

Polonsky Blog

Archives & Records Association UK & Ireland Conference 2016

Hilton London Wembley

31st August – 2nd September 2016

It has been three weeks since the last cup of tea was self-served, the last morsel of cake consumed and the sincere goodbyes to fellow colleagues said at the annual ARA Conference, held at Wembley. Many delegates left with minds crammed with new ideas, innovations and practical lessons to use back at work. I left with the strong impression that digital preservation within the recordkeeping community in the UK and Ireland has become part of the 'mainstream' in recordkeeping practice across a variety of sectors. The recordkeeping community has moved on from wanting to know what digital preservation is to how it gets involved and preserve digital collections for future generations.

I'm going to reflect upon the conference's digital preservation offerings, summing up each session for those not able to make the road to Wembley. Extended writes up of selected sessions will also be made available in the future. A full write up of the ARA whole conference is available on the ARA website and for those wishing to read it, the link is available here: <http://www.archives.org.uk/latest-news/647-reflections-on-ara-conference-2016-global-futures.html>.

Day 1 Morning Session

Digital Preservation: collective responsibility and technical innovation

Sarah Higgins, Lecturer, Department of Information Studies, Aberystwyth University
Sally McInnes, Archivist, LLGC/National Library of Wales

The morning session began with Sally McInnes (National Library of Wales) and Sarah Higgins (Aberystwyth University) presenting an overview of ongoing work started in 2009 to develop shared digital preservation provision in Wales. The project aimed for sustained access to Welsh records and in developing capacity for digital preservation; the approach would be incremental. Through a national, shared approach, a survey was first undertaken in 2009 to understand what difficulties were in implementing digital preservation across Wales. From the results, gaps were analysed (needing to address lack of confidence, skills, infrastructure, resources) and a business made. The work undertaken included first setting up project management structures, devising policy and procedures which then enabled the project to deploy approaches to addressing infrastructure and skills needs.

First, open source systems were reviewed, selected and tested. Archivematica was first selected, piloted and installed in local record offices. Unfortunately, some Record Office systems teams did not like what they considered to be undeveloped, not beta-tested software in their systems. The project's response was to have software cloud-based as a proof of concept through Arkivum's interfaces and storage services through Microsoft Azure and CloudSigma so that the local Record Offices could participate. Preservica was also piloted and the 12 institutions involved in the project received two full days of training in the systems. For skills and training, training grants were made available to project participants. The project

continues with the development of a National Digital Preservation Policy for Wales and continued scoping for acquiring a shared cloud-based storage and management system.

It is great to see that a sustained, national digital preservation project continues to evolve. Whilst it uses systems 'solutions,' it is clear that digital preservation is a continuous piece of work that not only reacts to technological change but also to the skills and training needs of the practitioners involved to build both their skillsets and their confidence.

Creating value with digital preservation: the living and dynamic digital archive

Mike Quinn, Commercial Director, Preservica

Mike Quinn opened his paper suggesting that digital preservation was “not an archives problem, but a records problem.” He referred to the benchmark report from the Information Governance Initiative (IGI) which reports on the need for organisations to have long-term protection and access to their digital information. Delegates were reminded that they need to remain flexible in relation to digital preservation challenges, nothing is guaranteed: [Apple ending support](#) for the .MOV file format demonstrates this.

In order to help tackle digital preservation, recordkeepers need to interact with the digital records creators. Recordkeepers need to encourage those creating the records to be 'archivists themselves' and referred to the [case study of the American Institute of Architects](#) who effectively delegated the work of the ingest process to the content creators.

Mike Quinn offered us a slide on the fast evolving world of digital preservation where there are new tools, communities, standards and processes, showing that technology changes all the time. He suggested that we are moving from a world of individual tools being available for digital preservation, to build your own solutions and now to 'out of the box' solutions where we can curate them in the right way. The digital preservation industry moving to a more automated way of working, so it is now becoming a process that is business as usual.

He also warned the delegates that there's a war going on in the storage market. Do not be locked in long-term deals as storage will continue to get cheaper. He also enthused that recordkeepers need to focus on delivering value, engaging their users and not to become a technologist, instead opting for a holistic approach.

Digital preservation: keep calm and get on with it!

Dr Matthew Addis, Chief Technology Officer, Arkivum Limited

For those wanting the slide and notes from the speaker himself, please do click through to [this link](#) and enjoy his clear and easy to understand approach to what Tim Gollins coined as 'parsimonious preservation'. If you do not have time to do this please read on, although I really do insist that you have a peek of the slides. They're really rather good.

