Preservation Planning for Personal Digital Collections

by Paul Wilson

Trials to find the best tools and guidance for planning preservation of a small collection

PAWDOC Collection

190,000+ scanned pages & 7500 born digital files (38 GB)

-174,000 TIF
-6300 Word
-3600 PDF
-550 Excel
-450 PowerPoint
-150 others
*300 originals retained

PERS Collection (7GB)

1450 documents and mementos

-1200 PDF
-170 HTML
-40 Word
-30 TIF
-20 JPG
-a few Excel and Powerpoint

*Some originals retained

PHOTOS Collection (49GB)

15,000 photos & a few movies

-15,000 JPG
-50 TIF
-10 Quicktime
-1450 Thumbs DB, PDF, Word, Excel, PNG, etc.

*all original prints retained

DOWNLOAD Toolset Documents

Scoping Documents: PERS, PHOTOS, Template
Preservation Project Plan Descriptions: PERS, PHOTOS, Template
Preservation Project Plan Charts: PERS, PHOTOS, Template
Preservation Maintenance Plans: PERS, PHOTOS, Template

http://dpconline.org/advice/case-notes/1641-case-note-prese-planning-ers-collections-pwilson

Related Resources

DPC 12-Step Questionnaire template

Abstract

This case study summarises the findings of the author when trying to find a suitable preservation planning process and associated documentation to apply to a large personal digital collection of a lifetime's worth of work documents (PAWDOC) comprising some 170,000 scanned hardcopy pages and 5000 born digital files. Since no such process could be found, the author obtained a slide set detailing a simple preservation workflow from the Digital Preservation Coalition, and used that as a base from which to establish an approach to the work. This was tested and refined on two personal digital collections, one of 1450 personal documents and mementos (PERS) and the other of 15,000 photographs (PHOTOS). Template documents were then derived from the results (these can be downloaded from the URL in the box above). While developed specifically for personal collections, these may be useful to anyone wishing to undertake a quick, initial digital preservation exercise on any small collection when unconstrained by organisational policies. A short description of this work appeared as a case study in the Technology Watch report on Personal Digital Archiving and further guidance on this subject can be found in that publication.

Keywords

Digital Preservation Planning; Preservation Planning Process; Document Templates; Personal Collections

Introduction

My lifetime collection of work documents (PAWDOC) was initially established in 1981 to gain practical experience of the emerging field of Office Automation when I was working

Paul Wilson
OFC, 2 Soames Close, Lavendon, Olney, Bucks, MK46 4EJ
pwilsonofc@btinternet.com

in the newly constituted Office Systems Division at The National Computing Centre in Manchester. The PAWDOC collection was created as a paper-based system with a card index because personal-scale document management systems were not available at that time. However, its approach (each item uniquely numbered with the number having no significance other than as a retrieval key) was deliberately designed to be digitised downstream (Wilson, 2001). Once established, the PAWDOC collection became, de facto, an integral and essential part of my working life. Most of my work documents were stored in it up to my retirement in 2012. All items have been numbered and recorded in an index. In 1987 the index was computerised using Filemaker software; and in 1996 I acquired a scanner and software and started digitising the paper documents as well as adding electronic application files to the collection. The scanned documents and all the born-digital files are managed by a document management system that is integrated with the electronic index to enable seamless search and retrieval. The index comprises some 17,000 entries representing about 190,000+ scanned pages (in 174,000 TIF files and 3,600 PDF files and 400 JPG files) , 6300 Word files, 550 Excel files, 450 PowerPoint files, around 2000 other files of various types and around 40 CDs containing a wide variety of material.. The index and all the electronic files (other than those on the CDs) are stored and accessed on a laptop computer running MS Windows. Some 350 special documents have been retained in their original physical form and are stored in two archive boxes.

