Personal Digital Archiving

Gabriela Redwine

DPC Technology Watch Report 15-01 December 2015

Series editors on behalf of the DPC
Charles Beagrie Ltd.

Principal Investigator for the Series
Neil Beagrie

DigitalPreservationCoalition

DPC Technology Watch Series

ISSN: 2048-7916

DOI: http://dx.doi.org/10.7207/twr15-01

First published in Great Britain in 2015 by the Digital Preservation Coalition.

Foreword

The Digital Preservation Coalition (DPC) is an advocate and catalyst for digital preservation, ensuring our members can deliver resilient long-term access to digital content and services. It is a not-for-profit membership organization whose primary objective is to raise awareness of the importance of the preservation of digital material and the attendant strategic, cultural and technological issues. It supports its members through knowledge exchange, capacity building, assurance, advocacy and partnership. The DPC's vision is to make our digital memory accessible tomorrow.

The *DPC Technology Watch Reports* identify, delineate, monitor and address topics that have a major bearing on ensuring our collected digital memory will be available tomorrow. They provide an advanced introduction in order to support those charged with ensuring a robust digital memory, and they are of general interest to a wide and international audience with interests in computing, information management, collections management and technology. The reports are commissioned after consultation among DPC members about shared priorities and challenges; they are commissioned from experts; and they are thoroughly scrutinized by peers before being released. The authors are asked to provide reports that are informed, current, concise and balanced; that lower the barriers to participation in digital preservation; and that are of wide utility. The reports are a distinctive and lasting contribution to the dissemination of good practice in digital preservation.

This report was written by Gabriela Redwine, Digital Archivist at the Beinecke Rare Book & Manuscript Library at Yale University. The report is published by the DPC in association with Charles Beagrie Ltd. Neil Beagrie, Director of Consultancy at Charles Beagrie Ltd, was commissioned to act as principal investigator for, and managing editor of, this Series in 2011. He has been further supported by an Editorial Board drawn from DPC members and peer reviewers who comment on text prior to release: William Kilbride (Chair), Janet Delve (University of Portsmouth), Marc Fresko (Inforesight), Sarah Higgins (University of Aberystwyth), Tim Keefe (Trinity College Dublin), and Dave Thompson (Wellcome Library).

Acknowledgements

I would like to thank Neil Beagrie for his encouragement, patience, and editorial support. This report would not have been possible without him. The DPC membership provided thoughtful feedback on an early outline of the report, which made the final product stronger and (I hope) more useful. The stimulating briefing day William Kilbride organized and hosted in April 2015 on the subject of personal digital archives provided inspiration and invigorating discussion. I would also like to thank Cathy Marshall, Charlotte Abney Salomon, and Paul Wilson for contributing such thought-provoking case studies. Finally, I owe many thanks to my peer reviewers for their substantive and thoughtful feedback which greatly improved the report. Any errors, omissions or inconsistencies are my own.

Gabriela Redwine

December 2015

Contents

1. Abstract 1
2. Executive Summary 2
3. A Brief History of Personal Digital Archiving 3
4. What Are Personal Digital Archives and Why Do They Matter? 7
4.1. What are personal digital archives? 7
4.2. Why do personal digital archives matter? 9
5. Key Issues and Concerns Related to Personal Digital Archives 10
5.1. What are the key threats to a personal digital archive? 10
5.2. How can you figure out what is in your archive? 14
5.3. What happens to a digital archive during a person's life and after death? 15
5.4. How can you take good care of your personal digital archive? 16
5.5. How much will it cost to maintain your digital archive? 17
6. Case Studies 19
6.1. Case study: Personal Digital Archiving 19
6.2. Case study: Institutional Archives of Social Media 20
6.3. Case study: Digital Preservation Planning 22
7. Conclusions 25
7.1. Reasonable steps forward 25
7.2. Managing expectations 25
7.3. Resource implications 25
7.4. A recap of steps you can take to preserve your own materials 26
7.5. Timescales 26
8. Current Activities and Resources Devoted to Personal Digital Archiving 27
8.1. Guidance 27
8.2. Grassroots groups and activities 28
8.3. Software resources 29
8.4. Professional/Academic Writing and Resources 29
8.5. Digital Legacy 30
8.6. Audio and video recordings 31
8.7. Web archiving 32
9. Summary of Key Points and Resources for Curators 33
10. Glossary 34
11. References 35

1. Abstract

Figure 1: 'What is Digital Preservation?'

