

DON'T PANIC

THE ARCHIVIST'S GUIDE TO DIGITAL PRESERVATION (with apologies to Douglas Adams)

Stefanie Davidson
Records Management & Digital Collections Coordinator
West Yorkshire Archive Service
PO Box 5
Nepshaw Lane South
Morley
Leeds
LS27 0QP
sdavidson@wyjs.org.uk
0113 393 9769

There is a bewildering amount of advice, guidance and research & development information available online in the fields of digital continuity and preservation. This document is intended to act as a signpost to assist in taking the first steps in understanding some of the issues involved. It is not meant to be a comprehensive guide, but to act as a starting point and introduction to the topic to help you find your feet.

Getting the basics:

Computers are the unknown

If computers are a complete unknown, you're really not sure what goes on behind that desktop on your screen and all that "2.8GHz dual-core, 4GB RAM, 750GB hard drive" stuff sounds like a foreign language then digital preservation can seem like a scary prospect. You really don't need to have a degree in computer science to get involved in digital preservation, but a little background can make you feel more at home in the world of IT. This series of online lectures from the Harvard Extension School gives a really good grounding in the basics. At the very least, the next time you're buying a new laptop, you'll understand what all those processing speed and memory size numbers actually mean.

<http://computerscience1.tv/2010/spring/> or <http://bit.ly/e0lifk>

Digital Preservation online tutorial

A free online Digital Preservation Tutorial from the US Inter-University Consortium for Political and Social Research:

http://www.icpsr.umich.edu/dpm/dpm-eng/eng_index.html or
<http://bit.ly/u95Xfb>

This is a very useful overview of some of the concepts involved in digital preservation. It includes an overview of both OAIS and TDR (see below also). The section on obsolescence and physical threats to digital information is also very helpful.

OAIS – Open Archival Information System

At the request of the ISO, the Consultative Committee for Space Data Systems (CCSDS) assumed the task of coordinating the development of archive standards for the long-term storage of digital data. To initiate this process, a reference model was developed to establish common terms and concepts, provide a framework for elucidating the significant entities and relationships among entities in an archive environment, and serve as the foundation for the development of standards supporting the archive environment. CCSDS's efforts resulted in the release of the OAIS reference model draft recommendation in May 1999.

OAIS – Reference Model for an Open Archival Information System can be downloaded here –

<http://public.ccsds.org/publications/archive/650x0b1.pdf> or <http://bit.ly/2nCh7d>
(this is the Jan 2002 Blue Book).

There is an Introductory Guide to the OAIS Reference Model by Brian Lavoie in the Digital Preservation Coalition Technology Report series which can be downloaded here

http://www.dpconline.org/component/docman/doc_download/91-introduction-to-oais or <http://bit.ly/iwpLUA>

And this blog post from Jan 2008 is quite helpful in getting to grips with the reference model - <http://alanake.wordpress.com/2008/01/15/oais-reference-model-general/> or <http://bit.ly/vBnFwe>

Terminology from OAIS has been widely adopted in the digital preservation world and you often hear about systems and workflows being designed to be “OAIS compliant” so it helps to have an understanding of OAIS when reading about digital preservation issues.

Trusted Digital Repositories

The Research Libraries Group and OCLC (a worldwide library cooperative) produced a report on the attributes and responsibilities of a Trusted Digital Repository. This report defines the organizational context for a digital preservation program. TDR embraces OAIS and demonstrates what adhering to OAIS will mean for an institution.

<http://www.oclc.org/research/activities/past/rlg/trustedrep/repositories.pdf> or <http://bit.ly/cMpecM>

There is an accompanying checklist for auditing and certifying TDRs.

http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf or <http://bit.ly/9hdFXf>

Starting the conversation:

Digital preservation takes resources, and finding out about what resources might be required and making the case for allocating them will require a lot of awareness-raising within your organisation. Some things which should help you express your requirements in a way management and stakeholders might understand are listed below.

Digital Continuity guidance from the National Archives UK:

The best place to start – comprehensive guidance on digital continuity from the National Archives.

<http://www.nationalarchives.gov.uk/information-management/our-services/digital-continuity.htm> or <http://bit.ly/uly5Co>

Of particular use are:

Understanding Digital Continuity -

<http://www.nationalarchives.gov.uk/documents/information-management/understanding-digital-continuity.pdf> or <http://bit.ly/vkAz9f> as an introduction to the topic.