My original objective to understand the impact of electronic filing on professionals has largely been achieved. However, there is one final major aspect which is left to explore – how to preserve the accessibility of the digital collection over decades in the face of incessant technological change. Of course, I have addressed aspects of this issue as I upgraded hardware and software throughout my career. However, in recent years I have become increasingly concerned that I may not be able to open some of the file formats I embedded in the collection. Furthermore, the cost of purchasing and implementing upgrades to the index and document management software may become prohibitive. Such issues are experienced by most digital collections – whether they are in the commercial, library, museum or research sectors – and much work has been done to find solutions. I discovered that this area of work is generally referred to as 'digital preservation'—a topic widely covered in the literature– see, for example, the special edition of the magazine for the European Research Consortium for Informatics and Mathematics (ERCIM News, 2010). The process of deciding how to undertake digital preservation is generally referred to as *preservation planning* so I duly set about 'Googling' and talking with people in the field to find a simple digital preservation workflow that I could apply to a personal collection like mine. The next section describes what I discovered.

Identification of Preservation Planning Resources

A widely cited model in digital preservation literature is the Space Data System Community's Reference Model for an Open Archival Information System (OAIS). This defines preservation planning as 'the services and functions for monitoring the environment of the OAIS, providing recommendations and preservation plans to ensure that the information stored in the OAIS remains accessible to, and understandable by, the Designated Community over the Long Term, even if the original computing environment becomes obsolete' (CCSDS, 2012) (N.B. The OAIS model is also published as an international standard [ISO, 2012]).

The document goes on to spend just over two pages discussing specific preservation planning functions (monitor the designated community, monitor technology, develop preservation strategies and standards, develop packaging design and migration plans), but none of this material made it clear to me the first steps I should take to preserve my personal document collection. The model is clearly aimed at providing a framework for large enterprises with producers and consumers and, therefore, does not provide the necessary guidance for preservation planning for small personal collections such as my own.

Another widely cited model is that created by the Digital Curation Centre. This explains that to undertake the preservation planning activity requires one to: 'Plan for preservation throughout the curation lifecycle of digital material. This would include plans for management and administration of all curation lifecycle actions' (Higgins, 2008). The document doesn't discuss preservation planning as such in any further detail but it does provide the

description of 'preservation action':

> '[to] undertake actions to ensure long-term preservation and retention of the authoritative nature of data. Preservation actions should ensure that data remains authentic, reliable and usable while maintaining its integrity. Actions include data cleaning, validation, assigning preservation metadata, assigning representation information and ensuring acceptable data structures or file formats' (Higgins, 2008).

I felt that the suggested actions might be relevant, but I was no closer to figuring out how to get started.

My search uncovered another document that focused specifically on preservation planning. It was produced by The NESTOR Group — 'a network of German organisations focused on Digital Preservation for libraries, archives and museums' (Nestor, 2015). Their guideline on preservation planning involves assessing Information Use and Information Type (text, audio, image etc.), according to the criteria of Financial Viability, Authenticity, Adequacy and Potential for Automation (NESTOR Working Group on Preservation Planning, 2014). While such an analysis may well be important for institutions with huge scale, such an exercise would not initiate practical steps to preserve my personal document collection.

Preservation planning has also been explicitly addressed in the PLANETS project (PLANETS, 2015), which has developed the following definition of what a preservation plan is:

> 'A preservation plan defines a series of preservation actions to be taken by a responsible institution due to an identified risk for a given set of digital objects or records (called collection). The Preservation Plan takes into account the preservation policies, legal obligations, organisational and technical constraints, user requirements and preservation goals and describes the preservation context, the evaluated preservation strategies and the resulting decision for one strategy, including the reasoning for the decision. It also specifies a series of steps or actions (called preservation action plan) along with responsibilities and rules and conditions for execution on the collection. Provided that the actions and their deployment as well as the technical environment allow it, this action plan is an executable workflow definition' (Becker, et. al., 2009).

This guidance sounded more promising, suggesting the creation of a plan document which describes a series of actions to be taken. Furthermore, Becker's paper describes the contents of such a document: Identification, Status and triggers, Description of the institutional setting, Description of the collection, Requirements for preservation, Evidence of decision for a preservation strategy, Costs, Roles and responsibilities, and Preservation action plan, with a description of each item (Becker, et. al., 2009). However, the subsequent descriptions again seem very complex and to target very large organisations. Furthermore, the paper explains how the PLANETS project has produced a freely available online tool called PLATO. This tool will automate some of these activities and will produce a complete preservation plan in both PDF and XML, which can then be executed by other parts of the PLANETS software suite (Becker, et. al., 2009; Plato, 2015). Overall, I felt that the PLANETS project provided helpful guidance for producing a written plan, but that the other activities and tools were too complicated to apply to my relatively small personal collection.