Illustration by Jørgen Stamp digitalbevaring.dk CC BY 2.5 Denmark

Personal Digital Archiving explores some of the key challenges individuals face as they struggle to manage and store their digital files. The report stresses the importance of preserving personal files, particularly as personal archives expand to include a combination of physical, digitized and born-digital materials. The problems and recommendations outlined in the report can aid individual creators and users of personal digital archives, as well as the curators who may be in a position to provide advice about what to keep and how to preserve it, in their efforts to preserve personal digital files for the long term.

2. Executive Summary

At the briefing day hosted by the Digital Preservation Coalition (DPC) in April 2015, one of the speakers suggested that the term 'personal digital archiving' be replaced by the more urgent exhortation, 'Save your digital stuff!'. The following report provides an overview of the work that cultural heritage institutions in the UK and elsewhere have been doing over the last few decades to capture and preserve personal digital files, and offers guidance and resources to help individuals save their digital 'stuff'.

The term 'personal digital archiving' refers to how individuals manage or keep track of their digital files, where they store them, and how these files are described and organized. People keep personal archives for reasons that may be simultaneously sentimental, practical, and even accidental. The ubiquity of personal computing devices and the ease with which files can be duplicated and stored in multiple locations mean that the same digital files may be stored on a local device, such as a mobile phone or laptop, as well as in the cloud. Personal digital archives contain files that capture both the mundane and the extraordinary, and represent moments that people may want to remember forever, as well as some that a person may wish had never happened. An important part of personal digital archiving involves determining, either individually or with the help of a curatorial institution, which parts of an individual's 'digital stuff' are worth trying to save permanently. The more actively one selects, manages and cares for personal digital archives, the more likely those are to survive the ravages of time on technology. Similarly, the more control an individual has over their personal archives, the greater their ability to save only what they intend.

This report provides an overview of the key issues related to personal digital archiving, arguing for the importance and urgency of preserving personal files while also acknowledging the difficulty of managing digital media and files that include a combination of digitized and born-digital materials. A short introduction to the role of cultural heritage organizations, such as libraries and archives, in the history of personal digital archiving, as well as current initiatives, sets the stage for resources and recommendations for individuals who want to be proactive about saving their own digital stuff. Each of the main sections concludes with a range of recommended actions and helpful hints that range from 'quick wins' requiring a minimum amount of effort, to more intensive and costly solutions. Three case studies provide more detailed exploration of individual personal digital archiving practices and consider the possible role that cultural heritage institutions might play in preserving the digital files posted on social media outlets such as Facebook.

The target audience for the report is twofold: (1) individuals who are creating and using growing personal digital archives and are concerned about how best to manage and preserve them, and (2) professional librarians, archivists, and curators, or volunteer archivists who may be responsible for advising people from all walks of life about how to take good care of their digital files and help them select which ones to preserve permanently. Although this report considers and summarizes the implications of certain suggestions for curators and collecting institutions, the guidance and recommendations offered are not intended to duplicate advice that is available for curators in *Born Digital: Guidance for Donors, Dealers and Archival Repositories* (Redwine *et al.*, 2013). The report does not address in detail concerns related to the financial appraisal of digital archives, intellectual property, specific technology issues, or professional standards but will point to relevant detailed coverage elsewhere in the professional literature.

