

Innovation during evolution: Document Supply Service digitises library collections

Margarita Moreno and Anne Xu

National Library of Australia, Canberra, Australia



Copyright © 2013 by **Margarita Moreno** and **Anne Xu**. This work is made available under the terms of the Creative Commons Attribution 3.0 Unported License:

<http://creativecommons.org/licenses/by/3.0/>

Abstract:

The Document Supply Service of the National Library of Australia has taken an innovative and strategic approach during its evolution in the digital age. While the section was providing efficient and client oriented interlibrary loan and document delivery services (ILL/DD), in 2003 it began digitising collections based on user demand, enabled by the introduction of the Copies Direct Service, an easy to use online ordering service. This bottom up approach to digitisation was taken as soon as technology to keep a scanned copy was developed by the Library. A decade later, while digitisation has now become an intrinsic part of the Library's strategic directions, the Document Supply Service is contributing about half of the in house digitised collection images as a by-product of its ILL/DD services.

The paper will cover the development of digitisation practices adopted by the Document Supply Service, discuss the convergence with the planned digitisation undertaken by the Library's Digitisation and Photography Branch and describe how the automated interlibrary loan and document delivery management system Relais has been developed and used to manage and report some of the digitisation activities. The National Library is currently undertaking the development of a books and journals digitisation program and the role of Document Supply Service in the Library's future digitisation program will also be discussed.

Background

The seeds for the Document Supply Service digitisation program were sown in 2000-1 when the then Interlending Services Branch Director was reassigned to set up the National Library's Digitisation section and its work program. Despite some differences, it was very evident from the beginning that there were many similarities between document

delivery and digitisation. Both processes require access to, handling of and caring for the collection; capturing/scanning collection items and delivering captured images. However while the digital images captured through digitisation are stored for the future, the images captured for document delivery are discarded.

The main drivers for establishing the Library's digitisation program were to:

- Enable users, regardless of location, to directly access and use a range of digitised materials relating to Australia and Australians;
- Promote an understanding of Australia and the Australian experience to those who may access the collections from overseas;
- Collaboratively build with other organisations a critical mass of digital items in particular subjects and formats that will form a significant component of Australian digital information in these areas; and
- Increase access to and help to preserve rare and fragile collections by providing digital surrogates.

In order to support the digitisation program, the Library purchased hardware and software where it was available and then built/developed software to fill any gaps. In 2001 the Library acquired a dedicated server and tape library to store its digital content, and then the digital archiving, digital collection management and the delivery systems were developed in house over a number of years.

The digitisation program commenced by selecting from the Library's rare and valuable collections, in particular items from the pictures, maps, music and manuscript collections.

The program began with music items and created files of data for the Music Australia service with works such as *I was dreaming* (<http://nla.gov.au/nla.cat-vn611517>), *Orotava* (<http://nla.gov.au/nla.mus-an3701052>), *Canberra's calling you* (<http://nla.gov.au/nla.cat-vn2061067>), *Aeroplane jelly* (<http://nla.gov.au/nla.cat-vn263241>) etc.

Digitising for Maps, and manuscripts followed with an emphasis on rare formed map collections: Rex Nan Kivell and Ferguson rare maps for the former, for example *Tahiti and Moorea* (<http://nla.gov.au/nla.cat-vn4534807>); and samples from manuscripts for the latter, for example the Barton papers (<http://nla.gov.au/nla.cat-vn2049726>); Wills' diary (<http://nla.gov.au/nla.cat-vn1299268>), Bligh's notebook and list of mutineers (<http://nla.gov.au/nla.cat-vn801340>), the tapa cloth samples (<http://nla.gov.au/nla.cat-vn1123278>) and selections from the Deakin papers (<http://nla.gov.au/nla.cat-vn2960190>).

