

You Can't Manage What You Don't Understand: The role of a digital systems stock take in protecting future local government born-digital archives

Natalie Dewson

Digital information at risk

'Born-digital' refers to information originally created in digital format. Born digital information is inherently fragile, which makes it a real challenge to manage. Unlike paper records that can sit untouched for decades and remain accessible and usable; records stored on digital media can become inaccessible in only a few short years through obsolescence or media degradation.

There are three types of obsolescence:

The physical carrier of the record becomes obsolete or damaged

For example, standard media of the 1980s such as floppy discs are no longer commercially available. Current media such as USB sticks, CDs and DVDs will also become obsolete. They can also degrade or become damaged.

The hardware needed to access the record becomes obsolete

Both the drives needed to read the media and the computers required to operate them have become obsolete due to rapid technology changes. For example, it is hard to obtain working disk drives. Most new computers have a usable life of only three to five years and are not fitted with floppy disk drives.

The software needed to access the record becomes obsolete

This includes both the software to read and write the record and the operating system on which the software ran. For example in the 1980s WordStar was widely used, but is no longer available. The fact that WordStar used a closed data format makes continuing accessibility even more problematic.¹

The last Government Recordkeeping Survey held in 2010 found that 49

per cent of local authorities held digital information they can no longer access:² this is information that is no longer available for business or historic purposes.

MW LASS Ltd and the Archives Central Project

MW LASS Ltd is a Council-Controlled Organisation established in 2008 by local authorities in the Manawatu-Wanganui region. The eight member councils collaborate on projects with the goal to improve services and reduce costs. Councils can choose their level of involvement based on their business needs.

All councils are responsible for making sure certain types (or classes) of information remain permanently available to their staff and the community as archives. These information types are identified in the Archives New Zealand List of Protected Records for Local Authorities.³ Councils may have additional business value information they also wish to keep long term.

This responsibility created an opportunity for MW LASS councils to join forces on a project to safeguard their permanent-value information. A database was launched in 2010⁴ followed by a Feilding-based, shared archives storage and research facility for physical records in 2012. Councils can choose their level of involvement (storage and database; or database alone). Councils outside the region are also able to opt in.

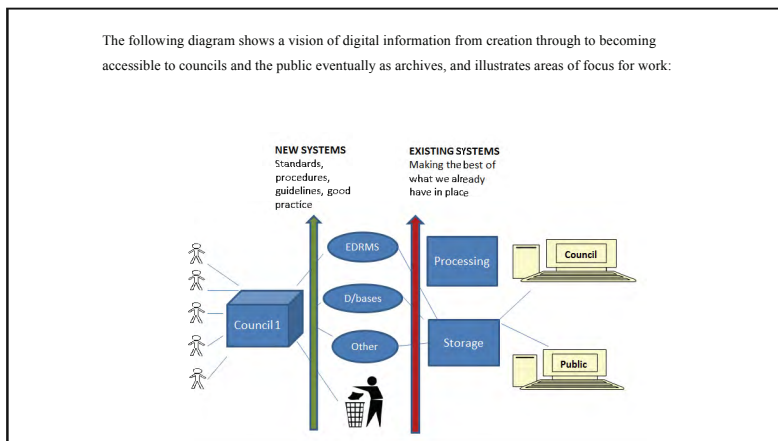
With the paper record now secure, attention turned to the archives of the future – the ‘born digital’.

Born-digital programme

Initial investigations at participating MW LASS councils corroborated the results of the 2010 Government Recordkeeping Survey. In 2011, a report commissioned by MW LASS Directors advised them on the archival implications of ‘born digital’ records maintained in each council. Referring to interviews with a selection of staff responsible for the creation of valuable born-digital records, the report concludes there is a very real risk that born-digital information won’t be available in the short term, let alone permanently.

MW LASS member councils decided to turn this challenge into an opportunity to collaborate on a programme that would safeguard born-digital records, and in 2012 I was engaged to co-ordinate work in this area. My brief was to source or create good practice standards for the creation and management of born-digital information; and to identify what a digital archiving future would look like for MW LASS councils.

The following diagram shows a vision of digital information from creation through to becoming accessible to councils and the public eventually as archives, and illustrates areas of focus for work:



The two arrows show intervention points where good practice can be applied to records – the first is when new systems are launched; the second is to those that already exist.