3. A Brief History of Personal Digital Archiving

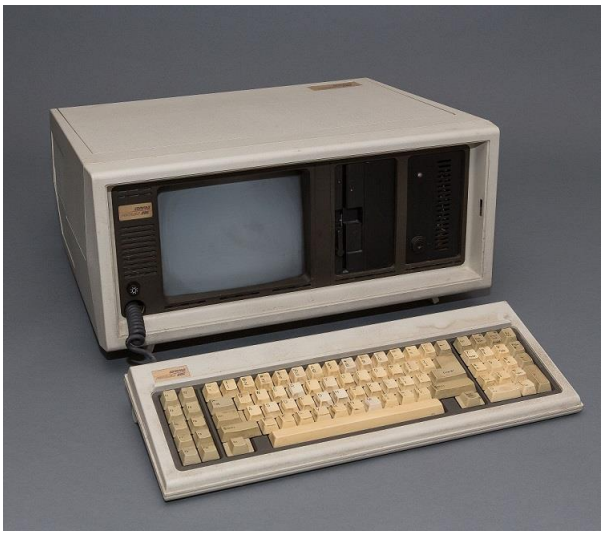

Figure 2: Compaq Portable 286, mid 1980s, an early example of a personal computer.

Larry Kramer Papers, Yale Collection of American Literature, Beinecke Rare Book and Manuscript Library © Beinecke Library 2014

In 'Ghost in the Machine', a 1989 episode from the *Inspector Morse* television series, Chief Inspector Morse and Sergeant Lewis investigate the death of a local baronet and art collector. Lewis discovers what appears to be a suicide note on the IBM PS/2 desktop computer in the victim's office, but Morse is sceptical and declares the note to be a fake. 'What?' asks Lewis, mystified. Gesturing toward the computer, Morse barks, 'Whoever wrote a suicide note on a typewriter, let alone one of these things?'.

Although Morse's investigative instincts proved correct – the note was indeed fabricated — his resistance to the idea of a word-processing file as an appropriate medium for something as personal as a suicide note highlights a pervasive undercurrent in the history of personal digital archiving: namely, the perception that items in digital formats hold less significance than their analogue counterparts and are therefore not worth saving. As personal computers became progressively smaller and more affordable in the 1980s, and the storage capacity of removable media such as floppy disks increased, it became feasible for individuals to save digital files in a more permanent way and even store a back-up copy for safekeeping.

Librarians, archivists, and curators at cultural heritage institutions, some of which were already collecting traditional personal archives, expressed concerns early on about the importance of capturing and preserving personal digital files as well. For example, web archiving has been driven by archives and libraries throughout the world – the list of initiatives below is by no means comprehensive — and continues to evolve and expand in both scope and ambition. Libraries in Australia and New Zealand began experimenting with web archiving in the late 1990s and have since developed robust programmes. In 1996, the US-based Internet Archive developed the Wayback Machine, a service that provides people with access to archived versions of websites, , and in 2005 launched Archive-It, a web harvesting service for libraries and archives. The British Library's early involvement in digital preservation and curation of personal digital papers began in 2001 with the DOMAIN.UK web archiving experiment that led to the development of an ongoing programme to selectively capture UK websites (Tuck, 2004). With the formation of the UK Web Archiving Consortium in 2003 and the subsequent development of the UK Web Archive, the UK emerged as a leader in web archiving (UK Web Archiving Consortium, undated). Today, both individuals and cultural heritage institutions can choose from a variety of tools and services to harvest web content, several of which are listed on the website for the International

Internet Preservation Consortium (IIPC), a global membership organization founded in 2003 that has dedicated itself to promoting and improving the capture, storage, and use of web archives.

In the first decade of the 21st century, a few key organizations and projects began to take shape and advocate for the importance of preserving personal digital files. In late 2001, the Digital Preservation Coalition (DPC) was established with the goal of 'develop[ing] a UK digital preservation agenda within an international context' (Beagrie, 2002). The DPC remains a vital community and resource for people who are engaged in personal digital archiving and digital preservation more broadly. From 2005 to 2007, the Bodleian Library at Oxford and the John Rylands Library at Manchester undertook the Joint Information Systems Committee (JISC)-funded Personal Archives Accessible in Digital Media (PARADIGM) project, a major study that explored the issues in preserving personal digital papers and produced a workbook detailing best-practice guidelines (Paradigm, 2005). In the US, the Library of Congress first offered personal digital archiving advice on its website in 2007; since then, it has continued to provide valuable guidance on *The Signal*, a blog devoted to digital preservation (http://blogs.loc.gov/digitalpreservation/); in published reports (NDIIPP, 2013); and by means of several different types of online resources (http://www.digitalpreservation.gov/).