In contrast to the very detailed and tool-based approach described above, I did encounter two rather more pragmatic and straightforward documents. One was a British Library presentation slide set on a Digital Preservation Case Study that describes a 3-stage preservation planning process comprising Collection Assessment, Risk Assessment and Prioritisation (Fay, 2013). The other was a very informative presentation slide set from the US Library of Congress Digital Preservation and Education Outreach programme that describes the preservation process in 6 stages: Identify, Select, Store, Protect, Manage, and Provide (Bunnell, et. al., 2014). Despite both being aimed at large institutions, they contain many practical points of relevance for personal collections.

It became clear that most digital preservation models and approaches are, unsurprisingly, designed specifically for the

professional curation of digital materials owned by organisations. As such, their detailed processes assume far greater levels of data volumes, time, budget, facilities, commitment and organisational constraints than individuals are likely to have. However, I did find two sets of material that include guidance specifically for the individual:

- The PARADIGM (Personal Archives Accessible in Digital Media) project, conducted by the research libraries of the Universities of Oxford and Manchester, helps facilitate the ingest of individuals' papers into the digital repositories of large institutions. The project provides extensive supporting materials publically on their website, particularly in the form of a workbook that provides information on things such as Collection Development, Appraisal and Disposal, Administrative and Preservation Metadata, Digital Repositories, and Digital Preservation Strategies. The workbook also contains an appendix of 'Guidelines for Creators of Personal Archives' that includes Eleven Top Tips for Preserving Your Personal Data' (PARADIGM, 2008). However, while this material is very informative and helpful, it does not appear to include any specific preservation planning guidance applicable to my personal collections.
- The Personal Archiving section of the website of The Library of Congress includes a document advising on 'How To Preserve Your Own Digital Materials' broken down into sections on Photos, Audio, Video, Email, Personal Digital Records, and Websites (Library of Congress, 2015). However, this is very general guidance, and does not include any detailed preservation planning processes.

In summary, none of the models, approaches or guidance I came across on the net provided a process description of preservation planning that I felt would meet my needs as an individual. Though the approaches and guidance they provide can certainly inform personal preservation planning, they do not offer practical step by step advice. Consequently, I started approaching people working in the field and finally found some relevant guidance I from the Digital Preservation Coalition (DPC). The DPC kindly provided me with a 12-step preservation planning exercise sheet used in a training programme on 'Preservation Planning: from theory to practice'. The accompanying Powerpoint slide set is available on the DPC website (Kilbride, 2013). This seemed to be the closest I could get to what I was looking for. DPC also advised me to take a closer look at the PLATO tool.

Developing an Approach

To explore what specific approach to apply to the PAWDOC collection, I decided to do some trial runs of the DPC guidance on two of my other, much smaller, collections:

PERS

A collection of some 1450 electronic files of personal documents, scanned mementos and photographed personal artefacts. The files include approximately 1200 PDF, 170 HTML, 40 Word, 30 TIF, 20 JPG and a few Excel and Powerpoint files. About 400 of the files were contained in 16 ZIP files. Total file size was about 7GB. Some physical items have been retained after they were digitised. The collection contents are recorded in an Excel index with some 700 entries i.e. one entry may represent multiple electronic files and physical items. The index file is held in the same folder as the PERS contents and numbered 0000 so as to appear at the top of the file list.

PHOTOS

A collection of some 15,000 photos (which have all been digitised) and a few movies. The electronic files took up about 49GB of storage, were held in folders in the Windows 'My pictures' library, and comprised about 15,000 JPG files; 50 TIF files; 10 Quicktime video files, 1450 Thumbs DB files; and various other pdf, Word, Excel, png and Microsoft Windows Shortcut files. Sets of photos are recorded in an Excel index and most of the physical photos are mounted in albums. The index file is held in the same 'my pictures' library as the photos.

First, I did a quick test run of both the Plato 4 tool and the DPC 12-step questionnaire using the PERS collection.