Managing digital records is a big challenge. So, where to start? As a first step I sought advice from the Digital Continuity Team at Archives New Zealand. A crucial point they made was that *we couldn't manage what we didn't understand*. Before we did anything we had to understand what born-digital information councils had and its technological dependencies through a preliminary systems stock take exercise.

Benefits of a stock take

A digital stock take has clear benefits. For instance, any future business case for a digital archive approach will have to be supported by strong business drivers and an understanding of the risks to digital records. Also, once key risks are known, appropriate standards and good practice can be identified to target them.

A stock take also informs work priorities. It is not achievable for participating councils to change how they handle all their digital information. The greatest chance of success is to deal with obsolescence one small chunk at a time dealing with the areas of greatest risk, first.

It is also an excellent opportunity to educate. Information

management has its own language that is not well understood by non-practitioners: stories cut through barriers and demonstrate through real world examples that obsolescence affects business activities now.

Getting started

I began by familiarising myself with the big picture of digital continuity, and understanding how digital information is broadly managed at each participating council.

I attended an Archives NZ training course to learn more about digital continuity issues. International archives institutions have also published useful background information; for example The National Archives of the United Kingdom have issued step-by-step guidance for managing digital continuity.⁵ I also examined New Zealand and international case studies of how other organisations manage and preserve their born-digital information to identify common themes.

Following this I undertook an environmental scan in consultation with council's Information Technology (IT) leaders. Because a number of councils within the MW LASS are small, IT leaders frequently have dual responsibility for born-digital information and the technologies that enable their access and use. Starting a dialogue with these people gave me the opportunity to win support for data collection, and it also enabled them to identify opportunities for their work. These include:

MW LASS Councils work together increasingly, and use open standards
Knowing what information we have, what it is stored on and how it
enables the business, opens up opportunities for data sharing
A collective voice holds vendors accountable for making systems easy
to exit.

Methodology

Due to time and resource constraints (no dedicated budget and data collection had to be juggled with competing work priorities) it was not possible to do an in-depth analysis of all business functions producing permanent-value information. I decided instead to take an in-depth look at the selection of functions that had been briefly looked at during the initial investigation in 2011. These were:

- Building consents
- Resource consents
- Cemetery management

- Infrastructure (water, storm water, roads)
- District plan.

Core digital systems, their accompanying paper records (where these existed) and information held on portable storage media were looked at. File formats were out of the scope of this study.

Questionnaire development

A reference group set up to act as an advice touchstone for the duration of the programme was consulted with to get agreement on the desired outcomes. These were:

What is our 'day one' for accessible digital information? I.e. when was our earliest recoverable information created? – this required identifying current line-of-business systems and the software needed to access and manipulate the information in them, querying whether there was still accessible information in them. I did the same for their predecessor systems and working backwards until records were created solely in paper format.

To what extent are these accompanied by paper records? I had an initial assumption that many early digital records would have paper equivalents, but this needed to be tested.

What are the major risk areas?

Can we tell which systems are obsolete or degraded? What is the feasibility of recovery, and what will it cost?

These desired outcomes informed the questionnaire design which covered areas such as the physical condition and age of the system and software, whether it is active or inactive, accessibility of the information inside it, existence of backups etc.

Data collection approach

In terms of how to locate, survey and priorities digital holdings, the resource I found the most helpful was an OCLC publication detailing the first steps for managing born-digital content received on physical media.⁶ Written for a non-expert audience, the methodology is designed for people and organizations with limited resources. I didn't follow it to the letter, but adapted it for my own needs.

This is the data collection approach:

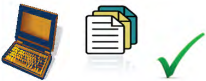
- a) All current official line-of-business systems were identified and broadly described, and their dates of commission recorded on a spread sheet. For each current system, their electronic predecessors were traced backward until recordkeeping systems became paper-based.
- b) Of the systems listed at each Council, at least two were selected for closer analysis, including descriptions of their supporting hardware and software.
- c) Discussions were held with Records Managers to find out more about the records stored in these systems: particularly whether they were accompanied by physical records.
- d) Staff who produced these records were then interviewed to get further contextual information, and to identify any records they were aware of that were stored on portable storage media.

Limitations

The findings apply only to the subset of council functions that were examined: an analysis of other functions may identify further issues we are not aware of. A small amount of born-digital information on portable storage media (e.g. VHS video, CD Rom, reel-to-reel film) was found; but note that a complete search of work spaces, paper files, storage areas etc. to find more was not undertaken.