Also in 2007, the British Library, in partnership with University College London and the University of Bristol, launched the Arts and Humanities Research Council (AHRC)-funded Digital Lives project (2007–2009). This major initiative explored the legal, ethical, technological, behavioural and access issues surrounding how people create, acquire and share the digital equivalents of personal papers (John, Rowlands, Williams, Dean, 2010). The 2008 NEH-funded project and report A*pproaches to Managing and Collecting Born-Digital Literary Resources for Scholarly Use* (Kirschenbaum *et al*., 2009) explored the challenges faced by archivists, administrators, and/or scholars at three US institutions whose collections include personal digital archives: the Manuscript and Rare Book Library (MARBL) at Emory University, Atlanta, Georgia, the Harry Ransom Center at the University of Texas at Austin, and the Deena Larsen Collection at the Maryland Institute for Technology in the Humanities at the University of Maryland College Park. Emory's use of emulation to preserve and provide access to the contents of one of author Salman Rushdie's computers in 2010 (Saltzman, 2010) built on earlier emulation research in the Netherlands, such as the Dioscuri project, and as part of the European PLANETS project (2006–2010) and expanded it into the realm of personal digital archives. The US Library of Congress's acquisition of the Twitter archive in 2010 spurred ongoing discussion about the role of institutions in creating archives of public/personal social media files, as well as the infrastructure necessary to support such initiatives (Osterberg, 2013; Editorial Board, 2015).

In some ways, the rise of community archives in the UK has overlapped with a growing awareness of the importance of personal digital archives, both those created by, and for, individuals as well as those preserved with a wider view of community history in mind. In a 2007 investigation into the impact of community archives, commissioned by the Community Archives Development Group (CADG; now the Community Archives and Heritage Group), researchers found that 50 per cent of community archives at that time were collecting documents and sound recordings (including oral histories), 70 per cent were collecting photographs, and 80 per cent were creating digitized versions of these materials (CADG, 2007). The report highlighted the lack of paid staff in community archives and a subsequent reliance on volunteers to run the archives. Because community archives often consist of the archives of individuals, and in addition are being cared for by one or more people with limited resources, they likely face many of the same challenges as someone trying to figure out how to manage their own digital materials.

Over 30 years after the personal computer first began to make its way into people's homes, the ubiquity of personal computing devices, coupled with the popularity of the Internet, has made it easier and less expensive than ever before for people to create and share digital files. Today all over the world billions of emails, text messages and instant messages are sent daily; writings, photographs and videos posted online; online calendars and blogs maintained; word-processing programs used to write diary entries; we conduct our banking online; wear devices that track our activities; and produce all sorts of other digital traces. Complicating matters further is the fact that not all computer files are natively digital, or 'born digital': some contain digital copies of older handwritten letters, art, photographs, and audio and video recordings that have been

reformatted for preservation reasons and now exist alongside born-digital items. As a result, personal archives often contain an assortment of paper, born digital, and digitally reformatted items, all of which exist in relation to each other. The relationship, or context, of these files to one another, and to the archive as a whole, can often be as important as their contents.

Some people create and save their digital files deliberately; in the early 2000s, Gordon Bell, a former researcher at Microsoft Research, undertook the MyLifeBits project to record every moment and interaction in his life and save this comprehensive digital record for posterity (O'Neill, 2013; Microsoft Research, undated). Others eschew digital technology, preferring instead to correspond by letter rather than email, use a landline telephone and not participate in social media. The vast majority of us fall somewhere in between these two extremes, using both digital and non-digital technologies to conduct our daily work and personal lives.