PLATO 4 is a free comprehensive web-based tool which, after registering, you can apply to any number of collections (PLATO, 2015). It takes you through a workflow

that includes inputting sample documents and defining experiments to test preservation options. However, it doesn't provide any guidance on what those options could be and the subsequent stages seemed too complex for my requirements.

The 12-step questionnaire is a much more flexible tool, which scales up to large institutional collections but is also easily adapted for use with small personal collections. I filled in the questionnaire as best I could for the PERS collection, and then reflected on what I had experienced using both the PLATO and the DPC tools. My findings were as follows:

- Neither tool provides specific technical guidance on what to do.
- The DPC 12-step questionnaire is easiest to adapt for use for personal collections.
- The 12-step questionnaire seems to result in more of a scoping document, which would be a precursor to a preservation plan.
- Neither tool mentions a maintenance schedule but both seem to assume that the preservation plan will incorporate both immediate and long-term actions. I decided it will be more effective to separate the two so that an initial project (which includes a task to produce a separate maintenance schedule) can be defined, completed and closed.

As a result of these findings, I decided I would develop the following three sets of preservation planning documentation:

- Scoping Document (adapted from the DPC 12-step questionnaire)
- Preservation Plan (to define specific actions to be taken immediately)
- Maintenance Schedule (to define actions to be taken over the lifetime of the collection)

In order to hone the approach and documentation, I decided to create them first by applying them to my PERS collection; and then to refine them by applying them to my PHOTOS collection.

I researched file formats as best I could online and concluded (rightly or wrongly) that the best preservation approach for my files would be to convert them to PDF/A files using a recent version of the PDF editing software, eCopy PDF Pro Office.

The First Trial

To create the first version of the scoping document, I made some minor wording modifications to the DPC 12-step questionnaire , and replaced Question 4 ('What is the collection? How does it break down in terms of technological dependencies?'), with some fields at the beginning of the document specifying the name, broad contents and digital components of the collection. It was straightforward to answer most questions in the PERS Scoping Document, though there were two that I was unable to complete immediately: 'What are your preferred preservation approaches?' and 'What tools are available to carry out the actions to meet the risks?'. To address these questions, I researched file formats as best I could online and concluded (rightly or wrongly) that the best preservation approach for my files would be to convert them to PDF/A files using a recent version of the PDF editing software, eCopy PDF Pro Office.

Having completed the scoping document, I set about creating a PERS Preservation Plan. To do this, I combined elements of project planning that I had experienced while working as an IT professional together with aspects of the preservation planning concepts I had come across in my reading and in the scoping document. The plan that emerged consisted of two documents: a PERS Preservation Project Plan Description and a PERS Preservation Project Plan Chart. The contents of the description document were as follows:

1. INTRODUCTION (Document purpose, Ownership, and Maintenance)
2. PROJECT BACKGROUND AND OBJECTIVES

3. COMPONENT SCOPE (Collection, Location, Organisation, Process, Data, Application, Technology)
4. PRINCIPLES, ASSUMPTIONS, CONSTRAINTS AND RISKS
5. PROJECT GOVERNANCE (Organisation, Reporting, Change Control)
6. PROJECT MILESTONES AND DELIVERABLES
7. PROJECT PLAN (Tasks and Resources)
8. PROJECT BUDGETS AND COSTS

Some of the elements of the Plan Description document may appear to be not absolutely necessary for the personal sphere (for example, 'Project Governance'). However, they were included because they are easy to bypass if necessary (by saying, for example, 'not needed'), and their inclusion may expand the utility and applicability of the document.

The Project Plan Chart consisted of an Excel spreadsheet listing all the tasks identified in the project plan together with salient information for each one (Ref No, Outcome, Owner, Duration, Start Date, End Date, Status and Notes).

Before implementing the preservation plan, I sent a draft copy to a number of institutions for feedback. The Wellcome Library kindly responded saying they thought the plans were thorough and that the decision to convert most documents to PDF or PDF/A was a good one. They also suggested keeping the original versions of any documents containing some processing components (such as spreadsheets), which may not be captured within the PDF format. Wellcome further endorsed keeping off-site copies.

