“A hard drive! Box it and put it on the shelf, Elizabeth.” First steps in digital forensics in a very small archives

ELIZABETH CHARLTON¹
Society of Mary – Marist Archives, Wellington, New Zealand

Abstract: Digital material cannot survive with benign neglect, it needs active management. Digital material is lurking in boxes, being actively collected in information heritage institutions – big and small, but no practical training is offered in-country in how to deal with it. Large, national institutions are working on their own processes, but are they scalable?

This study documents how it is achievable for very small archives to embark on managing digital material found on legacy media through the use of digital forensic practices and to use that knowledge to underpin acquisitions of current born-digital material.

Background

A naked hard drive arrived in our Archives from Fiji in 2013 and was simply put on a shelf. The elderly Marist brother who was then in charge of the Archives indicated that at least a decade would pass before it would be processed. I was volunteering at the time.

Experiences of the last decade – with theory and obsolescence of hardware and software – had attuned me to the precariousness of the digital world;

- I had an external hard drive for back-up indicating signs of failure. I was able to successfully move and retain all my files.
- My information studies had records management as a major focus. This emphasized the records continuum model created by Frank Upward at Monash University, Australia and the four characteristics of authenticity, reliability, integrity and usability specified in the records management standard, ISO 15489.
- Another paper expected us to make a video. My video camera requires firewire to download footage and my current laptop does not have one. The video editing software I was used to (and had) would not run on Windows 7. Luckily, I still had a usable laptop with Windows XP and a firewire port.

I had the feeling that just leaving the drive on the shelf was not proper management but I did not know what was.

Introduction

With this as my context, I continued with personal learning on the subject of digital preservation. From listserv subscriptions, I was introduced to the OCLC Demystifying Born Digital Reports series and learnt that the SAA Manuscripts Section had started the Jump In Initiative using the first report.² I followed

¹ The author would like to acknowledge the scholarship awarded by the Ian McLean Wards Memorial Trust to conduct this research.
² Manuscript Repositories Section, "Jump In Initiative," Society of American Archivists, last modified 2013, accessed September 25, 2016, http://www2.archivists.org/groups/manuscript-repositories-section/jump-in-initiative; R. Erway, You've Got to Walk Before You can Run: First Steps for Managing Born-digital Content Received on
along. This was starting small by creating an inventory of the legacy media in the collection. It at least meant that I found out what we had and gave an indication of the volume of storage that would be required if everything were to be retained and could be accessed.

In 2014, I organized a professional visit to the Library of Congress to learn more about the levels of digital preservation. During the meeting I was introduced to the model of hardware shared within a community at a public library. Admittedly this was for video, but the seed was planted as to whether something similar could work for floppies or hard drives that archives receive.

A few months later, at the joint ARANZ/ASA Conference in Christchurch, two digital analysts from Archives New Zealand spoke on "Legacy Digital and Outreach @ArchivesNZ". This outreach catered specifically for government departments. In particular, Mick Crouch asserted that it is important not to wait for a finished solution for digital preservation, but to deal with what we have now; using the tools we have now. The conference was closely followed by a presentation from Cal Lee from UNC Chapel Hill who introduced BitCurator to the New Zealand archival community.

Apart from one master’s paper in digital curation, no training was offered in New Zealand in digital preservation, until late 2015 when the first DPOE course was run with about 30 attendees. There is no practical guidance regarding digital archives pitched towards very small archives.

In fact, within the broader information management and services profession, managing digital material had been identified as a skill gap as indicated in the needs analysis conducted in 2014 for the New Zealand Qualifications Authority’s mandatory review of qualifications for Information Management and Services.

---


5 M. Crouch and R. Spencer, "Legacy Digital and Outreach @ArchivesNZ" (presentation, Connecting Past Present and Future. ARANZ/ASA Conference, Christchurch, New Zealand, October 1, 2014).

6 C. Lee, "All Bits Considered: Caring for Collections when Bitstreams Matter" (lecture, Archives and Records Association of New Zealand (ARANZ), Library and Information Association of New Zealand (LIANZA) and the Institute of IT Professionals New Zealand (IITP), National Library of New Zealand, Wellington, New Zealand, October 13, 2014).