Number of items in the NLA's digital collections by format 30 June 2012

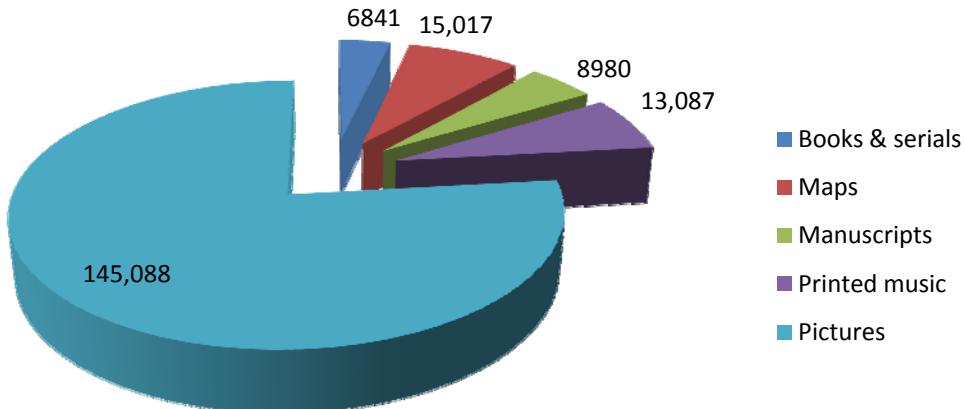


Figure 1 shows the number of items digitised from the libraries collections excluding newspapers as at 30 June 2012

This table shows the number of items in the Library's digital collections broken down by collection format. Although the number of digitised Books and Serials is quite modest (6,841 items) it actually represents over 200,000 images, well over half the total number of images in the Digital Content Management (DCM) system. Note this graph does not include any of the National Library's newspaper digitisation efforts which, in collaboration with other libraries, comprises over 10 million newspaper pages, digitised from microform, from over 480 newspapers and made accessible through the [Trove](#) Service¹ At about the same time the Library commenced its digitisation program, it also began automating ILL/DD services and expanding its offering to individuals with the development of the Copies Direct service. The Copies Direct Service allows users to place requests for material from the National Library collections online, supports electronic delivery of documents to their desktop and provides secure payment options.

¹ Trove is a discovery experience focused on Australia and Australians supplements what search engines provide in the field of social sciences, literature, local or family history, see trove.nla.gov.au/

Demand from the National Library's users is somewhat different than requests from the library community. On average 44% of requests from users are for material published prior to 1960 whilst for libraries this figure is 10%.

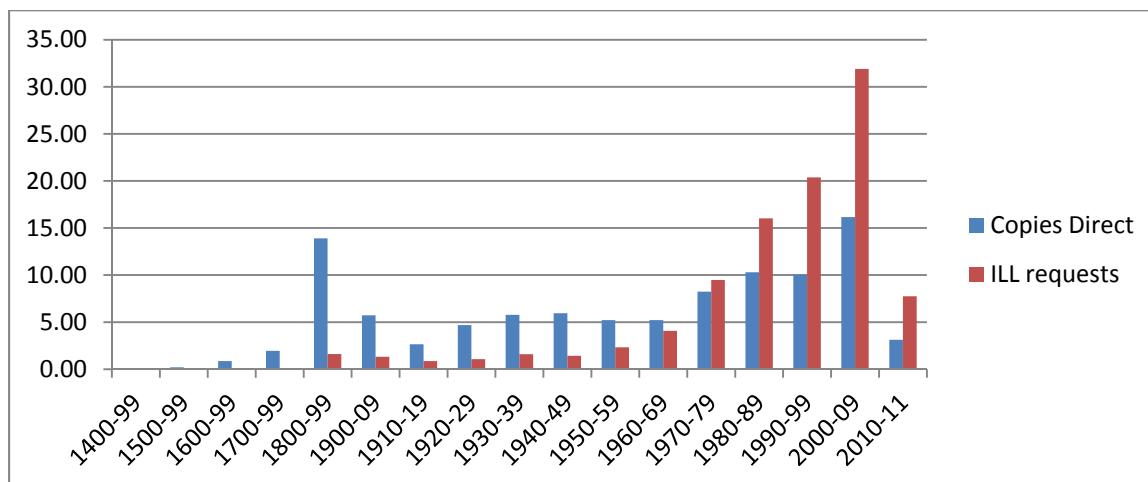


Figure 2 shows the percentage of requests by year of publications for Copies Direct and Interlibrary loan requests as a proportion of total requests received for each group based on 2011 data.

