we can select all pistols, of French, German, or Belgian (to be on the safe side) origin, with a date of manufacture earlier than 1916. I haven't done this particular search, but my guess is that the result would be a list of perhaps two to three hundred candidate items. Most of these would be in our armoury, so we could take our enquirer there and start, draw by draw, to examine each candidate. However, for reasons of security, we prefer to keep our armouries reasonably distant from the public gaze. Alternatively, if we had images of all of our pistols on the image retrieval system (which we do not as yet have), we could request these images.

Even if our written records were sufficiently detailed as to provide accurate descriptions of all the physical features of each handgun, which they are not, it would be touch and go that they would convey in the print sufficient idea of the weapon to allow the user to positively select a match.

Such records would be very lengthy indeed, and to read each one would take many seconds if not several minutes. Each record would have to be read almost completely to decide whether the item described was a viable candidate or could be rejected. To plough through 300 records would take almost two hours if the user could read on average one record every twenty seconds - and that is not allowing for any computer response time. Response time for a system holding large numbers of long record is usually quite slow.

By comparison, the mind can take in a picture of a pistol at least with sufficient detail to decide if it warrants close inspection very quickly - certainly in 3-5 seconds. The longest time required to have a selected image appear on the screen is less than 1.5 seconds, the average typically half that. So our user could scan our collection vicariously to make his identification either positive, or far closer to positive - in a little over thirty minutes.

Remember also that if your written records contain a lot of specific technical detail written in the language of the professional, then you may have to supply a professional - a curator - to act as interpretor of the written word. A classic example here would be the correct technical description of a coat of arms or other heraldic device. I will not dwell on the professional time required to create this level of detail in the records, for this is obvious.

Object images have, by and large, no such requirement for professional interpretation. Since our system is usable by most enquirers, even those with very minimal computer literacy, the image system has the potential of freeing many, many hours of staff time from user assistance to other activities.

I have used the example of a user enquiry to show how poor records of the collection can usefully be augmented with images, but the technique equally assists one's own staff who

may be identifying items in the collection for any purpose - exhibition design, conservation treatment, research resource, whatever.

This brings us to the second group of problems which I identified: access to the collection.

One part of access is the identification of required items, which we have just discussed. Recording their location is of course a routine entry on the item record. The actual physical access issue bears some further discussion. As with most museums, the vast part of our collection is not on display. Our storage facilities are quite some distance from the main Memorial building, and are basically not accessible to the public at any time except by prior arrangement, and not accessible at all on the weekends, so even to get near the items requires considerable effort on the part of the public and staff time to conduct visitors into the closed areas. Arriving at the near vicinity of the object is not the end of the story, however. For many items, actual inspection requires the removal of protective wrappings or even coatings; they may have to be unboxed, retrieved from special cases, or even assembled if stored in a demounted state. Stabilisation systems may have to be disengaged. Once the user has inspected the item, it has to be returned to proper storage conditions. It is not hard to see that considerable staff time is often involved in bringing the user face-to-face with the item - and if the item turns out not to be the one required, then this is wasted staff time.

We could, of course, simply refuse to make our collection available outside of the material on display, but this is not good museum policy. In our case, we have a special reason to make our collection as accessible as possible. A great part of the collection has been acquired by donation, and as you will appreciate given the nature of that collection, many of these items are of great sentimental value to the owner and his or her family. Can you imagine the effect on a family who has travelled perhaps 2,000 kilometers to get to the Memorial, of being told that they can't see their grandfather's medals and uniform simply because they are in storage? We would lose community support very rapidly indeed; we would also, in all likelihood, find that donations of material would dwindle or cease.

I believe that if we can provide an image, or images, of items in our collection readily for our users, we will satisfy a considerable proportion of access requests without having to go near the material, with significant savings in both staff time and potential deterioration / damage to the material itself. For access requests where ultimately physical access is required, the availability of an image will help reduce the occurrence of false leads, and staff time will at least be used productively.

In one special area, of course, the on-line image replaces contact with the original material almost entirely. I refer to our collection of photographs.