Matthews Addis suggests that managing digital records is really managing risk due to their fragility. Don't let the bits become derelict to rot away. "Apathy is the digital record killer," so starting from somewhere simple and working from there is the best way to tackle the digital preservation 'problem'. Addis observed that lots of organisations seem to suffer from a

“digital preservation paralysis” and fear getting it wrong. However, he advised those assembled that “doing nothing is the worst choice”. As seen in the slides, he gave the delegates an overview of the tools available to start a collections audit or get a better understanding of digital holdings, including DROID and the Digital Asset Framework (DAF). As well as looking at file format identification, he also suggested that it is important to consider bit-level fixity such as using checksum tools once you’ve got the files ‘in’. When looking at the maturity of your systems, it is good to benchmark them against maturity models. Re-iterating what he said earlier, it’s always good to start from the lower levels and build maturity. Allied to looking at technical solutions, staff skills and policy must also be developed for the digital preservation a programme to be maintained.

In reference to OAIS, it needs to be remodelled but is very useful as a starting point when thinking about digital preservation but it should be adapted to suit specific circumstances. He suggested that standards should evolve so to build in flexibility to digital preservation models. The OAIS model is under review of the [DPC are hosting a wiki](#) to gather input on the OAIS review.

Day 1 First Afternoon

First steps in digital preservation and online access to the BT Heritage collection

James Elder, Archives Manager, BT Archives

Jenny Hunt’s tweet neatly sums up feeling of the recordkeeping community wanting to learn more and share digital preservation experiences: “Second year of #digitalpreservation stream and great to see the room is at capacity #howthingschange.” James Elder continued the strongly attended digital preservation session for the afternoon with a packed crowd sharing his experiences.

James Elder outlined that for the BT Archives, “a good way of getting to grips with practical digital preservation” for the uninitiated is to start off with digital surrogates as means of learning the systems and software involved in the processes. The BT Archives started with preserving digital surrogates as they were a priority at the time after a series of large digitisation programmes.

The BT Archives maintain TIFFs as the digital preservation copies of the photos through Preservica systems. He found that after trials, it was necessary to have a PC with large processing power dedicated to the ingest process. For storage and faster ingest, they had moved from external hard drives (HDD) with USB2 ports to solid state drives (SSD) with the faster USB3 ports. Once he had batch, he would create a SIP using SIP creator in Preservica and then “zip that SIP” using a Windows tool rather than the SIP creator. The zipped SIP would then be transferred into the Transfer Agent upload folder, which then automatically starts to upload. To avoid slowing down networks, he recommended that uploads are done overnight or over weekends. Once ready, the ingest process is controlled from the Preservica dashboard. Documents were created as TIFFs only but access JPGs were also produced as a part of the SIP.

His take home message was to, “plan, plan, plan.” When running a project, work out what you are trying to achieve and how you are going to do it. “Flowcharts and checklists are your friends”. Preservica has its own processes, however you will need to manage your own use of the product. And finally, he restates his opening with, “starting digital surrogates is a good introduction” but admits that it’s not using Preservica to its full potential, but you do have to start somewhere.

The path to developing a digital preservation programme: the view from the bottom
Kirsty Lee, Digital Preservation Curator, University of Edinburgh

Kirsty Lee’s relatable, easy to understand approach began by calming any digital preservation anxieties amongst the delegates in that “you don’t feel you need to be an expert.” She then structured her paper and slides through a series of ‘tips’ for developing a preservation programme.

The first piece of, in the words of the speaker, “simple but vital” advice was to understand where your organisation is in terms of digital preservation work and work out what it is you want to do with digital preservation. When starting out, she said that Adrian Brown’s book on *Digital Preservation* became her *vade mecum* as the Digital Preservation Capability Maturity Model (DPCMM) was really useful in working out the level of maturity of digital preservation within her organisation. For Lee, she found that the DPCMM was a useful, clear visual guide to work out where you are with digital preservation.

She advocated Agile project management methodologies for digital preservation and agreed with James Elder’s take home message of, “plan, plan, plan.” PLATTER’s Repository Planning Checklist and Guidance tool was also of great use and when used with the DPCMM they can set out the OAIS and PREMIS requirements and identify what gaps need to be filled.

When it comes around to testing, Lee’s tip was to make sure that you systematically document your tests (for example on compression level testing) when using Archivematica or similar so that you can refer back to them for future use. For digital preservation, workflows and processes must be devised as a basis to work from and adjusted over time. Mind mapping can also be a useful tool when working on this.

Tools, practitioners should keep an eye out for developing tools, although there is the issue of the sustainability of them so you should not rely on them or build them into workflows. The Community Owned Digital Preservation Tool Registry (COPTR) was referred to (http://coptr.digipres.org/Main_Page). The take home tip was to document everything you do and share it as much as you can through blogs and papers.