Almost immediately after the implementation of the Copies Direct service, the Library began receiving requests for copies of complete books that were frequently out of copyright. For in copyright material the user was advised of alternative options for accessing the item, e.g. interlibrary loan, but for out of copyright material, the item was scanned and delivered. It was recognised from the beginning that the scanned files after delivery could be re-used or accessed at a later stage. Document Supply Service commenced storing these digital files and so digitisation based on user demand, i.e. 'digitisation on demand' began. Over the last decade digitisation on demand undertaken by the Document Supply Service has evolved and changed in many ways, impacting on digitisation standards, workflows, practices and procedures as well as equipment and technology. These changes will be discussed in three phases – the beginning phase, current processes and future convergence.

The beginning phase

Initially, retaining files from digitisation on demand came about as a measure to take advantage of work already completed and avoiding re-scanning the same collection item when requested again. A further benefit was that keeping these digital files would assist in preserving the Library's collections if a digital surrogate was accessible.

Although these benefits were recognised, it was clear from the beginning that digitisation on demand and retaining the resulting file was only a by-product of the ILL/DD process and therefore not a 'core' function of the Service. This underlying principle impacted on the digitisation standards and workflows adopted by the Section and the standard ILL/DD practices for scanning and delivery were used to create these files.

The first challenge related to concerns in accessing special collections, i.e. material classed as rare, valuable and sometimes irreplaceable, for document delivery and digitisation. These ‘special collections’ were not available for ILL/DD, because collection managers had concerns about Document Supply Service staff handling the items, meaning they were only available for use on-site.

In 2002, the National Library developed new strategic directions that recognised the importance of providing broad access to all of the Library’s collections for both on-site and off-site users whilst still safeguarding them. “Our particular goal for 2000-2002 and beyond is that: all Australians, at their place of choice, have direct seamless access to the print and electronic sources of information”². This shift in focus gave digitisation on demand legitimacy and allowed the Document Supply Service to work collaboratively with collection managers and Preservation Services to establish guidelines on the use of the special collection material for document delivery and digitisation. It also enabled the IT team to develop simple solutions to upload these documents to the Library’s Digital Content Management System

The second challenge was in developing a ‘digitisation on demand’ workflow specifically for these requests, taking into account the standard turnaround time for ILL/DD. In Australia the Interlibrary Resource Sharing Code recommended five working days (later changed to four) for the delivery of a standard ILL/DD request.

For ‘digitisation on demand’ requests, in addition to assessing items for document delivery, the digitisation workflow requires careful assessment for copyright compliance as well as the item’s physical conditions against preservation guidelines for copying and the likely turnaround time to capture the material. So when a copy request for a complete item is received, it is identified at the preliminary sorting stage and passed to a supervisor for copyright assessment. If the material cannot be copied in accordance with provisions in the Australian Copyright Act, the request is declined and a message sent to the client to that effect, where appropriate suggesting alternative methods of access, for example interlibrary loan request through their local library. If the item can be copied, it will be retrieved, and its physical condition assessed against preservation guidelines, as well as a decision made on the equipment to be used to capture the item.

The digitisation process usually takes longer than normal ILL/DD scanning of journal articles or parts of monographs. There are a number of reasons for this, including extra care needed in scanning and the size of task, which can range from a small number of pages to over a thousand pages. In order to complete the digitisation work it is shared among staff to ensure the job can be done within required turnaround time. Sometimes negotiation with clients on turnaround time is required for big or complex copying jobs. As the Copies Direct service became popular, more requests were received for digitisation of complete works.