Findings

Through the investigations we learnt the following:

Pre 2000's	Post 2000's
	
Digital information is gone, but there are paper equivalents.	Information is created increasingly digitally only. Once it has gone, it's gone forever.

Some information is lost or nearly gone

For example, there is a lot of CCTV footage used to manage infrastructure assets that is degrading or can't be played.

Councils keep everything but you can't find what you need

Systems are locked-down by vendors meaning the information they hold is inaccessible. This has happened where software licenses used to manipulate data have expired or information has been saved on unusual file formats and can only be read on the old system. Also metadata is inconsistently applied to some types of information.

*Not all information is coming across when data is moved**Information is lost*

Information is routinely moved from old to new systems, but not always successfully or completely. Consequences experienced by councils include: partial transfer of information, or complete loss of subsets of information.

Lessons learned

I underestimated the time required for data collection. Initially timetabling two weeks for each of the five participating councils (two months overall), the process took nine months in total. To a great extent this can be attributed to the 'locked shed' analogy. Where physical records are stored in a locked shed, access is a minor hurdle – you simply call a locksmith. In the digital world, people's heads are the 'locked sheds'. They have work of their own to do and can pose a greater challenge to extract information from than a physical storage environment. Patience, communication and persistence help.

Staff interviews didn't reveal as much about the systems as I hoped. Along with the inevitable staff turnover and resulting loss of institutional knowledge over the years, people's recollections tended to focus on the information itself rather than the systems they were stored in. Corporate paper files held at Archives Central did help fill in some gaps, although few records of this type had been retained.

I wanted to identify a 'burning platform' that would spur councils into action as councils are not subjected to the information audits that might highlight these. While councils have a legal obligation to care for the information on the Protected Records List, it would have been useful to compare this with what the councils themselves identify as business-critical; using the findings to highlight opportunities for cost savings and work efficiencies in a more meaningful way.

Outcomes

We learned that you don't need to invest a lot of money or have a team of experts on hand to get meaningful results from a stock take; but you do need someone who can dedicate time to see it through.

Now that MW LASS councils have a good sense of where the key recordkeeping systems reside and the main risks posed to the quality of the business information in them (metadata, migration and vendor lock-in), they are using this knowledge to make improvements to ensure it remains usable and accessible long term.

To begin this first stage, national and international good practice in migration and metadata has been identified and signed off by councils as recommended standards, and they are currently looking at ways to embed their consideration into system procurement projects. Metadata gap analysis tools for Electronic Document and Records Management Systems (EDRMS) and business systems have been developed, and the Archives Central team is supporting councils in its use. The team is also advocating the development of exit strategies from systems and software: this is about getting information out of old systems with the minimum of stress and cost.

This work will be an incredibly useful basis on which councils can prepare for a digital archiving future. An important first step towards this ultimate goal will be to build interest in digital continuity among the local government sector. MW LASS looks forward to its part in this step in the journey.

Endnotes

1. National Archives of Australia. "Obsolescence – a key challenge in the digital age". National Archives of Australia. Accessed May 15, 2014. <http://www.naa.gov.au/records-management/agency/preserve/e-preservation/obsolescence.aspx>.
2. Dudding, Amanda, and Kris Mayo. "Government Recordkeeping Survey Report 2010 – Local Authorities". Archives New Zealand. 2010. Accessed May 7, 2014. http://archives.govt.nz/sites/default/files/Publication_Local_Authorities_Report_0.pdf. P3.
3. Archives New Zealand. "List of Protected Records for Local Authorities". Accessed May 8, 2014. http://archives.govt.nz/sites/default/files/list_of_protected_records_for_local_authorities_0.pdf.
4. MW LASS Ltd. *Archives Central*. MS LASS Ltd, <http://www.archivescentral.org.nz/>
5. Digital Continuity Project. "Managing Digital Continuity". The National Archives, 2011. <http://www.nationalarchives.gov.uk/documents/information-management/managing-digital-continuity.pdf>
6. Erway, Ricky. "You've Got to Walk Before You Can Run: First Steps for Managing Born-Digital Content Received on Physical Media". OCLC. 2012. <https://www.oclc.org/content/dam/research/publications/library/2012/2012-06.pdf>