It was around this time that I started participating in the UCL Online Digital Curation course (DPC Online, 2014), which introduced me to the National Archives DROID tool (The National Archives, 2015). DROID (Digital Record Object Identification) is a freely available software package that provides a detailed report of the types and numbers of files in a specified directory. It is an extremely useful tool when dealing with large numbers of diverse files. I used DROID to good effect in this first trial and subsequently included it as an integral part of the refined process for the second trial.

My initial decision to convert all files to PDF/A format didn't quite go as planned. There are in fact many different versions of PDF/A :

- PDF/A-1a
- PDF/A-1b
- PDF/A-2a
- PDF/A-2b
- PDF/A-2u

My updated e-Copy software claimed to support all of these. Broadly speaking, PDF/A-1b seems to be the most basic level of conformance required and aims to achieve a reliably rendered visual appearance. PDF/A-1a supports additional features such as tags and language, while PDF/A-2 also ensures that layers, transparency and embedded files are preserved. eCopy claims to be able to check whether a document conforms to these standards, a function I used to confirm that the documents I was converting to PDF complied with PDF/A-1b. On almost every occasion, even though I was using the eCopy software to convert the documents into PDF, the compliance check returned errors, including 'bad dates', 'device-specific colour space used but no output intent defined for file', and 'missing PDF/A identifier'.

eCopy also provides a 'Fix' function that in most cases cleared the errors, though only if the resulting file was saved with a different file name. In some cases, even the 'fixed' file still had errors that were only cleared by a further 'Fix' and saving the file to yet another file name. This process proved far too time-consuming so I decided to ensure compliance with PDF/A-1b only for those files I was converting to PDF from other formats. The remaining 800+ existing PDF files would have to stay as they are.

I have recounted my PDF experiences above not to alert others to specifics about PDF (about which I know very little) or the eCopy software (which I am generally very pleased with), but to illustrate how complicated and time-consuming work on file formats can be. As a consequence, one of the key findings from this first trial was that it is vital to fully understand the file formats you are targeting, and to become very familiar with the software you intend to use, before creating the preservation plan. To ensure that

this, and any other unquantifiable areas of work, are addressed before the preservation plan is created, I decided that the second version of the scoping document would have to be modified to include a separate section explicitly to identify the pre-work required before creating the plan.

I have recounted my PDF experiences not to alert others to specifics about PDF (about which I know very little) or the eCopy software (which I am generally very pleased with), but to illustrate how complicated and time-consuming work on file formats can be.

Other findings from this first trial were:

- Put work on revising backup arrangements at the end of the schedule immediately before the taking of backups. That way the revised backup documentation will itself get backed up.
- The preservation project plan description document needs modifying to include the risks to the collection (which are already defined in the scoping document)
- The htm format was the only format that didn't convert accurately at all into PDF - this is something to watch out for. Very wide spreadsheets also didn't translate well, though that was more of a pagination issue. I have subsequently been told that PDF may not be appropriate at all for spreadsheets and other applications which don't work with fixed page sizes such as project plans, CAD files etc. – another example of the need for research before creating the preservation project plan.
- While there seems to be general agreement that PDF/A1-b is the best basic level version of PDF-A to use , there appears to be no such general consensus on the other versions of PDF/A (such as 1-a, 2-a, 2-b, 2-u, 3). These have particular features which one must match to one's own specific requirements.
- All versions of PDF and PDF/A have the same extension at the end of the file: '.pdf'. Thus there is no way of immediately and easily recognising or confirming what version of PDF the file is in, and this is a nuisance. My way of dealing with this issue was to put '- PDF-A1-b' right at the end of the file name immediately before the '.pdf'.

The final activity in the preservation plan for this first trial was to create a PERS Preservation Maintenance Plan. This was the first time I had attempted to create such a document, and it entailed envisaging the actions I would need to be taking 4 or 5 years hence in a step-by-step process description. However, it turned out not to be so difficult after all; most of the activities had already been documented in the scoping document and preservation plan, and the findings from the first trial indicated where adjustments needed to be made.

The Second Trial

The second trial was undertaken on the PHOTOS collection. Before starting work, I used the insights gained from the first trial to make the following changes to the scoping document :

- A question was added relating to the hardware and software environments being used because these are critical components in enabling long term access to the collection.
- The question relating to risk activities was modified to ensure that any necessary up-front research on risk mitigation is identified in the scoping document.
- A question was added to explicitly collect together all the pre-plan activities.
- A question was added to specify what documentation will be produced to plan and manage the work.