² Direction for 2000-2002/National Library of Australia

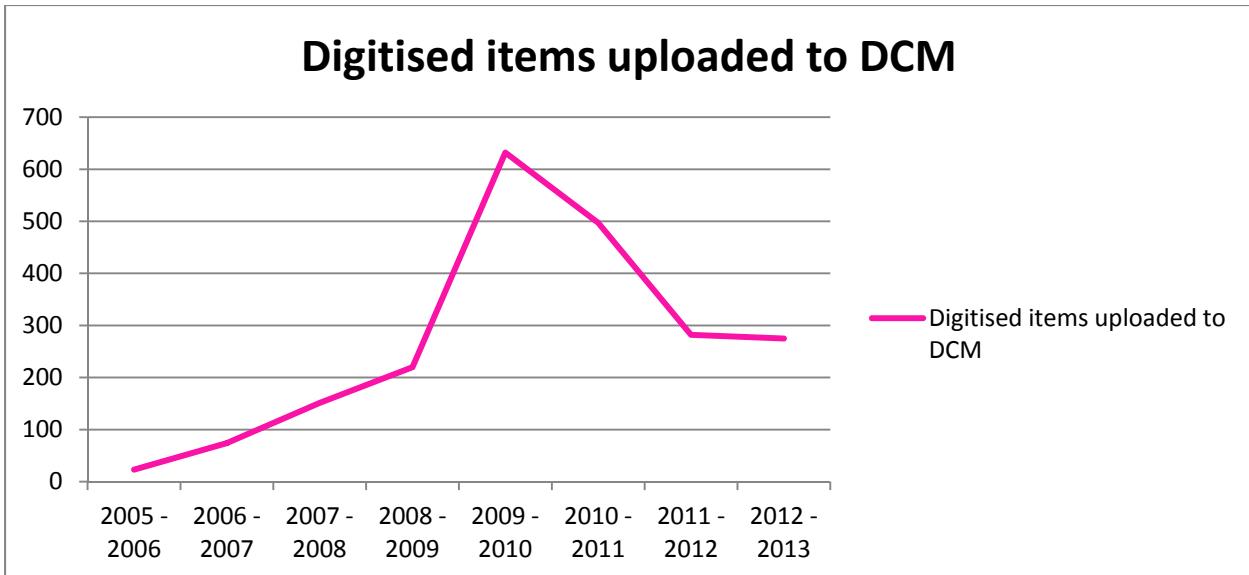


Figure 3 shows the increase in demand for copies of out of copyright material, digitised and uploaded to the Library's digital content management system DCM

Although preservation assessment for standard ILL/DD practices is an integral part of the process, the requirements for digitisation on demand are much more stringent, particularly for material from the special collections. The workflow involved greater collaboration with colleagues in other sections of the Library, in particular the Preservation team. It was agreed that the majority of requests would be assessed by trained Document Supply Service staff based on the preservation guidelines and Preservation staff would handle assessment for more complex materials.

Although a general training program organised by Preservation Services existed for routine ILL/DD copying, the digitisation on demand program resulted in more focused training, in particular how to assess and handle special and fragile collection items. If staff in Document Supply Service had any questions regarding a particular item before or during scanning, they would contact Preservation staff who would come to assess the item on the spot. As a result, Document Supply Service staff quickly developed knowledge and expertise in handling special collections material.

The third challenge involved the technical requirement for storage, equipment and software as well as the need to set digitisation standards. Document Supply Service first had to determine where to store the digital files after they were delivered to users and then how to make these files accessible longer term. In the first instance the digital files were kept on a network drive as there were no other options at the time. In order to access these files easily, thought was given to developing an appropriate naming convention. It was decided to use a combination of keywords from the title and the call number for the collection item. Before returning the material to the shelves, a bookmark made of acid free paper stamped with the date of digitisation and the Document Supply Service name was inserted, to alert staff and onsite users to the existence of a digital copy, promote its use and avoid re-digitising the item.

Document Supply Service has always been concerned about safeguarding the collection. For that reason, the Library used a combination of photocopiers, flatbed and overhead scanners. Preservation Services staff assessed the equipment used by Document Supply Service and advised that special collection material should not be scanned on the flatbed scanners and that the overhead scanners were more appropriate for this type of material. Document Supply Service had two Minolta 7000 overhead scanners at the time, which were capable of scanning material up to A2 size and only produced black and white images. There were also other limitations and intermittent problems with the equipment, for example, at times staff could not get a clean scan even though a thin perspex sheet was used to hold the book open. Sometimes staff had to use their gloved hands and fingers to hold an item in order to get a good image.