In the second edition of *Preserving digital materials*, Ross Harvey wrote,

> Discussion about digital preservation has too frequently been couched in terms that apply only to large well-resourced institutions, with little regard for the requirements of smaller organizations and of individuals.\(^\text{10}\)

Harvey goes on to discuss the issue of scalability of solutions and ends by urging for the creation of guidelines for small institutions;

> These institutions urgently need guidance in the form of precise and concise directions that can be readily implemented, perhaps as workflows applicable to day-to-day operations.\(^\text{11}\)

I did not see that any solutions were forthcoming in New Zealand; so to work out what to do I would have to do it myself and then start sharing this information with the wider community. This gave me the impetus to apply for the 2014 scholarship from the Ian McLean Wards Memorial Trust, which I was awarded.

The project, to answer a professional problem, was of a personal nature in that it was conducted outside of work hours, though the Society supported my absences for training. While I reported regularly to the Trust on my progress, I had no direct supervision.

**Problem Statement**

The problem posed was quite simple – what do I need to do to manage the hard drive sitting on the shelf and how do I do it?

I believed that I could not be the sole very small archives in New Zealand to hold such material and that by working out a solution for me; I would consequently also provide a way for other very small archives. That is how I came to set the project’s parameters. I wanted to answer

- Can we manage legacy media? What do we need to be aware of if we are going to take up this challenge?
- If we do not know what is on the legacy media, how can we appraise without making irreversible changes?
- If the material is of value, how do we remove it without making irreversible changes?
- What could be an equipment set-up that could be shared among smaller institutions?

Institutional needs differ and so approaches differ. I wanted to create something that would be workable for very small institutions that do not have regular IT support and staff with little digital knowledge.

I determined that the following aspects would be out of scope;

- Arrangement and description of selected material
- Access (by users)
- Long-term storage systems

**Methodology**

As a former teacher I was familiar with the principles of action research and this is what I selected as it is useful in addressing issues that arise from professional practice and it is a practical method of research.

---


\(^{11}\) Harvey, *Preserving Digital*, 214.
Action research is appropriate “when you wish to achieve understanding and change at the same time”.  
According to Wallace, action research “… nearly always arises from some specific problem or issue arising out of our professional practice… It is very problem-focused in its approach and very practical in its intended outcomes”; the research is directly relevant to the participants and results in improving practice.

It offered a structure that was particularly applicable to the iterative nature of this research in that it both entails a process of planning, acting and evaluating against the project goals and allows for flexibility in adjusting project parameters.

From presentations and readings I was convinced that current tools and processes would fit my needs and that it would be relatively straightforward to implement. To answer the questions I initially framed my research in three stages;

1. The selection of the equipment (hardware and software) required to access and process the digital information for preservation.
2. The testing of the preservation process.
3. The creation of workshop materials documenting the preservation process

The purpose of the first stage was to ensure that the equipment selected was the most up-to-date available at that point in time; this reflects Mick Crouch’s recommendation mentioned earlier. The second stage’s focus was deepening my understanding of the processes involved in digital preservation, while the last stage was to embed the improvement into my own practice - the best way to learn is to teach - by creating the materials for and facilitating a practical workshop for interested parties who also wanted to better their understanding of this domain.

Action research offered a structure that was particularly applicable to the iterative nature of this research, in that it both entails a process of planning, acting and evaluating against the project goals and allows for flexibility in adjusting project parameters.

After visiting Archives New Zealand and the National Library of New Zealand, I realized that working with digital objects on legacy media was not as frequent or as widespread in New Zealand as I had originally thought. The two institutions share some digital forensic equipment such as a Kryoflux. While the National Library’s digital preservation program is well established, conversation indicated that its model was not scalable downward to the amount of resourcing and knowledge available in very small archives.

I thus decided to widen my discussions to include digital archivists at Georgia Institute of Technology, Georgia State and Emory Universities, the Library of Congress and the Smithsonian Institution Archives in addition to attending DigCCurr Professional Institute.

---

15 Crouch and Spencer, "Legacy Digital."
16 S. Knight, interview by the author, Wellington, New Zealand, May 6, 2015.
Furthermore, upon surveying members of a long-standing informal group of lone arrangers in Wellington and meeting with a few of them to gain more in-depth information about their practices, I realized that there was a lack of awareness amongst colleagues not just about legacy media but also more general computer knowledge, such as an understanding of network drives.

These aspects meant that addressing points 1 and 3 of my initial research framework would entail more focus and detail than the scholarship funding would cover, so I concentrated my efforts on the testing of a preservation process that would be manageable for very small archives.

Results

I achieved the aim of action research that is understanding and change at the same time. I also answered the questions I had posed.

To deepen my understanding of the processes involved in digital preservation, I first attended DigCCurr Professional Institute, where I was presented with an opportunity to try out software and equipment and to discuss optimal hardware configurations for use as a shareable digital forensics station. With that information in hand, I was able to answer my last question about a shareable equipment set-up, purchased upon my return. Hands-on training followed in the Digital Forensics: Advanced class at the 2015 Annual Meeting, held in Cleveland, OH.

