New Zealand

National Library of New Zealand Te Puna M□tauranga o Aotearoa – Wellington

1. General

Legal deposit legislation

Legal deposit legislation in New Zealand is administered by the National Library of New Zealand (NLNZ) through the National Library Act 1965. This Act requires that three copies of every book or publication published in New Zealand are deposited at the National Library of New Zealand.

In May 2003, the 1965 Act was revised with the passing of the National Library of New Zealand (Te Puna Mātauranga o Aotearoa) Act 2003. Legal deposit as applied to printed copies now requires 'a publisher of a public document to give the National Librarian, at the publisher's own expense a specified number of copies (not exceeding 3) …'. The National Library of New Zealand Act 2003 has mandated NLNZ to collect, preserve, and make available not only New Zealand's traditional paper documents, but also material in digital form (including websites and digital journals).

The legal deposit provisions apply to any person, group or organisation that publishes books, magazines, newsletters or any other work, for sale or free of charge, to any section of the public. One copy is made accessible through the general collections of NLNZ, and the other is preserved for use by future generations as part of the Alexander Turnbull Library, New Zealand's premier research library and a part of the National Library of New Zealand.
Digital legal deposit is expected to come into force in late 2005.

Digital preservation in NLNZ

Organisational embedding

NLNZ consists of seven units: three of these units: Electronic Services, Collection Services and the Alexander Turnbull Library (ATL) are involved in digital preservation. Each unit is managed by a director. The directors form the Senior Leadership Team and report to the Chief Executive/National Librarian.

Electronic Services provides national access to library and information resources by providing application and systems environment support to clients, and new ventures in digitisation and digital preservation. Collection Services provides access to the general collections, is responsible for collection development and collection management, and produces NLNZ's bibliographic services.

© Koninklijke Bibliotheek / IFLA / Saur - 2006

The ATL focuses on collecting and providing access to its research collections. Policy and Strategic Development supports NLNZ in the development of research and policy advice in its role as a key advisor to the Government on information management and delivery.

The Library's Electronic Services business unit is responsible for the development, establishment and implementation of the National Digital Heritage Archive (NDHA) Programme (see below under funding). There is a core team of approximately eight people, including staff seconded from other areas of the Library and contractors, working through the issues related to the NDHA.

At this stage it looks as though NLNZ will have distributed responsibilities for the range of activities associated with digital objects and their preservation. Once NDHA is implemented, a business unit with ongoing responsibility for its management will probably be set up. This will probably be within Electronic Services. The whole-of-domain or broad crawling of the .nz domain will probably take place within Electronic Services. Selection, acquisition and cataloguing/indexing of published digital material will probably be the responsibility of Collection Services. Appraisal, acquisition and arrangement/description of unpublished digital material will probably be the responsibility of the Alexander Turnbull Library. In Collection Services and in the Alexander Turnbull Library some environmental functions, for example, virus checking will to be undertaken prior to those functions. Selective and event web-harvesting will probably be undertaken within the Alexander Turnbull Library. It is likely that decisions on particular preservation strategies and their impacts (e.g., notions of acceptable loss) will be a shared activity between Electronic Services and the Alexander Turnbull Library. These activities have still to become fully embedded practice. See also Section II.2 for the organisational chart.

Funding

NLNZ is a government department and therefore all of its activities are funded by the New Zealand government. In May 2004, as part of its annual funding bid to government, the National Library received government funding to build a trusted digital repository now known as the National Digital Heritage Archive (NDHA) programme. An important component of this funding is that at the end of the NDHA project, NLNZ will receive an increase in baseline funding to help ensure scalability and sustainability of its digital preservation activities.

© Koninklijke Bibliotheek / IFLA / Saur - 2006

- provide enhanced access to digital information for New Zealanders, e.g., online databases, digital journals, and especially New Zealand content;
- collect digital resources, especially those relating to New Zealand and New Zealanders;
- ensure the long-term storage and preservation of New Zealand's online heritage;
- provide enhanced access to the Library's collections through digitisation.[86]

The Library has chosen a broad definition of the 'digital library' to encompass all the services and resources delivered in a digital environment along the lines of the Digital Library Federation definition of digital libraries as 'organisations that provide the resources, including the specialised staff, to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of, and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities'.

A Digital Strategy Implementation Committee has been established to monitor the implementation of the Digital Strategy and the Library's annual programme of work related to the Digital Strategy.

Recently, a Business Requirements Specification has been developed along with an interim Object Management System for the handling of digital material while the digital repository, the NDHA, is being developed. This includes issues of redundancy in a country where issues of geological stability are key. A Functional Requirements Specification will be developed in the second half of 2005. The current activity is a logical extension of previous work on a range of Digital Library activities including digital preservation (e.g., preservation metadata schemes/data model, persistent identifiers, structural and rights metadata, OpenURL etc).

Services
The mission of the NDHA will be to retain the digital objects that are collected by NLNZ in perpetuity in a structured, expansible and secure environment. The repository will provide long-term storage, preservation and access. It will also provide metadata management and digital object management. The repository is expected to be fully operational by 2008.

When operational, the repository must be fully accepted as a trusted repository for digital objects. It is expected that this will include not only material that reaches NLNZ through legal deposit or from donations of unpublished materi-

[86] Ross, Seamus. 2003.

© Koninklijke Bibliotheek / IFLA / Saur - 2006

agreements and defining the most suitable methods for processing digital objects for deposit at NLNZ.

Software and OAIS
For the purposes of digital preservation and the management of 'distributed' Archival Information Packages (in OAIS terms) the design of the NDHA is unlikely to be delivered in terms of one single solution. The library has core resource discovery and digital application software both of which run on an Oracle backend and are not currently designed for the extra functionality required to support digital preservation. Components of preservation metadata are currently in an SQL Server repository and the current object storage strategy uses standard UNIX file management tools in a SAN environment.