I mentioned that we have some 800,000 photographs. These were mostly taken by official war photographers, but we have some private material here as well. For the official war photographs, we have the negative; for private material we usually have only a copy. We will not place all 800,000 photographs on disk. We have identified a "core collection," cur-

rently numbering around 120,000, of these photographs, and these have all been captured on laser-disk.

The advantage for the user of access to the photographic collection on-line is mainly one of speed: many candidate images can be retrieved and examined very quickly to select the required one. Given that the record of the image - which relies on the caption - may not indicate clearly just what the image really contains, it is often necesary to look through many images to find one which is "the best" for the user's purpose. Selected images can then be printed out on the spot using laser printers. For many users, the quality of print from the laser printers will be sufficient for research purposes. However, where the intention is to publish, we will continue to insist that we supply a photographic reproduction of the original to ensure the highest possible quality.

This brings me to the question of the quality of images required in the image storage and retrieval system, which is a fairly natural way to conclude what I want to say here today.

For the storage and retrieval of images, quality is largely a matter of resolution, both for capture and reproduction. The functional characteristics of the system you require really predicates the sort of resolution you need. To a considerable degree the resolution you need also dictates the technology you can use. When we started out looking for an image system, we had no fixed idea of what we required. We examined laser-disk options, which as most people are aware is analogue technology of fixed resolution, and of course digital technology, which can range from very low to extremely high resolution. It became rapidly obvious that we had to determine exactly what we wanted our system to do, that is, what functional characteristics we required. I don't want to go into any great detail, but some of the salient points were:

- speed of retrieval
- minimal influence on the computer network
- ongoing cost of capture and storage
- reasonable cost of reproduction
- acceptable resolution.

We certainly did not require image manipulation, and even the ability to zoom in on parts of an image, while quite desirable, was not considered to be worth significant extra expense or load on the communications network.

The reason we settled on laser-disk technology is that it was the best 'fit' to our requirements. The speed of retrieval is better than virtually anything that digital technology could offer - maximum seek and read time is less than 1.5 seconds. Through use of the video-bus network, the performance of the computer network is not affected by the transmission of images. The impact on the network of the very large file sizes required for high-resolution images is a very serious consideration for a digital system.

The ongoing cost of capture and storage is acceptable. Each laser-WORM disk costs about A$600, and each disk stores 36,250 images per side. The glass master disks cost about $A2000 including transcription costs, and store 54,000 images. All up, using in-house capture, it is currently costing us around A$1 per image to capture and store each image from the photo collection. Each player costs about A$2,500. The recording stations comprising a broadcast quality camera, and the recorder and image processor units, cost around A$45,000 each. The cost of our two stations amortised over an expected 300,000+ images comes to less than A$0.30. For reproduction of the images, the user has a choice of a small thermo-fax copy, which is of fairly low but acceptable resolution for identification or basic research purposes, at about A$0.30, or a larger and higher-resolution laser-print costing around A$2.00. These latter images are A5 size, and of sufficiently good quality to be used as reference prints or even personal record prints.

Finally, acceptable resolution. The laser-disk technology provides resolution of 625 lines for the frame, which works out at about 80 dpi. for an original of about A4 size. While this is not high by digital standards - resolution of 1000 dpi or better is currently possible using digital techniques - it is surprisingly reasonable and certainly adequate for our intended use. If you need greater resolution, then you have no option but to use digital techniques - but you must remember the cost of storage. A full-colour, sixteen-bit image of about A4 size, captured at 300 dpi, occupies somewhere around 34 Mb. The effect on storage, and on network performance, of a large database of such images is very, very significant.

I should comment that the Memorial is also moving to laser-disk technology for audio-visual presentations in the galleries, to eliminate the problems of tape-based presentation technology, but this will be quite separate from our collections management facility.

If there is any message from our experience to date, it is this: image storage and retrieval can be a serious tool for the management of collections, which is, after all, the bread-and-butter of museum existence. As an aid to better documentation, it is likely to be highly cost-effective by comparison with upgrading written records. As an aid to client access to your collections, it has equally strong, if not even stronger, claims. Look beyond merely using image storage and retrieval for exhibition purposes, and the wider role for the technology is readily apparent.