As many computer users have discovered, creating digital files is the easy part; storing and managing those files requires more time and effort. Some people actively organize and back up their files, while others forget or lose heart when confronted with what seems like an insurmountable task. In her 2008 investigation of personal digital archiving strategies, former Microsoft researcher Cathy Marshall described the casual approach most people take to managing their personal artefacts as 'benign neglect' – a hands-off strategy that can have disastrous results. Often, people feel they ought to do more to manage their digital files but struggle to find the time. The field of personal information management (PIM) has developed to provide people with resources and practical strategies for organizing and managing personal files. This important field of study and practice is related but distinct from personal digital archiving, and will not be addressed in detail in this report. For a helpful overview of PIM, see Christopher A. Lee and Robert Capra's 2011 chapter 'And Now the Twain Shall Meet: Exploring the Connections between PIM and Archives' (Lee and Capra, 2011).

As personal items increasingly exist only in digital format, the needs of those who are creating and saving digital files have encouraged a new understanding of what letters, photos and other keepsakes look like in the digital age, and raised important questions about how to manage these personal digital items for use today and how to preserve them for future generations. In the last decade, the number of services designed to help individuals save, aggregate, share and store digital files has exploded. Just a few of the many popular services available over the last several years include Picasa, Flickr and Google Docs for managing and sharing photographs and other files; YouTube for sharing videos; Dropbox, Box and Google Drive for sharing and storing files of all types; and social media services such as Facebook, Instagram, MySpace and Twitter which provide a way for users to create, share and annotate different kinds of content. All of these services have one thing in common: cloud storage. When a person uploads their files to Google, or YouTube, or any other online service, those personal files are stored on remote computers owned or leased by the company that offers the service. It is important to keep in mind the fact that although storage 'in the cloud' is invisible to the average user, these remote servers are physical entities that are located in space and time and are subject to law.

In addition to file management and sharing services, a number of companies – Amazon, Google, Apple (iCloud), Microsoft (OneDrive) and Dropbox are a few – have begun to include cloud storage as part of the package they offer. Increasingly, privacy and security questions have been at the forefront of conversations about personal digital files, with regard to cloud storage services in particular. What rights do people have related to personal digital information in the public domain? What is saved and shared, and by whom? Who controls files stored on a company's server? Can law enforcement order a company to turn over an individual's files? The benefits for an individual of moving their digital photographs and other documents to a cloud-based storage service might outweigh the risks, but an awareness of the issues at stake will help people make an informed decision. A related conversation currently taking place in the public sphere has to do with the fate of digital files after a person's death. A need has arisen for tools and services to help people make arrangements for their digital files in anticipation of death, as well as to help family members and friends manage the digital afterlives of the deceased.

The work that cultural heritage institutions have undertaken over the last three decades to preserve personal digital archives has been supplemented recently by a rising awareness among individuals of the importance of their own digital files. Many of the services and tools currently available can help people be pro-active about managing their own digital 'stuff,' but these resources also raise increasingly urgent questions about privacy and security. Personal digital archives are vital to preserve, but people need to proceed thoughtfully when choosing which tools and services best meet their needs.

4. What Are Personal Digital Archives and Why Do They Matter?

Personal digital archives are first and foremost personal, and therefore shaped by the idiosyncrasies of our interests and desires. Understanding the challenges presented by digital formats is important, whether we want to preserve the traces of an active online social life, save digital copies of medical and business files for practical reasons – or not save anything at all.

4.1. What are personal digital archives?

'Personal digital archives' is a formal term for the 'digital stuff' we create and save every day. These digital files can be as banal as a text message confirming a meeting time, or as significant as a digital video of a baby's first steps. The meaning of digital files can change over time. The text message that initially seems inconsequential may take on vital significance if it ends up being the last communication from a loved one. A digital photograph of a painting for insurance purposes is not a substitute for the original, but may become incredibly significant if that artwork is stolen. This fluidity hints at one of the key challenges of preserving personal digital files: given the enormous volume of digital files being created, how can we know which ones to save? What if something that seems insignificant now becomes vitally important in five or 50 years?

In addition to items in a person's possession, personal digital files include things that were created by or about an individual but perhaps are no longer held by them, such as digital photographs stored in the cloud, Facebook and Instagram posts shared online (and hosted on a remote server), texts and emails sent to other parties, or even an old family photograph that has been digitally scanned and shared among relatives. The distributed nature of personal archives in a digital age requires a shift in thinking about what constitutes an archive. How far does the reach of the personal extend? What is the relationship between the files saved on a personal computer or disks and those shared online and stored remotely by a third-party service provider?