Preservation metadata is extracted programmatically from the objects, using a locally developed stand-alone application. This application provides data for the metadata repository and has been successfully integrated with the Library's Object Management System. The Library hopes to undertake a process shortly, which will align the Harvard University file format characterisation tool JHOVE with the Library's preservation metadata extract tool.

The Library is currently using the National Library of Australia's PANDAS web-harvesting tool for selective web-harvesting. It is hoped that a 'next generation' application for web harvesting will arise out of the International Internet Preservation Consortium activity and that a first iteration of this will be developed in 2006. The increasing use of Content Management Systems (CMS) and Digital Asset Management Systems (DAMS) for storing digital objects that institutions deliver dynamically to the web, will make it more difficult to capture web publications in the not too distant future. This will require significant investment in new technologies and necessitate the establishment of formal deposit arrangements with information creators and providers to help facilitate the collection process.

For the new system, NLNZ assessed the potential use of key emerging digital repository software, including DSpace and Fedora. However, it was uncertain whether these would scale to provide an enterprise class solution to a National Library's requirements to provide long-term storage and preservation 'in perpetuity'. As a result of this assessment NLNZ decided to go out to RFI for a preferred software supplier to develop a commercially viable and sustainable digital preservation and management software application. This process is still underway.

While NLNZ has yet to decide on the mechanics of being a third party host for the preservation of other agencies' digital materials, there is another wider

© Koninklijke Bibliotheek / IFLA / Saur - 2006

Metadata and metadata schemes

Metadata collection will largely be automated to make the process as consistent, reliable and auditable as possible. As far as possible metadata relevant to preservation is programmatically extracted from objects as XML, and output according to NLNZ preservation metadata schemes developed for the preservation metadata repository. However, it is still likely that some preservation metadata will need to be produced manually, although it has not yet been formalised which will be obtained automatically and which manually. At this stage NLNZ asks contributors to provide only minimal metadata with their contributions as part of the digital legal deposit process.

NLNZ's current thinking allows for preservation, descriptive and other metadata to be stored separately from the digital objects, currently in standard XML-enabled relational databases. NLNZ underlines the importance of having clear definitions as to what preservation metadata is and what it comprises, as opposed to other sorts of metadata and where the demarcation lines are to be drawn. Rights and permission metadata, bibliographic/descriptive metadata, and structural metadata (including METS) for example, are not considered to be preservation metadata. The approach NLNZ schemes use is fairly rigorous in defining what constitutes preservation metadata. The metadata considered necessary for preserving objects, is only a part of the total metadata known about an object.

The repository will store information on logical object level, file level, bit stream level and metadata (mainly to document changes to the metadata record for an object). No information will be stored on collections (except to relate collection objects to each other) and on non-digital source objects (which will be held as part of the descriptive metadata).

Preservation as an activity will rely on all metadata and not that deemed to be unique to preservation metadata. All metadata will have to be available in a unified and meaningful manner for all processes (resource recovery as well as preservation). The whole metadata record of an object is considered important, not just for supporting preservation, but also for providing information for continual and ongoing resource discovery and administrative management.

NLNZ has developed its own preservation metadata scheme, data model and XML scheme definition. This work has been mapped to the OCLC/RLG Preservation Metadata Working Group[89] and the NISO Z39.87 Data Dictionary,

[89] OCLC/RLG Preservation Metadata Working Group : See: Research Libraries Group. 2001. Preservation metadata for digital objects: A review of the state of the art. A

© Koninklijke Bibliotheek / IFLA / Saur - 2006

the type of material and the permission given. All kinds of access might be possible in future, although paid access is not being considered at this stage.

Under the new legal deposit legislation NLNZ may provide a specified number of copies but no more than three of a deposited document for use by the public (within or outside the library) but is not allowed to make the document available on the Internet without the express permission of the publisher/copyright owner. However, if a publication has been made accessible on the Internet without any authentication or commercial considerations NLNZ can make that publication available for access and use for the public via the Internet.

3. Preservation strategies

NLNZ is developing a long-term strategy for preserving its digital material in line with the requirements of its Act to preserve in perpetuity. NLNZ is aware that there are, as yet, no existing practical tools for the long-term management of digital objects, such as tools for emulation. Any given strategy depends upon reliable tools that can be used to implement that strategy. Without tools, the minimum preservation strategy for an object is to hold some form of agreed preservation master in anticipation of the future availability of tools, and/or have an original iteration of a digitally-born object to act as a reference copy for use with the tools.

NLNZ's current position is that the first priority must be given to 'the mechanics of possession' whereby NLNZ must be able to ensure the capture and storage, including some level of curatorial endorsement of the look and feel of an object. By their very nature preservation strategy issues must be secondary to the mechanical process at this stage.

Future strategies of NLNZ will be based upon automated processes carried out on bulk types of objects. It may be possible to restrict submission formats in order to normalise objects, but this cannot cover all of NLNZ's material. Normalising and migration have their limitations in a preservation sense, in that an object may ultimately become so removed from it's original self as to lose all relevant, or desirable look and feel. In order to respond to a dynamic and changing external environment any chosen preservation strategy must in itself be flexible and dynamic, if it is to be an adequate and enduring response. Possible strategies for NLNZ will be: restriction on submission (under consideration, because it is not yet clear what the implications might be for NLNZ and the donors); normalisation (under consideration, but only for text materials); migration and emulation. Migration on demand will only be applied as the implications become clearer. Current strategies are to a certain extent based on what is currently appropriate for implementation while allowing NLNZ to engage with

© Koninklijke Bibliotheek / IFLA / Saur - 2006