Managing Digital Continuity - <http://www.nationalarchives.gov.uk/information-management/our-services/dc-step-by-step-guidance.htm> or <http://bit.ly/v4UaUR>

Which takes you through the 4 stage process of managing digital continuity:

- 1) Plan for action
- 2) Define your digital continuity requirements
- 3) Assess and manage risks to digital continuity
- 4) Maintain digital continuity

Also very useful is the Self-Assessment Tool which can be downloaded in Stage 3 of Managing Digital Continuity – this enables you to assess your organisation's risks in relation to digital continuity and suggests actions to mitigate identified risks.

The concept of digital continuity has been taken up in Australia by the Queensland State Archives, who have recently published their own digital continuity strategy at <http://www.archives.qld.gov.au/digitalcontinuity.asp> or <http://bit.ly/upeA6X>

The **Blue Ribbon Taskforce** looked at the economic sustainability of digital preservation and their report can be downloaded from the publications section of their website - <http://brtf.sdsc.edu/>

It can be useful for expressing the benefits of digital preservation if you are working to justify the expenditure of resources in order to achieve it.

Chris Rusbridge is doing some work to develop a reference model out of the BRTF work which he's presenting as a work in progress on his

Unsustainable ideas blog at

<http://unsustainableideas.wordpress.com/economic-sustainability-ref-model-page/> or <http://bit.ly/pkMjDO>

The **Keeping Research Data Safe** (KRDS) cost/benefit studies, tools and methodologies can be seen on Charles Beagrie's website at <http://beagrie.com/krds.php> These were primarily designed to assess, define and express the benefits of ensuring long term access to digital data in academic research projects, but could potentially be useful for expressing the benefits of ensuring long term access to any digital records.

Advocacy, advice, training and resources:

The Digital Preservation Coalition (DPC)

<http://www.dpconline.org/>

The DPC is a not-for-profit membership organisation whose primary objective is to raise awareness of the importance of the preservation of digital material and the attendant strategic, cultural and technological issues.

They have a number of useful resources on their website including:

The on-line version of the Digital Preservation Handbook

<http://www.dpconline.org/advice/preservationhandbook> or <http://bit.ly/uQdetk>

Their Technology Watch Reports and Case Study notes – available in the Advice section of the website at <http://www.dpconline.org/advice>

They run training and information days on a variety of different subjects. I recently attended one relating to Email preservation and the notes from that can be seen here - <http://www.dpconline.org/events/details/32-preserving-email-directions-and-perspectives?xref=34> or <http://bit.ly/pXonDh>

The Digital Curation Centre (DCC)

<http://www.dcc.ac.uk/>

From the academic world – The DCC is a leading hub of expertise in curating digital research data. While this may seem far removed from digital preservation in archives the strategies and approaches in ensuring effective digital data management throughout the research lifecycle are similar or identical to maintaining access to digital records throughout their lifecycle.

The resources section of the website is very useful.

JISC (Joint Information Systems Committee)

The JISC Beginner's Guide to Digital Preservation can be found here -

<http://blogs.ukoln.ac.uk/jisc-beg-dig-pres/> or <http://bit.ly/dNDbos>

“It has been written for those working on JISC projects who would like help with preserving their outputs. It is aimed at those who are new to digital preservation but can also serve as a resource for those who have specific requirements or wish to find further resources in certain areas.”

There are a number of useful resources in this guide, and it's also a helpful reference point for specific issues and topics.

The **JISC Digital Media** site has some very useful resources for digital audio visual materials, ranging from advice on digitisation projects to preserving digital audiovisual materials over the long term.

<http://www.jiscdigitalmedia.ac.uk/>

The Digital Preservation Training Programme (DPTP)

<http://www.dptp.org/>

Described by a former course attendee as “a veritable Swiss army knife of tools, models, maps, trends and critical thinking” and run by the University of London Computer Centre this 3 day intensive course is extremely useful in providing an understanding of the main issues involved in digital data management and providing a road map for developing strategies and solutions for your organisation. This link provides an overview of the course run in London in Oct 2010:

<http://www.dptp.org/course/>

Unfortunately it may be beyond the current training budget of many organisations.

