



NSLA Digital Collecting Forum Report

State Library of Victoria

Tuesday 19 – Wednesday 20 November 2013

Background

Without exception, NSLA libraries are in a state of flux with regards to digital collecting – adapting and improving workflows, acquiring or decommissioning equipment, adopting new technical infrastructure, cultivating new skills, undergoing restructure and review. At the July 2013 NSLA meeting it was agreed to hold a forum for staff from NSLA libraries to consider issues relating to digital collecting. The forum was designed to:

- Develop a shared understanding of digital collecting and its scope, including born digital publications, e-legal deposit, digital ephemera and digitised content
- Share information on progress, skills and approaches to digital collecting already underway in NSLA libraries
- Link strategic planning on digital collecting issues across NSLA libraries and project groups (particularly Digital Collecting, Digital Preservation, Pictures, Maps, Archival Collections, Copyright and Collaborative Collections)
- Disseminate and discuss the outcomes and proposed outcomes of the NSLA Digital Collecting Project (Stages 1 and 2)
- Identify priorities for collaborative work to feed into work planning for NSLA groups in 2014; and consider internal and external communication priorities

The forum ran for two days in November 2013 and was hosted by the State Library of Victoria, with co-chairs Justine Hyde (Director, Library Services and Experience) and Kate Irvine (Executive Officer, NSLA). There were 33 participants representing nine NSLA libraries.

Day one was comprised of a series of updates from represented libraries and project groups to provide context for joint discussion. A ten minute video update was provided by the British Library as a contribution to the forum. Day two included small group workshop sessions to tease out some of the issues raised and a final whole group session to agree upon priorities for collaborative work.

This report sets out the major themes of the forum and corresponding proposals for NSLA. These will be incorporated in the Digital Collecting Project's Stage 2 report to NSLA in February 2014.

1. Strategy: From Acquisition to Access

Digital collecting strategies have been implemented by NSLA libraries to varying degrees. The National Library of Australia's new Access & Digital Services branch, for example, has set out a seven-point agenda encompassing deposit standards, digital preservation, skills training, digital systems, discovery and delivery, digitisation programs and trial workflows. The National Library of New Zealand has appointed a Digital Collection Strategy Leader with a specific mandate to work with other areas in the library to adapt and improve approaches to digital collection, and to formulate workable policy. The library has adopted a nationally coordinated approach to the collection of born digital items, supporting other cultural institutions in their own collecting endeavours.

The British Library has identified 2020 as the year by which scholars will be routinely using its digital content in their work and practice as a primary evidence base. Its four-part strategy begins with bringing digital content into the library and cleared for use; engaging with digital scholars to establish their requirements; introducing services and tools to meet these needs; and finally raising staff skill levels in order to engage effectively with researchers.

Not all NSLA libraries have a comprehensive strategy set out in this way, though the State Library of New South Wales is currently developing a whole-of-library digital collecting strategy. The State Library of South Australia's strategic plan for 2014-2017 includes provision of robust technical infrastructure and a commitment to statutory responsibilities to collect born digital material, and prioritises development of a digital collection policy. The State Library of Victoria is set to commence major reviews of its service delivery and collection development areas, including its electronic content strategy. The State Library of Queensland has drafted a content-focused and format-agnostic Content Development Policy.

On an operational level, the Northern Territory Library has included "digital transformation" in its business plan for the first time. The State Library of Queensland has established a Digital Library Project to look at digitisation, digital preservation, and possibilities for distributed collections and future-focused research. The State Library of Western Australia is reviewing processes and infrastructure with a long-term goal of more automation and less intervention.

In some libraries, approaches to digital collecting are connected to – or offset by – priorities for digitisation. Many libraries have undergone mass digitisation projects with a focus on photographs and oral histories. LINC Tasmania has drafted its digitisation plan with an emphasis on early and vulnerable formats.

Proposal from the forum:

To draft an overarching NSLA Digital Strategy encompassing digital collecting, digital preservation, and staff skills and capability.

Relevant NSLA project work underway:

The Stage 2 report from the Digital Collecting Project will be presented to NSLA in February 2014 and will include recommendations regarding NSLA-wide approaches to digital collecting.

2. Staff Skills and Capability

Without doubt, this was the area of primary interest and concern for participants in the forum. Broadly speaking it was felt that library staff are not yet equipped with the skills, confidence and knowledge to select, collect, acquire, store and preserve digital material effectively. Results from the 2013 State of the Public Service survey, for example, indicated that only 67% (two thirds) of staff at the National Library of Australia felt confident in their digital capabilities and only 27% felt their capabilities had improved in the last twelve months.

