*Title:* **PAWDOC Preservation Plan Project DESCRIPTION**

*Abstract:* This document defines the Preservation work that will be undertaken on the PAWDOC collection by Paul Wilson.
**Start Date:**  17Oct2021
**End Date:** 03Dec2021
**Person Days:** 42
**Cost:** £10

**Author:** Paul Wilson

**Collection Owner Approval:** Paul Wilson

**Version and Date of this document:** 1.0, 16Oct2021

**Doc Ref:** PAW-DOC-9189-08 PAWDOC 2021 Preservation Project Plan DESCRIPTION, v1.0, 16-10-2021.docx

**Document Amendment Record**

| **Version #** | **Date** | **Page #** | **Change** |
| --- | --- | --- | --- |
| 1.0 | 16Oct2021 | n/a | This is the first version of the document. |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**Contents List**

1. Introduction 4

1.1 Document Purpose and Scope 4

1.2 Document Ownership and Maintenance 4

2. Project Background and Objectives 4

2.1 Background 4

2.2 Objectives 4

3. Project Scope (CLOPDAT) 5

3.1 Collection Scope 5

3.2 Location Scope 5

3.3 Organisation Scope 5

3.4 Process Scope 6

3.5 Data Scope 6

3.6 Application Scope 8

3.7 Technology Scope 9

4. Principles, Assumptions, Constraints and Risks 10

4.1 Principles 10

4.2 Assumptions 10

4.3 Constraints 10

4.4 Risks 11

5. Project Governance 11

5.1 Project Management structure 11

5.2 Project Reporting 11

5.3 Change Control 11

6. Project Milestones and Deliverables 11

6.1 Project Milestones and approach to achieving them 12

6.2 Deliverables 12

7. Project Plan 13

7.1 Task Breakdown 13

7.2 Resource Requirements 13

8. Project Budgets and Costs 13

8.1 Budget 13

8.2 Estimated Costs 14

9. References 14

# Introduction

## Document Purpose and Scope

This document defines the scope, timescales and costs of a project which will take preservation measures to ensure that the PAWDOC collection stays accessible in the long term and in the face of continually changing technology. The purpose of the document is threefold: first to ensure that, in the course of constructing the document, all necessary activities are identified and planned to produce a successful result; second, to provide the information upon which a decision to proceed or to shelve the preservation project can be taken; and third, to enable the project to get off to a quick start should it be approved.

## Document Ownership and Maintenance

This document is owned and maintained by Paul Wilson.

# Project Background and Objectives

## Background

The PAWDOC collection was initiated in 1981 to gain practical experience of Office Automation when the owner was working in that newly emerging field at the National Computing Centre in Manchester, UK. Once established, the PAWDOC collection became, de facto, an integral and essential part of the Owner's working life, and most of his work documents were stored in it up to his

retirement in 2012. By then it was becoming clear that parts of the collection would become inaccessible unless action was taken to preserve their contents. In order to understand how to perform digital preservation work on the PAWDOC collection, an outline workflow from the Digital Preservation Coalition was tested and refined on two smaller sets of personal documents. From this work emerged the processes and template documents which were used to undertake the first digital preservation exercise on the PAWDOC collection in 2018/19 [1]. The Maintenance Plan that was created at the end of that exercise [2] specified that the next Digital Preservation exercise on the PAWDOC collection should start on 1st September 2021.

The work at the beginning of September focused on examining the ‘Possible Future Issues’ listed in section 3 of the Maintenance Plan [2]. Actions taken were recorded in the document ‘Notes on PAWDOC Preservation Maintenance 2021’ [3] (this document will continue to be added to until the end of the project). This was followed by running the DROID analysis software across all the collection’s files to identify the file types present in the collection [4, 5]. Where possible, two or three examples of each file type were opened, and, from this examination, the tasks documented in this Project Plan DESCRIPTION document were derived.