Document Supply Service also experimented with a camera to capture some of the more difficult material, such as oversize material like broadsheet newspapers that could not be captured by overhead scanner.

Over time the standards for digitisation on demand have evolved. In the beginning, the decision on digitisation standards was relatively straightforward, as the digital copies were a by product of document delivery processes and therefore the document delivery standards were adopted. The images were scanned in black & white at 300 dpi resolution and saved in multi-TIFF format. Minimal image cleaning up and editing was performed.

The Ariel software was used to scan images for document delivery and so it was also used for digitisation. Paperport image editing software was adopted for straightening and rotating images, adding missing pages or deleting duplicate pages, as well as performing minimal cleaning up of images i.e. erase thick black borders. As the required format for document delivery is multi-TIFF format, items were scanned or imported into PaperPort as a single file in multi-TIFF format and unstacked if inserting or deleting pages was required.

During this period staff learnt through trial and error how to digitise effectively. The workflow was efficient and enabled Document Supply Service to undertake digitisation without compromising the delivery of its core ILL/DD service.

Current processes

It took time, system development and workflow changes before Document Supply Service evolved to its current digitisation capacity. As mentioned earlier, the Library formed its Digitisation and Photography section (DAP) and built its technical platform in 2002. It consisted of the Digital Object Storage System (DOSS) and the Digital Collection Management System (DCM) built in-house to support digitisation processes. The DCM manages image uploading processes, access to the delivery systems as well as access to images on the DOSS. When an item is digitised it is uploaded to DCM, and the catalogue record is updated with the Persistent Identifier (PI) for the digitised item in order to provide access through the Library's online catalogue. The system was initially developed to handle special collections material, including pictures, maps and sheet

music, and therefore the system specifications developed were based on the requirements of these special collections, such as adopting single TIFF as the file format and digitising images in colour. It was not possible to upload digital files produced by Document Supply Service without further development.

The DCM was not ready for books until 2005. Document Supply Service had identified the need to be able to keep digital files in a secure environment but it took some time for the Library to be in a position to enhance the system to permit uploading books. The enhancements included the capacity of loading digitised books in both single and multi-TIFF formats. The digital master images are saved as uncompressed TIFF files – single TIFF digitised by DAP and multi TIFF digitised by Document Supply Service. At the time the Library took a pragmatic approach to these different standards and it was agreed the Library would support a digitisation standard as well as an access standard. DCM was enhanced to enable Document Supply Service to save the multipart TIFF file on the DOSS as a digital master and the system would generate a PDF copy for public access. A link in the respective catalogue record would lead users to the DCM delivery systems and instead of a list of images appearing users would see a thumbnail image marked PDF. By selecting the icon users could access the PDF which they could then download or view online.

Users began reporting issues with downloading large files and at the same time Document Supply Service experienced difficulties in uploading large multi-TIFF files, i.e. files larger than 1 GB. The Digitisation and Photography section also participated in a number of small scale book and journal digitisation projects and had identified a number of changes required to the DCM system.

Further development of the DCM was undertaken in 2010-11. The changes introduced allowed the uploading of multi-part TIFF files and conversion to single page TIFF. This meant that all files in the system had a similar structure however there were unintended consequences. Although the Library scanned master images in colour at 300dpi, the display version and downloadable PDF version is in a much lower 72dpi resolution. For Document Supply Service this meant that some of the pages were not easily read. In these circumstances, a copy of the master file could be downloaded and sent to a client.