Many libraries have established working groups to tackle specific areas relating to digital collecting. The State Library of Victoria has merged its digital collecting and digital preservation working groups, and set up a copyright working group. The State Library of New South Wales has formed a working group to address the scope, speed and scale of digital collecting, with participants from across the

library including original materials, learning services, public libraries, digital library services and access and information. The State Library of Queensland has appointed an e-Content Team dedicated to describing born digital material. The National Library of Australia has established a Digital Collecting Support Team to trouble-shoot digital operational problems encountered by staff on a daily basis, and has expanded the role of catalogue staff to include digital skills training based on the European Commission's DigiCurV package.

According to the British Library, the current skills gap includes not only the practicalities of digital collecting and making digital material accessible, but an understanding of how library users are and will be using digital content. The British Library offers a digital scholarship training program for staff, drawing on both internal and external instructors to deliver a set of fifteen courses ranging from digital mapping to social media, presentation skills, crowd-sourcing techniques, working with wikis, data visualisation and geo-referencing. Take-up has been very high. It was noted that "the advances we've been able to make in digital scholarship were only possible with a dedicated digital research and curation team – that mixture of fundamental curatorial understanding combined with quite sophisticated digital capabilities has been absolutely critical".

Workshop groups on the second day of the forum began to identify some of the skills necessary for digital collecting and how they might be acquired. These included an understanding of:

- Vocabulary and key terms
- Formats and specifications
- Legal issues/copyright and intellectual property laws
- Digital preservation (including carrier transfer and format obsolescence)
- Algorithms and linked data
- Collection content

Essential characteristics for future roles were considered to be: confidence in problem-solving and capacity to deal with ambiguity; content expertise; librarian or archival skills (to varying degrees); core ICT skills; and a capacity to build networks.

Part of the challenge for libraries comes with the need to constantly adapt workflows to deal with the changing type and volume of material coming in – changes which often require training and funding. The British Library, in order to fund the development of scalable workflows for the collection, preservation and delivery of digital material, recently "top-sliced" half a million pounds from its acquisitions budget. Given that many libraries are facing budget constraints it was felt that cross-organisational roles may be more viable than new roles in every institution. Staff exchanges and digital scholarship courses were popular suggestions, drawing upon the expertise already within NSLA. Possible new or shared roles might include:

- Data quality/standards
- Digital archivists/curators
- System management
- Preservation planning (including technology watch and format analysis)
- Linked data
- Delivery innovation
- Social media curation

Proposal from the forum:

To develop a NSLA digital scholarship course or shared modules; to draw on expertise already within NSLA libraries and form a community of practice; to establish a program of staff exchange and consultation; to investigate the possibility of shared roles in specialised areas.

Relevant NSLA project work underway:

The Pictures Project will be submitting a business case to NSLA in February 2014 for employment of a business analyst to investigate current workflows in NSLA libraries for processing large born-digital pictorial collections, and to make recommendations for more streamlined workflows. These recommendations will likely be applicable to other collection areas and format types. The work package was proposed and endorsed by NSLA in November 2013.

The Digital Collecting Project, as part of Stage 2, is undertaking an analysis of workforce requirements and staff confidence and capability with regards to digital collecting. Findings will be included in the Stage 2 report to NSLA in February 2014, complementing an organisational capability matrix and report from the Digital Preservation Project.

3. Technical Infrastructure and Difficult Formats

Effective digital collecting calls not only for apposite skills but for tailored, reliable infrastructure to enable preservation, storage and delivery of digital material. Most libraries identified storage and delivery of digital content as a major issue.

The State Library of South Australia noted that storage capacity was a problem particularly for re-formatted video, film and pre-press digital newspaper files (discussions are underway with APN News & Media to establish a workable model). The Northern Territory Library has calculated that between fifty and one hundred times the amount of material currently stored digitally will need storing in the next twelve months. Provision of a stable platform for digital content is a priority for the NTL in 2014, particularly given its role as host of multiple 'digital keeping places' for Indigenous community content and given the rapid degeneration of hardware in very hot and humid climates. Cloud storage has been utilised as part of a disaster preparedness plan.

The National Library of New Zealand is notable for its comparatively sophisticated digital infrastructure incorporating the Indigo system for ingest and the Rosetta system for preservation. The National Library of Australia is mid-way through its Digital Library Infrastructure Replacement (or DLIR) project, with a Safety Deposit Box feature due to be installed by the end of 2013 – allowing for fully automated ingest and preservation of some digital materials. Digital and physical preservation teams are now co-located in the library.

All NSLA libraries are encountering difficulties with specific formats such as CAD drawings, games, databases, email accounts and social media which require customised preservation and remain inaccessible without specific (often outdated) software or hardware.

The State Library of New South Wales is a forerunner in the collection of social media, partnering with the CSIRO as part of the Vizie trial alongside ABC Innovation, the National Mental Health Commission, police and counter-terrorism departments in New South Wales. The Vizie program searches multiple platforms for defined keywords and flags relevant tweets, blogs and posts – the

library is the only collecting institution involved in the trial, and uses a collecting framework for social media based on its broader Collection Development Policy.