## Objectives

The objectives of this project are to:

* make the PAWDOC collection accessible and usable in the long term;

# Project Scope (CLOPDAT)

## Collection Scope

| **Component Name** | **Description** | **Modify** | **Add** | **Remove** | **No change** | **Notes** |
| --- | --- | --- | --- | --- | --- | --- |
| PAWDOC - Digital contents | The documents comprising the collection, all in digital form | x |  |  |  | Some files may need converting to modern, readable, formats. |
| PAWDOC - Physical contents | A subset of the digital documents comprising the original hardcopy or disk versions of the documents |  |  |  | x |  |
| PAWBACK – the PAWDOC Backup disks | The CDs and DVDs that contain backups taken over the years, of the collection and its digital control files |  |  |  | x | The contents of these disks are now held on the same laptop as the main PAWDOC Digital contents. The disks still exist but could be disposed of in the future. |

## Location Scope

| **Component Name** | **Description** | **Modify** | **Add** | **Remove** | **No change** | **Notes** |
| --- | --- | --- | --- | --- | --- | --- |
| Paul's study | This is where the Digital and Hardcopy contents reside |  |  |  | x |  |
| Paul's wardrobe | This is where the box of backup disks resides |  |  |  | x |  |
| 27 High Street, Great Houghton | This is where the UK remote backup resides |  |  |  | x |  |
| Bethell's Rd, NZ | This is where the NZ remote backup resides |  |  |  | x |  |

## Organisation Scope

| **Component Name** | **Description** | **Modify** | **Add** | **Remove** | **No change** | **Notes** |
| --- | --- | --- | --- | --- | --- | --- |
| None |  |  |  |  |  |  |

## Process Scope

| **Component Name** | **Description** | **Modify** | **Add** | **Remove** | **No change** | **Notes** |
| --- | --- | --- | --- | --- | --- | --- |
| Accession process | This defines how to add new material to the collection. |  | x |  |  | The Owner is still adding material. Update the User Guide to explain how this should be done. |
| Search and Access process | This defines how to search for collection documents. It is defined on the front page of the User Guide under the heading ‘Quick Start Guide’. | x |  |  |  | Add bullets defining how to access the hardcopy of a document. |
| Preservation Maintenance process | The Preservation Maintenance Plan defines the ongoing preservation processes to be performed in the future. | x |  |  |  | Make minor changes to the Maintenance Plan template, as identified by experience of using the template in this preservation exercise. |
| Backup process | This process defines what backups will be taken of the digital components, when they will be taken, and where the backups are to be stored. | x |  |  |  | Review the current Backup arrangements and make any changes that are necessary. Changes will be required to the description of the Cloud Backup service which has recently been upgraded to a new version. A new User Guide document will need to be produced to make space in the checklist on the back page for future backup activities. |

## Data Scope

| **Component Name** | **Description** | **Modify** | **Add** | **Remove** | **No change** | **Notes** |
| --- | --- | --- | --- | --- | --- | --- |
| Digital content files | Around 110,000 content files in approximately 16,200 folders, are managed in the Windows folder system. Of these, about 80,000 files are held in roughly 1300 zip files; and many of the content files are multi-page TIF documents. | x |  |  |  | Those files that won’t open should be converted to an accessible format, if possible. |
| File Title structure | The file titles are structured as follows: Ref No…Name, Date of accession into the digital filing system, Number given to it by the Document Management System (DMS). Extension For example: PAW-DOC-5368-01 Discussions with Chris Moyer, 14-04-2001, D149.TIF |  |  |  | x | Should the file be updated, the word ORIGINAL is placed before the Date in the original file; and, in the new file, the word UPDATED followed by the date of Updating, replaces the original date and any number allocated by the DMS. For example,PAW-DOC-4311-01 Bulk data capture, ORIGINAL, 26-05-2002, 193F.docandPAW-DOC-4311-01 Bulk data capture, UPDATED, 27-03-2018.pdf |
| Filemaker Set Index | This records details of all the different sets of information that are to be found in the Filemaker Content Index. The sets are specified by the first two parts of the Reference Number, for example, DOC in PAW-DOC |  |  |  | x |  |
| Index fields | These are fields used to record details of each of the items in the collection. The fields are: Reference No, Title, Movement Status, Publication Date, Creation Date, Date Last Accessed |  |  |  | x | Index fields are described in more detail in the PAWDOC User Guide. |
| Filemaker Content Index Database | This is the definitive list of the contents of the collection; and it provides a means of searching for a document and identifying the reference number by which the actual document itself can be retrieved. |  |  |  | x |  |
| Backup data  | This comprises backups of the Filemaker Content and Set indexes; and of all the files comprising the PAWDOC collection. |  |  |  | x | Backup arrangements are documented in the PAWDOC User Guide  |