In 2010-11 as part of a redevelopment of Copies Direct, the Library decided to use the interlibrary loan management system Relais to handle the workflow for digitisation on demand and tracking movement of material within the Library. The system was developed to enable staff to forward requests to other sections of the Library for action. Workflows were created to forward requests to Preservation Services for assessment and conservation treatment, to collection managers for consideration of digitisation if required and to DAP if photographic quality copies or prints were required. As a result the workflow for digitisation on demand was standardised and automated to a high degree. In addition this workflow was not only available just to Document Supply Service staff but also to all of the special collection areas handling copy requests.

Document Supply Service administers and manages the Relais system³ and produces statistical reports on behalf of all collection users.

During this period there were progressive improvements in standards, workflows and equipment. Firstly, the Library implemented its cataloguing policy for electronic resources and adopted the practice of creating a second record for the electronic version of a title. This increased workload combined with the requirement for specialist cataloguing skills not available in the Document Supply Service meant this extra step in the digitisation workflow could not be performed by the staff in the section. Staff in the Collection Management Division took over the task of creating the catalogue record as well as the responsibility of uploading the digitised items to DCM. The workflow changes freed up staff resources and enabled Document Supply Service to concentrate on document delivery and digitisation.

The Library refurbished and re-configured the work area of the Document Supply Service in 2008. The Document Supply Service was relocated proving the opportunity to improve workflows. Prior to the move staff shared equipment which lead to bottlenecks in processing requests. The Library procured new flatbed scanners for each of the team members. Having a scanner at their workstations enabled staff to manage work processes and routines more effectively, removing bottlenecks and increasing the quantity of work completed. The Library purchased two SMA 21 overhead scanners to replace the Minolta 7000 scanners and a Cobra Book2net scanner for specialist scanning. This equipment was able to handle more complex/difficult material such as tightly bound items as well as rare and valuable material. The upgraded equipment meant the Document Supply Service was able to digitise collection material to higher standards, more efficiently.

Document Supply Service reviewed its digitisation standards in 2010 and made a further change by digitising book covers and illustrations in colour while the rest of the book would be scanned in black and white. Quality assurance (QA) conducted by peers was introduced to ensure the work uploaded to DCM was consistently of a good quality.

These technical changes allowed Document Supply Service to increase the quantity of digitisation by adopting higher standards while staffing levels remained the same.

Future Convergence

In 2011-12 the Library began a very ambitious project – the Digital Library Infrastructure Replacement (DLIR) - to “build a new digital library infrastructure to acquire, preserve and deliver digital collections” (Strategic Directions 2012-2012, National Library of Australia). The Library spent some time developing a set of requirements and went to

³ The Relais system is used by the National Library of Australia to manage its ILL/DD traffic. It handles all ILL/DD requests from libraries participating in Libraries Australia Document Delivery, Te Puna Interloan, OCLC and the National Library website as well as Copies Direct Requests. It manages the entire workflow from receiving the request to delivery of the material. The Relais system was progressively rolled out between 2005 and 2008.

market in 2012. Although two products were identified that will form part of the technical infrastructure, Tessella Safety Deposit box for archiving and DocWorks for OCRing, again the Library will need to develop additional components as part of this project.

The first phase of the DLIR project will redevelop the books and journal component of the Library's digital infrastructure. As part of this phase from December 2012 to July 2013 a project was undertaken to develop the standards and specifications; selection policies and procedures and new workflows for books and journals digitisation. The Library for the first time recognised not just the importance of the planned digitisation program undertaken by DAP based on curated selection, but also acknowledged the digitisation contribution by Document Supply Service, based on user demand. As a result the project examined both digitisation streams.

In the current environment, while book digitisation standards under the planned digitisation program are for colour and single page TIFF format, those for digitisation based on user demand are predominantly in black and white and multi-page TIFF format. These two separate standards have led in a few cases to re-digitisation of some material and convoluted workflows. Any books digitised by Document Supply Service which fall within the planned digitisation program will be re-digitised in colour. Due to these different standards, for some special collections material, collection managers prefer DAP to digitise the material to the preservation standard rather than Document Supply Service, as it means the material won't be handled and digitised a second time, leading to convoluted workflows. It has become apparent that adopting one set of standards for both streams of digitisation makes good sense and will be better for preserving collection material as well as improving workflows.