The British Library has begun collecting born digital archives and manuscripts – the archive of poet and broadcaster Wendy Cope, for example, containing 40,000 emails – as well as born digital video and audio recordings. Geospatial data is collected directly from ordinance surveys and displayed using the cloud-based service ThinkWare. Doctoral theses and metadata are harvested from institutional repositories across the UK via the Electronic Theses Online Service (EThOS), with 70% of all published theses collected to date.

Many libraries are placing digital items on in-house servers without necessarily having a strategy for long-term storage, preservation and access. Unsurprisingly, funding and lack of expertise were identified by participants as the primary barriers to acquisition and use of effective, reliable technical infrastructure for digital collecting.

Proposal from the forum:

To continue development of a system-agnostic technical format registry to assist in making decisions about which file formats to collect and how to preserve them, regardless of the technical infrastructure in use; to share expertise regarding the collection and preservation of difficult formats.

Relevant NSLA project work underway:

The Digital Preservation Project has reached ‘Stage 3’ of its technical registry project, and is developing a full business case including a proposed funding model to be presented to NSLA in February 2014. The project group has also discussed collaborative possibilities with its CAARA counterpart, the Australasian Digital Recordkeeping Initiative, which is undertaking research into preservation of difficult formats and supports the development of a technical registry.

4. Legal Deposit and Permissions

Issues around copyright, intellectual property, creative commons, privacy and legal deposit are an intrinsic part of digital collecting policy. The sheer volume of digital material and the impetus to collect it as soon as possible – given its generally short lifespan online – means that some libraries are forced to collect first and establish copyright ownership or access conditions later. It also means that a great deal of material which would traditionally have been deposited at the library in hard copy as legal deposit is being produced in digital form and not deposited at all.

The State Library of Western Australia was successful in a bid for digital legal deposit legislation this year, though specific regulations are yet to be finalised. Likewise the State Library of South Australia has been advised by the Crown Solicitor that the state’s Legal Deposit Act can be interpreted to include born digital material, but that there is a need for new regulations. Legislation in Queensland can also be interpreted to include digital materials. The National Library of Australia remains hopeful that e-legal deposit legislation will be taken up by the federal government.

In the UK legal deposit legislation was passed for digital publications in April 2013, allowing the British Library to collect e-books, e-journals, and digital newspapers. The Finch Review of 2012 also recommended that all UK research outputs be made available through open licences. The British Library has embarked upon its first crawl of the UK web domain (approximately one billion pages) and will repeat the crawl annually. Agreements have been signed with three major publishers to

collect e-journal articles, with a further ten to fifteen publishers to sign this year and 120,000 articles collected to date. The library is conducting pilots for collection of mobile app material, social media, and documents embedded within websites (primarily PDF, Word and Powerpoint).

Preliminary results from the NSLA Collaborative Collections Project legal deposit survey indicate that digital legal deposit is in its early infancy in most NSLA libraries, though the National Library of New Zealand is well ahead. The State Library of Queensland has established a Deposit Module in DigiTool, allowing publishers to self-deposit files, select metadata and choose access conditions – over 460 legal deposit e-books have been ingested to date. The survey found that publishers expect libraries to be collecting digital legal deposit material, despite current practice, and noted that the role of PANDORA at the National Library of Australia needs clarification in order to delineate between local and national responsibilities for collecting websites – government and otherwise.

Web harvesting and archiving was raised as a particular issue, both practically and legislatively. The State Library of Queensland is seeking blanket permission for web archiving, as is the State Library of South Australia for state government websites. It was noted that the impact of AusGOAL has been minimal to date. The National Library of Australia is conducting a prototype trial for providing access to gov.au domain material this year, and will collect feedback to scope requirements for the DLIR to harvest this material in future. The British Library reported collecting 14,000 websites since 2004 on subjects ranging from politics to the environment and the arts.

Proposal from the forum:

To develop a shared approach to e-legal deposit, with shared messages to communicate its importance. To consider increased support for AusGOAL.

Relevant NSLA project work underway:

Members of the Copyright Group have contributed to submissions from the Australian Digital Alliance (ADA) and Australian Libraries Copyright Council (ALCC) to the recent ALRC Inquiry into Copyright and the Digital Economy, with strong advocacy for introduction of a fair use provision. The Collaborative Collections Project is drafting a report based on its recent legal deposit survey across NSLA libraries, for submission to NSLA in February 2014.

For full details of individual library updates, contact the NSLA Program Coordinator at blemon@slv.vic.gov.au

Report author:

Barbara Lemon, NSLA Program Coordinator

3 December 2013