CAIS (Dundee University Centre For Archive and Information Studies)

CAIS offer single module distance learning courses for continuing professional development, including one on the Management and Preservation of Digital Records. See

http://www.dundee.ac.uk/cais/cpd/modules_for_cpd.htm#mpdr

or

<http://bit.ly/vu5XMa>

Prestocentre

The Presto Centre in Europe concentrates specifically on the preservation of digital audio visual materials. Their main site can be found here

<http://prestocentre.eu/>

The Library section is very useful.

Digital Preservation Europe has an online video training page here

<http://www.digitalpreservationeurope.eu/video-training/prague-2008/?media=2>

or <http://bit.ly/rIIU61>

The **Planets** (Preservation and Long-term Access through Networked Services) project website at <http://www.planets-project.eu/> can be used to access their video training materials here

<http://www.planets-project.eu/training-materials/> or <http://bit.ly/9qwB4h>

The **US Library of Congress Digital Preservation** website can also be used to access a number of different training videos and resources.

<http://www.digitalpreservation.gov/>

Online bibliography for further reading

An online bibliography for Digital Curation and Preservation is available here –

<http://digital-scholarship.org/dcpb/dcpb.htm> or <http://bit.ly/ddymoA>

Wikipedia!

Is a really useful resource. If I need a quick answer on what an acronym stands for, information on a file format or an open source digital preservation tool I'll check Wikipedia first –

See entries on PRONOM and DROID from TNA -

http://en.wikipedia.org/wiki/PRONOM_technical_registry

PDF/A file format - <http://en.wikipedia.org/wiki/PDF/A>

Web-archiving - http://en.wikipedia.org/wiki/Web_archiving

And whatever else you can think of.

Keeping in the loop:

Social Media/Web 2.0

If you've never ventured much into the world of Social Media or Web 2.0 the following self-directed online training course for Archivists might be helpful –

<http://23thingsforarchivists.wordpress.com/>

I certainly couldn't cope without using RSS feeds in a feedreader to keep track of all the blogs and websites I'm interested in (and if that sentence doesn't mean anything to you, you should definitely look at this site for an explanation!)

The majority of funded projects, advisory organisations and individuals involved in digital preservation work make use of social media to share information.

Blogs

Chris Prom's **Practical e-Records** Blog, in which he has posted practical evaluations of software and conceptual models that archivists and records manager might use to identify preserve, and provide access to electronic records, is here:

<http://e-records.chrisprom.com/>

Alexandra Eveleigh (who started WYAS down the path of digital preservation) still keeps her hand in at **Around the World in 80GB** at

<http://80gb.wordpress.com/>

The Signal: Digital Preservation blog from the US Library of Congress is always interesting at <http://blogs.loc.gov/digitalpreservation/>

Archivists in the **North West Region Digital Preservation Group** at

<http://nwrpdg.wordpress.com/>

There are many, many more and if anyone *is* interested in my entire blog roll for digital preservation, let me know and I'll share.

Discussion lists and forums

The DPC operate an open discussion list through the jiscmail.ac.uk service. It is free to subscribe to the list and this is a very useful place to find out about new projects, developments and publications and relevant training events.

<https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=digital-preservation> or <http://bit.ly/cwsoWs>

Twitter

In addition to blogs (such as Practical e-Records) you will find useful information on recent developments in the field through Twitter:

For example

Open Planets on Twitter - <http://twitter.com/openplanets>

Library of Congress National Digital Information Infrastructure and Preservation Program on Twitter - <http://twitter.com/ndiipp>

Virtually being there

While it may not be possible to attend conferences and training events in person it is well worth remembering that most organisations involved in digital preservation research and activities (particularly from the academic world) lean towards openly sharing information and resources in order to facilitate collaboration.

As an example – the majority of papers from the **IPres 2010 Conference**, which took place in Vienna-

<http://www.ifs.tuwien.ac.at/dp/ipres2010/> or <http://bit.ly/WTIRG>

can be downloaded from the conference schedule pages here
<http://www.ifs.tuwien.ac.at/dp/ipres2010/schedule.html> or <http://bit.ly/8YpMWd>

Many training events and conferences provide a **Twitter hashtag** so that you can follow events through the tweets of attendees even if you cannot be there in person.

Adding presentations from events to sites such as **slideshare.net** is common. See <http://www.slideshare.net/SteveHitchcock/keepit-course-1-digital-preservation-tools-for-repository-managers> or <http://slidesha.re/rqx11X> for an example.