To digitise entirely in colour to some extent also simplifies the process. Staff will no longer need to think about when to digitise in colour and when not to. However this change of standards presents a challenge for the Document Supply Service as the software that has been used for digitisation for over a decade can no longer meet these requirements requiring a new solution.

With respect to workflows, changes are also inevitable. The DLIR development will offer opportunities to simplify many of the current practices. It is expected that it will remove the dependence digitisation has on cataloguing record creation and modifications, by automating the creation of the catalogue record for the digital copy as a by product of the file ingest process. In effect this will mean Document Supply Service will digitise and upload the images to the new Digital Content Management system and simplify the uploading process. The other change relates to delivery where different options are being explored. It may be that instead of digitising directly in multi-page TIFF format or manually compiling single page TIFF files into a multi-page TIFF document and then converting it to a PDF copy for delivery, Document Supply Service may digitise in single page TIFF format, uploading these images to DLIR. Then the new Digital Content Management system might be used to compile the images and produce the required files formats for delivery.

This project is just beginning, however it will provide opportunities to streamline workflows and produce digital content more efficiently and effectively.

The convergence of the two digitisation streams in terms of standards and workflows will provide opportunities to utilise staff resources more strategically to support both document delivery and digitisation. Once the barriers from utilising different standards and workflows have been removed, potentially the two areas could work together more closely to achieve better turnaround times for document delivery as well as meet the target of planned digitisation for books and journals. It won't be straightforward or happen overnight, but at least it will open up the possibilities for more efficient workflows and work arrangements across different areas of the Library.

This project will also migrate existing digitised content to the new platform and ensure all of the work Document Supply Service has undertaken in digitising content over the last 10 years is available for future use.

Lessons Learnt

This has been a long evolutionary story, over ten years in the making, from when Document Supply Service first considered the idea of digitisation on demand and took opportunities to experiment. There have been many lessons learnt during the period and many more are expected as the Library moves further into the digital library world.

The challenge is to seize opportunities to innovate, particularly when proposed activities or services help achieve the strategic goals of the Library. To digitise based on user demand as a by-product of ILL/DD contributes to building the Library digital collections and helps in preserving the Library's valuable collections.

It is important to take opportunities to influence outcomes and not to give up because of slow uptake of ideas and solutions, or obstacles encountered along the way. It took Document Supply Service four years before items they had digitised could be uploaded to the Library's digital repository. It also took a similar length of time to acquire more efficient equipment. The key is to take a long term view while dealing with short-term challenges.

Innovation cannot happen without experiment. Keep experimenting with new and different approaches as well as technologies as in the long run better solutions will surface. During the process, don't be afraid of making mistakes and discarding what does not work. Document Supply Service threw out quite a few files digitised in the early days that were not of sufficient quality. Through trial and error, a great deal was learnt along the way and knowledge and skill base in the section grew as a result.

Finally never stop looking for opportunities to improve services, think of the user, not just the patron who initiated the request, but consider other possible users and take opportunities to change and innovate. Document Supply Service has never stopped improving the services it provides to its users, both individuals and libraries, and is

continually striving for improvements to workflows and providing excellent customer services.

References

National Library of Australia. 2000. Direction 2000-2002., Canberra: National Library of Australia.

Jeong-Reuss, Lisa and Engram, Chesley. 2002. Preparing collections for digitisation: the National Library's Digitisation Project: A paper presented at the AICCM Book and Paper Symposium, Melbourne, 17-19 April 2002; located at
<http://pandora.nla.gov.au/pan/21336/20021011-0000/www.nla.gov.au/nla/staffpaper/2002/engram1.html>

Boston, Tony. Building a digital library at the NLA. 2002. Paper given at the OzeCulture 2002 conference, May 2002. <http://www.nla.gov.au/nla/staffpaper/2002/boston1.html>

Cathro, Warwick and Boston, Tony. 2003. Development of a digital services architecture at the National Library of Australia. Paper given at the EduCause 2003 conference.
<http://www.nla.gov.au/nla/staffpaper/2003/cathro1.html>