Using that newly gained knowledge, I tested “guinea pig” legacy media and wrote an article documenting the process. 17 This experiment answered my questions on how to appraise and how to move material of value without making irreversible changes. Practicing on my own floppies assisted me in estimating how

much time it would take to image and complete an initial processing. I then applied the process at work, allowing me to check the accuracy of my time estimation. It was within predicted range. With the quantity of digital material I have awaiting processing, I believe that this volume is manageable and can be integrated in my workload without compromising other tasks. The benefits of appraisal in managing my institution’s legacy material made the cost of learning advanced skills worthwhile. These skills are transferable in that the processes prepared me to work also with modern acquisitions.

Keeping the legacy media focus, I drafted workflows and created instruction sheets for individual processing steps using the selected equipment. I was then able to adapt the legacy workflow to apply to born-digital content now arriving in the Archives.

The project afforded opportunities to work through issues that arose and options to resolve them. Since I worked on this project without any immediate support, I demonstrated that it is possible for very small institutions to embark on managing digital archives in a sustainable manner both in terms of staff time and resourcing. In fact, I proved that it is possible to implement digital forensic practices in a lone arranger environment thereby answering affirmatively my first question whether lone arrangers can manage legacy media. However, the greatest benefit from this exercise was the transferable skills gained.

Findings

The findings from the research fall into three categories, those applicable to lone arrangers in general, those applicable to my repository and lastly those applicable to the New Zealand context.

Lone Arrangers

My journey through defining what is achievable for me and my institution has highlighted challenges for lone arrangers in their management of born-digital archives:

- Acquisition of technical knowledge
- Isolation
- Budgetary constraints
- Time constraints

While some of these challenges are not unique to lone arranger environments, they become more salient within this context.

The benefits of appraisal in managing my institution’s legacy material made the cost of learning advanced skills worthwhile. This newfound knowledge will assist with clearing the backlog and has prepared me to work with modern acquisitions. Since the digital environment is here to stay, lone arrangers cannot be left as ostriches with their heads in the sand. If lone arranger institutions include born-digital material within their collecting scope, then these skills are just as necessary as all the other facets of an archivist’s knowledge base.

Due to the nature of their positions, lone arrangers work mainly in isolation unless they are part of a wider community of practice.\(^{18}\) Overcoming this isolation requires management support for professional development opportunities and time to apply what has been learnt. Furthermore, management also needs to be made aware of the complexities of digital preservation and make provision for its implementation. Acquiring requisite technical knowledge is a barrier to taking action in the born-digital sphere. Lone arrangers are confronted by all the aspects of archival administration, so it can be challenging to prioritize the learning of new skills especially when those skills are applicable to only a small proportion of the work load. In larger institutions, the tasks are divided amongst many positions (e.g. reference archivist, processing archivist, digital archivist, etc.). These institutions are now looking at ways to spread the load

of digital processing to non-“digital” positions, with the digital archivists taking the lead in upskilling colleagues.\textsuperscript{19} This method of knowledge dissemination is unavailable to lone arrangers.

\textit{My repository}

The main awareness that came out of the project revolves around storage. While I had initially indicated that storage was out of scope, I soon came to realize that storage was in fact one of the key preparatory steps in taking on digital archiving. In addition the topic of calculating storage capacity requirements comes up infrequently. Finding \textbf{affordable} storage solutions that follow at least some of the principles of good digital preservation (eg more than one copy; monitoring for fixity) has not been easy. Communicating digital preservation requirements to IT professionals in an environment where a focus on digital preservation is the exception not the rule is my next challenge as I prepare for a server upgrade.

This initial foray has made my administration more aware of the extent of digital material currently held by and expected to arrive in the Archives and what will be required to manage this material in terms of technology. This exercise has also finally caused my administration to consider the management of active digital records with the development of a provincial retention and disposal schedule, a much-welcomed unanticipated gain from my scholarship research.

\textit{The New Zealand Context}

Attendance at DigCCurr Professional Institute where participants are encouraged to make a six-month action plan made me amend my goals in particular to add communicating about digital preservation in New Zealand.

Despite having a legislative function to “exercise a leadership role in facilitating and co-ordinating archival activities”, many archivists believe Archives New Zealand does not currently fulfil this responsibility.\textsuperscript{20} The Responsiveness to Maori and Community Archives Group that provided support and training to small archives was disbanded in 2014. When I found that results of work conducted at Archives New Zealand was to be found on personal blogs, I felt that I had to conduct some advocacy for lone arrangers grappling with digital preservation. I had meetings at Archives New Zealand to emphasize why this legislative function cannot be sidelined. I also wrote to members of Parliament and had a meeting with my representative.