Having made the above changes, I set about producing the PHOTOS Scoping Document. In contrast with the PERS scoping document which didn't have a section for pre-planning activities, the pre-planning task list that emerged in the PHOTOS scoping document was as follows:

- Decide what different types of backup arrangements are to be put in place.
- For each file type: 1) open up a few example files, 2) decide what application is preferred to open the file type in, 3) define what conversion action, if any, is to be taken, 4) decide if a conversion tool is to be used and, if so, become familiar with its operation.
- Decide what folder structure the files should be retained in going forwards.
- Decide what cross-referencing should be included in each of the different types of components in this particular collection (component types included Electronic Files, Electronic Index document, Physical Albums, Suitcase containing negatives)
- Decide whether to discuss the collection with the potential future recipients.

The work on the file types was the most demanding but was greatly assisted by the DROID tool. I soon discovered that what may look like a fairly coherent set of files with a limited number of file extensions in Windows Explorer, may be a lot more complicated when a tool like DROID actually looks at the internals of each file. For example,

- Windows Explorer failed to find 9 Thumbs.DB files and 6 jpg files, that DROID found.
- 52 tif files were identified by MS Explorer but 54 were found by DROID. The two extra turned out to be HP files which are present when a scanner is installed in the system.
- Explorer found 22 MP4 files and DROID found none. Most were files converted by the VLC video conversion software and were reported by DROID as unidentified.

This is just a small sample of the variances that had to be reconciled between the Windows Explorer analysis and the DROID results. Each variance represents a challenge to understand and explain, and once one starts to deal with more than a few hundred files, that challenge becomes considerably greater. I was able to get assistance from the very helpful DROID support team when I became really stuck, but it still took me six iterations of DROID analysis, comparison with Windows Explorer's numbers, and fixing/re-categorising files before I was finally able to nail down what all the 16,500 files were. The effort required to undertake this initial stocktaking should not be underestimated.

I was able to get assistance from the very helpful DROID support team when I became really stuck, but it still took me six iterations of DROID analysis, comparison with Windows Explorer's numbers, and fixing/re-categorising files before I was finally able to nail down what all the 16,500 files were.

Having completed the pre-planning work, I then produced a PHOTOS Preservation Project Plan Description and PHOTOS Preservation Project Plan Chart. This was considerably easier the second time around, not least because I was able to cut and paste from the versions produced in the first trial. I was also able to improve/ introduce some tables to make the information easier to record and read.

The actual performance of the plan was completed successfully without any further process or documentation changes being identified. When it came to producing the PHOTOS Preservation Maintenance Plan, only one substantive change was made to its contents: Item 14 (Implement the conversion action on each file) was expanded to include advice about retaining the originals to mitigate against any reduction in quality introduced during format conversions.

Production of Templates

In the belief that the document formats derived in the course of this work may be of use to others, the project-specific contents of the second trial versions of the scoping, plan, chart and maintenance documents were stripped out and guidance notes inserted where appropriate. The resulting templates are provided as Word and Excel documents on the DPC website.

Anyone who uses the templates are encouraged to contact the author with feedback on their experiences so that the templates can be improved.

Acknowledgements

I would like to thank the following people for the help they have provided in the course of this work: Neil Beagrie of Charles Beagrie Limited for introducing me to several people in this area of work; William Kilbride of the Digital Preservation Coalition for providing initial advice and the 12-step preservation planning exercise sheet on which the Scoping Document is based; Chris Hilton of the Wellcome Trust for providing feedback on the first drafts of the Preservation Planning Description and Chart documents; Jenny Bunn of the Department of Information Studies, University College London, for suggesting I might take part in the UCL Online course on Digital Curation, and for providing very useful comments on an initial draft of this paper; Dave Thompson of the Wellcome Library for advice on file formats and on Digital Curation in general; and Sara Day Thomson of the DPC who provided valuable guidance and performed extensive editing in order to make the paper suitable for publication on the DPC web site.