## Application Scope

| **Component Name** | **Description** | **Modify** | **Add** | **Remove** | **No change** | **Notes** |
| --- | --- | --- | --- | --- | --- | --- |
| Microsoft Windows 10 | This is the operating system under which the collection’s software applications run. |  |  |  |  |  |
| Filemaker Pro 18 Advanced | This is the database system in which the collection's Content Index and Set Index reside.  |  |  |  | x | This currently opens files with the extension .fmp12.Although no change is envisaged in the course of this project, the possibility of eliminating dependence on the Filemaker product (which keeps having to be upgraded at £200-£300 a time) by moving the indexes into Excel, will be recorded in the Preservation Maintenance Plan. |
| eCopy PDF Pro Office 6.2 | This is the PDF reader with which the majority of the collection’s files can be accessed |  |  |  | x | Opens files with extensions: .pdf, .tif, and .tiff |
| Microsoft Home and Student 2019 | This provides three of the core applications in which the Collection’s contents can be accessed: Word, Excel and PowerPoint |  |  |  | x | Opens files with extensions:.doc, .docx, .ppt, .pps, .pptx, .xls, .xlsx, .csv |
| Microsoft Photos | This enables images and photos to be viewed |  |  |  | x | Opens files with extensions:.jpg, .jpeg, .png, .bmp, .gif |
| Microsoft Films & TV | This plays Films and videos |  |  |  | x | Opens files with extensions: .mp4, mpeg |
| Windows Media Player | This plays videos and sounds |  |  |  | x | Opens files with extensions: .mp3, mpeg |
| Firefox Browser 92.0 (64-bit) | This displays web sites and other material constructed in HTML. |  |  |  | x | Opens files with extensions: .htm, .html  |
| HTML Help Control version 10.0 | This opens content structured as Help Files. |  |  |  | x | Opens files with extensions: .chm  |
| Notepad | This displays text documents |  |  |  | x | Opens files with extension: .txt |
| BT Cloud Backup Software | This is a Cloud backup service which stores copies of specified documents or specified directories, to a cloud server. | x |  |  |  | The Backup section in the User Guide will need updating to reflect a recently updated version of this application. |

## Technology Scope

| **Component Name** | **Description** | **Modify** | **Add** | **Remove** | **No change** | **Notes** |
| --- | --- | --- | --- | --- | --- | --- |
| Chillblast Leggera 14" Student Ultrabook | Notebook i7 computer with 8Mb Ram and a 1 Tb Solid State Drive |  |  |  | x |  |
| Physical backup CDs and DVDs | The contents of these disks are all stored on the laptop in a folder called APAWBAK, therefore there is no real need to retain them. However, they are still stored in Paul’s wardrobe in case they are required in the future. |  |  |  | x |  |
| Seagate 1Tb External Hard Drive | This is the dedicated backup device for the collection. |  |  |  | x |  |
| Maxtor 500Gb External Hard Drive | This is the remote UK backup device for the collection. |  |  |  | x |  |
| One new 500Gb Memory Stick | This will act as the new remote New Zealand backup device for the collection. |  | x |  |  | This will be acquired as and when someone is travelling to or from New Zealand. |
| BT Cloud Service | This is a Cloud backup service which stores copies of specified documents/ directories, to a cloud server. |  |  |  | x |  |
| BT Broadband | This is the Wi-Fi and broadband service over which the BT Cloud service operates. |  |  |  | x |  |

# Principles, Assumptions, Constraints and Risks

## Principles

The Principles below have been followed in the construction of this Project Plan, and will be applied throughout the performance of the project:

* No action will be taken which will increase the cost or effort required to maintain the collection
* Backup arrangements are considered to be significant factors in ensuring the longevity of a collection and will therefore be included as an integral part of this preservation project plan.
* All Word, Excel and PowerPoint files are to be at a version of the software not earlier than 1997.
* No attempt will be made to link files which are converted in this project, into their correct positions in web sites. Instead, they will be stored as separate files, or in a zip file, in the relevant Reference Number folder, with text along the following lines in their file titles: ‘This is an updated version of a file in the XXX website’

## Assumptions

The Assumptions below have been followed in the course of constructing this Project Plan:

* Document Management application and configuration files that were included in the collection do not need to be able to run in the future as they do not contain content information. Therefore, files related to those applications can be ignored in the exercise to open two or three files of each type as reported in the DROID analysis. This includes the applications Filemaker, PaperClip, Fish, SQL Server, and Envision. The mere presence of the files in the collection is sufficient.
* None of the versions of HTML, including those pre-dating HTML 2.0, are in immediate danger of being unreadable; and, therefore, no preservation work is required on any of the Collection's HTML files.