One big issue in New Zealand appears to be the lack of awareness from those in the major institutions of how limited the non-government archival sector is in terms of repository staffing and resourcing.

Peter McKinney from the National Library of New Zealand led a workshop at iPres 2012 in which he presented a challenge to the digital preservation community through an epochal framework;

\begin{quote}
The hobbyist is the beginning of the digital preservation journey. While becoming rarer, hobbyists are altruists, driven by an awareness of loss that no-one in their organisation has understood. They are individuals in their organisations, undertaking work that is not within their job descriptions. They have recognized that information is, or has been lost and are trying to do
\end{quote}


“something”, anything, about it for the good of future generations. Generally, they have no official resources and have neither the tools nor skills to be successful.21

I disagree with Peter that hobbyists are becoming rarer. Digital preservation is still well off the radar in most lone arranger institutions (and many larger ones) in New Zealand. As a lone arranger, there is no-one else to take on the challenge, so I am proud to be a hobbyist because at least something is being done with the material in my institution – and as Nancy McGovern poses “What is good enough in your situation?” In fact I am an early adopter of digital forensics for archives in New Zealand. I am dealing with what we have now; using the tools we have now.

A survey of born-digital collections in cultural heritage institutions conducted earlier this year by the National Library of New Zealand indicates that few institutions are taking action in the domain.22. There was a 24% response rate to the survey.

- 21.4% of the respondents are transferring born-digital archival content from physical media
- 50% of those transferring born-digital archival content from physical media sometimes create disk images
- 3.4% use write blockers
- 5.2% use digital forensic software such as BitCurator or FTK Imager

So it is clear that I will be continuing in my role as cheerleader of digital forensics in the archives for a while longer yet.

**Conclusion**

During the journey I identified that command line would also be practical as I continue my incremental technical development. I was fortunate to attend the “Command Line Interface” class at the 2016 Society of American Archivists’ Annual Meeting in Atlanta, GA. While I have not done anything with this material as yet, knowing I have worked through practical exercises and that I have these resources to hand will allow me to attempt it with confidence when the need arises.

In terms of what I set out to discover about how to manage the hard drive sitting on the shelf, the project is now complete. The biggest stumbling block remaining before putting into practice all that I have learnt is the server upgrade. Not only have more hard drives arrived in the meantime, I have also identified analogue audio-visual material that will require digitization and consequent digital curation. This material is still under appraisal so I have not yet been able to ascertain how much storage space at least for the first couple of years will be required on the new server and its back-up.

What is required is the establishment and continuation of a community of practice for digital preservation in New Zealand. I have found it surprising that there are no nationwide conversations or questions about what do we do. The Digital Preservation Practical Implementers’ Guild, which would fulfil this need, is currently in abeyance. It was set up to address the goal of understanding presented in the Archives New Zealand Digital Continuity Action Plan.23

---

22 J. Moran, "Born Digital in New Zealand" (working paper, 2016).
Harvey posits that collaboration is intrinsic to digital curation.\textsuperscript{24} This is evidenced in the ongoing success of open source projects such as BitCurator. There is growing awareness in the New Zealand community that digital preservation affects more than just the mandates of Archives New Zealand and the National Library of New Zealand and that a range of actions needs to be taken.\textsuperscript{25}

Advice and hands-on training offerings need to be presented at a level which is pitched particularly at non-academic and institutions even smaller than those targeted by the POWRR project.\textsuperscript{26} To address this, the Archives and Records Association of New Zealand is looking at holding workshops in 2017 in regional centers to provide hands-on training and to spread the message to a widely dispersed audience.

What has really been brought home to me is Luciana Duranti’s statement “Records preservation starts from the moment of creation”.\textsuperscript{27}

\begin{itemize}
\item \textsuperscript{26} "POWRR: Preserving (Digital) Objects with Restricted Resources," POWRR, accessed October 26, 2016, http://digitalpowrr.niu.edu/.
\end{itemize}
Resources


Lee, C. "All Bits Considered: Caring for Collections when Bitstreams Matter." Lecture, Archives and Records Association of New Zealand (ARANZ), Library and Information Association of New
Zealand (LIANZA) and the Institute of IT Professionals New Zealand (IITP), National Library of New Zealand, Wellington, New Zealand, October 13, 2014.