Future Work

I am currently seeking a permanent repository for the PAWDOC collection and shall offer to assist the receiving repository in applying the findings of this paper to the collection. However, if a repository is not identified, I intend to undertake preservation planning work on the PAWDOC collection myself, and would be pleased to discuss collaboration on this activity with any individual or institution that wishes to engage with this work as a research or learning opportunity. PAWDOC was originally conceived as an office document Test-Set so it is appropriate that it should serve as a mechanism to trial digital preservation methods. Please contact me if you are interested in these opportunities.

References

Becker, C, Kulovits, H, Guttenbrunner, M, Strodl, S, Rauber, A, Hofman, H, 'Systematic Planning for Digital Preservation: Evaluating Potential Strategies and Building Preservation Plans.' International Journal of Digital Libraries, DOI 10.1007/s00799-009-0057-1 (Published online December 19, 2009). http://www.ifs.tuwien.ac.at/~strodl/paper/becker-ijdl2009.pdf

Bunnell, K, Kingsland, V, Krynicki, L, Reece, N, Seale, R, 'Digital Preservation Planning: Just Do It!', Library of Congress Digital Preservation Outreach and Education (March 7, 2014). http://www.slideshare.net/valariek/digital-preservation-planning-just-do-it-32025600

CCSDS, 'Reference Model for an Open Archival Information System – Magenta Book.' The Consultative Committee for Space Data Systems, CCSDS 650.0-M-2 (June 2012): 4-1 & 4-2. http://public.ccsds.org/publications/archive/650x0m2.pdf

DPC Online, 'Introduction to Digital Curation: An open online UCLeXtend Course,' University College London. Digital Preservation Coalition web site – 2014 Digital Preservation Awards – Advocacy (2014). http://www.dpconline.org/advocacy/awards/digital-preservation-awards-2014/1285-introduction-to-digital-curation-an-open-online-uclextend-course-university-college-london

ERCIM News, Special theme on Digital Preservation, Published by the European Research Consortium for Informatics and Mathematics80 (January, 2010): 10-49. http://ercim-news.ercim.eu/images/stories/EN80/EN80-web.pdf

Fay, E, 'Digital Preservation Planning Case Study,' Presented at a meeting on 'Getting started in Digital Preservation' organised by the Digital Preservation Coalition (May 13, 2013). http://www.dpconline.org/component/docman/doc_download/863-2013-may-getting-started-london-planning-case-study-ed-fay

Higgins, S, 'The DCC Curation Lifecycle Model.' The International Journal of Digital Curation, 1, no.3 (2008): 134-140. http://www.ijdc.net/index.php/ijdc/article/viewFile/69/48

ISO, 'ISO 14721:2012(en) Space data and information transfer systems — Open archival information system (OAIS) — Reference model', International Standards Association (September 1, 2012). http://www.iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=57284

Kilbride, W, 'Preservation Planning: from theory to practice,' Digital Preservation Coalition (November, 2013). http://www.dpconline.org/component/docman/doc_download/937-gettingstartednov2013kilbrideplanning

Library of Congress, 'How to preserve your own digital material' (2015). http://www.digitalpreservation.gov/personalarchiving

The National Archives, File Profiling Tool (DROID). http://www.nationalarchives.gov.uk/information-management/manage-information/policy-process/digital-continuity/file-profiling-tool-droid

Nestor, http://www.dnb.de/EN/Wir/Kooperation/nestor/nestor_node.html

Nestor Working Group on Preservation Planning, 'Guideline for Preservation Planning. Procedural Model and Implementation,' Version 2.0, Nestor Materials 15 (October, 2014). http://files.d-nb.de/nestor/materialien/nestor_mat_15-eng.pdf

PARADIGM, The PARADIGM Workbook (January 2, 2008). http://www.paradigm.ac.uk/workbook/index.html

PLANETS, http://www.planets-project.eu

PLATO, Preservation Planning Decision Support Tool (version 4). http://www.ifs.tuwien.ac.at/dp/plato/intro

Wilson, P, '20 years in the life of a long-term empirical personal electronic filing study,' Behaviour & Information Technology 20, no.5 (2001): 395-409. http://www.tandfonline.com/doi/abs/10.1080/01449290131718849