## Constraints

This project may be limited by the following constraints:

* No automated tools are available for undertaking conversions of large numbers of files; and the use of macros has been discounted as being too error-prone and risky. Therefore, all the Preservation work defined in this Project Plan has to be undertaken manually by a single individual.

## Risks

There is a risk that:

* The Zamzar service may be unable to convert some of the files submitted to it, or that are passed to its experts for an appraisal. Mitigation: record the need to take further actions on specified files in the future, in section 3 of the Preservation Maintenance Plan, OR add information to the Index Title field and/or Movements field to the effect that the files concerned are unreadable and cannot be converted to an accessible format.

# Project Governance

## Project Management structure

Paul Wilson - Collection Owner, Sponsor, Project Manager, Project staff

## Project Reporting

The Project Manager will create a progress report every weekend and place it in a Progress Report sub-folder. The progress report will have 6 headings:

* ***Overall Status:*** either Red (serious problems), Amber (key decisions needed), Green (OK), Blue (Complete)
* ***Achievements this week:*** deliverables produced, milestones achieved, tasks completed.
* ***Plans not achieved:*** deliverables/milestones/tasks that were planned to be completed but have not been, and the actions necessary to regain the lost ground.
* ***Activities next week:*** deliverables/milestones/tasks that will be worked on next week.
* ***Key Issues:*** status of key issues.
* ***Key Risks:*** status of key risks.

## Change Control

The baseline for this project will consist of:

* This Preservation Project Plan DESCRIPTION document [6]
* PAW-DOC-9189-08 PAWDOC 2021 Preservation Project Plan CHART, v1.0, 16-10-2021 [7]
* PAW-DOC-9189-08 DROID analysis (Report) of PAWDOC on 12Sep2021 [5]
* PAW-DOC-9189-08 DROID analysis (full details) of PAWDOC on 12Sep2021, 16-10-2021 [4]
* PAW-DOC-9189-08 Changes made in the 2021 PAWDOC Preservation Exercise [8]

Changes to the CHART and DROID analysis and CHANGES MADE documents will not be controlled. Changes to this Preservation Project Plan DESCRIPTION document will be recorded in the Amendment Record at the front of this document.

# Project Milestones and Deliverables

## Project Milestones and approach to achieving them

The key project milestones are listed in the table below:

| **#** | **Milestone** | **Approach to achieving the Milestone** |
| --- | --- | --- |
| 1 | Bulk conversions completed | Word and Excel files will be converted using existing MS Office software; and Help files will be converted using freeware. |
| 2 | Conversions using Zamzar completed | A one-month subscription to Zamzar will be taken out to undertake conversions for which I do not have the necessary software. |
| 3 | Records relating to conversions updated | Additional tasks related to converted files will be performed; the index entries for files that have been converted will be checked to make sure they are correct; and text for the ‘Possible future issues’ section of the Maintenance Plan will be produced. |
| 4 | Documentation produced | The User Guide will be updated; and a new Preservation Maintenance Plan will be Produced. |
| 5 | PawdocDP project completed |  |
| 6 |  |  |

## Deliverables

The deliverables listed below will be produced. The approver will be the collection Owner.

1. 1023 updated Word documents in docx format
2. 343 updated Excel documents in xlsx format
3. 28 new Help files in chm format
4. New versions of 8 Photoshop, Lotus Organiser and Lotus ScreenCam files
5. New versions of 28 MS Project files in xlsx format
6. New versions of 16 database files in xlsx format
7. New versions of 165 Visio files in pptx format
8. 12 updated versions of PowerPoint files in pptx format
9. Updated Index entries for all changes made to the collection in this exercise
10. Updated User Guide
11. New Maintenance Plan for the next four years
12. A finalised set of reports documenting the work done in this preservation exercise
13. Updated local backup of the PAWDOC collection
14. Updated remote backup of the PAWDOC collection.