E-ARK Project: Progress and Outcomes

Dr Andrew Wilson, Senior Research Fellow, E-ARK Project, University of Brighton

The European funded E-ARK project (<http://www.eark-project.com/>) has attempted to fill the gaps in current digital preservation and interoperability and aimed to provide practical solutions “to real digital preservation issues.” The presentation showed to the delegates that

digital preservation is affordable and not an expensive vendor based solution. This was an especially valuable lesson to those in smaller organisations with smaller purses.

The project was a collaboration and had a duration of three years. The first year consisted of a state of the art review, the second year saw the development of specification and tool development and the third year saw the beginning of pilot projects. The message seemed to be that to do digital preservation properly and sustainability, time must be taken in the planning stages.

A part of the outputs of the project included common SIP, AIP and DIP formats, pre-ingest, preservation and access workflows. Open source tools were developed for the implementation of specifications which were scalable, modular, robust and adaptable. The tools could be implemented individually or as an integrated reference implementation. In particular, the project developed specifications for relational databases (SIARD 2.0) in collaboration with the Swiss Federal Archives. The E-ARK project is in discussion with the OPF and the DLM Forum for the ongoing maintenance of these tools.

This was a highly informative paper which was crammed full of information on the project and some of the deliverables involved. To find out more about this project which will soon be finishing (and we hope will continue as E-ARK2!) do visit <http://www.eark-project.com/>.

Day 1 Second Afternoon Session

“Would like to know more” – Digital preservation training and professional development

Edward Pinsent, Digital Archivist, University of London Computer Centre

Stephanie Taylor, Senior Consultant, University of London

Sharon McMeekin, Head of Training and Skills, Digital Preservation Coalition

The final afternoon session of the day summarised the findings of the ‘Digital Archiving and Preservation Training Needs Survey’ led by the University of London’s Computer Centre (ULCC) in collaboration the Digital Preservation Coalition (DPC) and the Digital Curation Centre (DCC). The lively and engaging delivery of Ed Pinsent began the session through introducing the survey which seemed to ask the critical question, “what do you want from digital preservation training?”. The survey saw a good response and there were two main issues that came out of it. The first, being the title of the session, was that respondents would ‘like to know more’. Secondly, it was the ‘confidence’ in being able to address digital preservation issues. ‘Confidence’ was not a word that was used in the wording of the survey, but looking through the qualitative data it was a reoccurring word. Another issue that was raised was the ‘fast-changing’ nature of digital preservation environments as the techniques and tools used in digital preservation now will not be fit for purpose in the future. Ed Pinsent neatly presented the findings of the needs survey as follows:

1. People want to learn about strategy and planning, not exclusively DP theory, not exclusively IT;
2. People are clear that Digital Preservation training will bring them benefits directly related to their job/organisation/collections;
3. People want to learn by doing;

4. Everybody wants to know more; and
5. Everyone wants to feel confident about digital preservation.

The session continued with Stephanie Taylor's view on how training could be developed as a result of meeting the needs of the survey. The team will be "working hard" to address these needs through the delivery of new courses and updating their existing courses. The survey seems to suggest that digital preservation in the UK is 'maturing' as a discipline. As a result, specialisms within digital preservation are beginning to emerge. In terms of delivering the training, one day face-to-face courses and online training offerings were a popular need and they will develop courses in this way as a result. As a taste of the online courses, the ULCC Digital Archive Training (DART) Blog contains a link to their free OAIS introductory course which they introduced last year. You can sign up for it here: <https://dart.blogs.ulcc.ac.uk/2015/11/02/free-oais-online-course/>.

Sharon McMeekin's paper even went as so far to suggest the idea that perhaps digital preservation is a specialism in its own right (cue debate!). The DigCurV framework was recommended as being very helpful when thinking about digital preservation skills and development, although actual digital preservation skills are only a small part of the framework. In the training needs survey, they found that most people working in digital preservation have at least a Masters qualification, so the practitioners working in this field are highly qualified. For training, people would like to come away from their courses feeling that they have either enhanced their competency in particular digital preservation skills areas as well as having something to show for it. Training and workshops seem to focus more on these practical skills and competencies.

As a part of the close, Stephanie Taylor advised that for digital preservation training, there was no 'magic answer' or a 'right path' in providing training, but you do have to accept that ongoing review and starting from anew is a part of the practice.

It was a shame that Laura was unable to attend as I was looking forward to hearing how digital preservation training could be delivered for 'non-scientific' audiences, possibly through means of adjusting the technical language for the practitioner from an arts and humanities background. It would be good to hear Laura's thoughts on this.

For the Digital Preservation at Oxford and Cambridge project, these conclusions and lessons from the ULCC led survey will certainly be interesting to compare once the initial Training Needs Survey has been carried out at the two respective institutions. This talk was highly interesting and instructive as to the state of digital preservation training in the UK to date.