# Project Plan

## Task Breakdown

| **Task No** | **Task Name** | **Resource (all PAW)** | **Deliverable Produced** | **Start Date** | **End Date** |
| --- | --- | --- | --- | --- | --- |
| **1** | **Perform bulk conversions** |  |  |  |  |
| 1.1 | Convert 28 Help files to chm | 4 | 28 .chm files | 17Oct21 | 20Oct21 |
| 1.2 | Convert 317 Excel docs to xlsx | 5 | 317 .xlsx files | 21Oct | 25Oct |
| 1.3 | Convert 974 Word docs to docx | 15 | 974 .docx files | 25Oct | 09Nov |
| **2** | **Convert files using Zamzar** |  |  |  |  |
| 2.1 | Convert 3 Photoshop files | 0.2 | 3 .jpg files | 06Nov | 10Nov |
| 2.2 | Convert 1 Lotus Organiser file | 0.1 | 1 new file | 06Nov | 10Nov |
| 2.3 | Convert 4 Lotus ScreenCam files  | 0.2 | 4 new files | 06Nov | 10Nov |
| 2.4 | Convert 28 Project files to xlsx | 1 | 28 .xlsx files | 09Nov | 10Nov |
| 2.5 | Convert 16 .DBF files to xlsx  | 0.5 | 16 .xlsx files | 11Nov | 15Nov |
| 2.6 | Convert 168 Visio files to pptx | 5 | 168 .pptx files | 11Nov | 15Nov |
| 2.7 | Convert 13 PowerPoint files to pptx | 1 | 13 .pptx files | 16Nov | 17Nov |
| **3** | **Update conversion records** |  |  |  |  |
| 3.1 | Perform Change Sheet Other Actions | 2 | Completed checklist | 18Nov | 19Nov |
| 3.2 | Write up Possible Future issues | 1 | Text for Maintenance Plan | 20Nov | 21Nov |
| 3.3 | Check index is correct for conversions | 1 | 3 completed checklists | 22Nov | 22Nov |
| **4** | **Produce documentation** |  |  |  |  |
| 4.1 | Update the User Guide | 1.5 | User Guide v2.0 | 23Nov | 24Nov |
| 4.2 | Print the UG & put copy in archive box | 0.5 | UG v2.0 in archive box | 25Nov | 25Nov |
| 4.3 | Produce the next Maintenance Plan | 2 | New Maintenance Plan | 26Nov | 28Nov |
| 4.4 | Put date in preservation control sheet | 0.1 | Updated control sheet | 28Nov | 28Nov |
| **5** | **Close down the Preservation Exercise** |  |  |  |  |
| 5.1 | Finalise PAWDOC DP2021 docs | 1 | v1.0 of all DP21 Docs | 29Nov | 30Nov |
| 5.2 | Place DP docs in PAWDOC | 0.3 | DP21 Docs in PAWDOC | 30Nov | 30Nov |
| 5.3 | Update PAWDOC local backup | 0.3 | New backup on Su's laptop | 01Dec | 01Dec |
| 5.4 | Update PAWDOC remote backup | 0.3 | New backup on Maxtor HD | 01Dec | 03Dec |
|  |  | 42 |  |  |  |

## Resource Requirements

| **Initials** | **Type/Name of resource** | **Total days** | **Cost** | **From** | **To** |
| --- | --- | --- | --- | --- | --- |
| PAW | Collection Owner / Paul Wilson | 42 | n/a | 16Oct2021 | 03Dec2021 |
|  |  |  |  |  |  |

# Project Budgets and Costs

## Budget

£10 has been allocated by the Collection Owner as the budget for this project

## Estimated Costs

| **Cost Element** | **Internal Costs** | **External Costs** | **Total** |
| --- | --- | --- | --- |
| Zamzar subscription for 1 month |  | 10 | 10 |
| **Total Project Expenditure** |  | **10** | **10** |
|  |  |  |  |
| Filemaker upgrade every three years  | **-** | £300 | £300 |
| **Total Ongoing Expenditure** |  | **£300** | **£300** |

# References

[1] Wilson P A, Preservation planning for personal digital collections, a Case Note published by the Digital Preservation Coalition on its website, 11Apr2016,
<http://www.dpconline.org/advice/case-notes>

[2] PAW-DOC-9189-05 PAWDOC Preservation MAINTENANCE PLAN - v1.2, 11March2019.docx

[3] PAW-DOC-9189-08 Notes on PAWDOC 2021 Preservation Maintenance exercise

[4] PAW-DOC-9189-08 DROID analysis (full details) of PAWDOC on 12Sep2021, 16-10-2021.csv

[5] PAW-DOC-9189-08 DROID analysis (Report) of PAWDOC on 12Sep2021

[6] PAW-DOC-9189-08 PAWDOC 2021 Preservation Project Plan DESCRIPTION, v1.0, 16-10-2021.docx

[7] PAW-DOC-9189-08 PAWDOC 2021 Preservation Project Plan CHART, v1.0, 16-10-2021.

[8] PAW-DOC-9189-08 Changes made in the 2021 PAWDOC Preservation Exercise

**ENDS**