Thinking of a personal collection of digital files as 'archives' places emphasis on the larger context within which those digital files exist. Much like incoming and outgoing letters in a correspondence or emails in a thread, digital files exist in relationship to each other and derive additional meaning from the physical computing environment in which they were created. Taking good care of a personal digital archive means paying attention to the location, volume, health and well-being of individual digital files, while also being mindful of this larger context. Common challenges to organizing and managing one's own digital files include the acquisition of a new computer (and subsequent transfer of files from an older machine), a hard disk crash, the theft of a computer or smartphone, a cross-country move, the transfer of email from one account to another, and the adoption of new types of software and services.

4.1.1. Who creates personal digital archives?

A personal digital archive can be created by anyone. Sometimes digital files are created and saved deliberately, but just as often they are created casually or inadvertently, a by-product of life in the 21st century. Digital files can be public or private, vital or inconsequential. In today's world children and teenagers are active consumers of technology, using it to digitally document their lives and communicate with friends and family. Like adults, they are creating a body of personal digital files, some of which they or a family member might want to save for the long term. An album of digitized photographs, a diary written in Microsoft Word, digital tax returns and a teenager's Facebook page are all examples of personal digital archival materials. Like paper archives, digital files might include memories of a life-long friendship; life stories that communicate a particular social, political, cultural, or religious perspective; health or financial information that may be of interest to both immediate family and future historians; or all of the above. The power of a personal digital archive lies in an individual's ability to assess the personal value of their own files and determine a course of action for safeguarding them and deciding the extent to which they want to share them with family, friends, a cultural heritage institution, or even directly online.

4.1.2. What can be in a personal digital archive?

As mentioned above, the contents and characters of personal digital archives are as varied as the individuals who create them. That said, most types of material fall into predictable categories that expand as new technologies are developed. The files a person accumulates typically include older and new formats – for example, documents in an early version of WordPerfect and Word 2010, or AOL email files and Gmail – as well as duplicate files.

The contents of a personal digital archive might include:

- email and letters
- websites and blogs
- diaries, recipes, and other writings
- drawings and other art
- photographs (from digital cameras, smart phones) and photo albums
- music, video, and voice recordings
- social media output
- Internet search histories
- text messages, instant messages
- contact lists and calendars
- spreadsheets, presentations, and databases
- personal records created online or received digitally (bank statements, bills, taxes, home inspection reports, deeds)
- medical records
- mementos (digitized versions of physical items)
- unexpected items – anything is possible!

One important distinction is between items that were born digital, that is, began life in digital form, and digital files that function as surrogates of paper materials. A Word document, an email and a photograph taken with a digital camera are common examples of born-digital items. A digital scan of a Polaroid photograph and an MP3 audio file of a digitally reformatted cassette tape from the 1990s are examples of digitized versions, or digital surrogates, of non-digital items.

Sometimes a single digital file can fall into both categories – for example, a digital photograph of a favourite necklace is simultaneously a born-digital item and a digital surrogate for the original necklace. The important thing to keep in mind about this distinction is that digital surrogates can often be recreated – it may be possible to take another photo of the original necklace – but a born-digital item is an original file that exists only in digital form and cannot be recreated. This overlap between digital and non-digital items means that born-digital and digitized items often co-exist on an individual's hard drive or other storage medium.

Items in personal digital archives can be shared and received by a variety of methods, including by postal mail (CDs, flash drives), over a network (shared cloud storage, Dropbox), or posted to public or private websites (Facebook, Instagram, Picasa, blogs, YouTube). Key questions to ask when choosing which items to keep or add to a collection of digital files include:

- What about this letter, photograph, email, recording, text message, or other type of file is significant?
- How are the different files in a folder or on a computer related, and are those relationships important?
- What descriptive information (e.g., names of people in photographs, story behind a family heirloom, author of a Word document) needs to be added so that the significance of a saved file will be apparent 20 years from now?

Sometimes a person creates a digital archive but chooses not to share it with anyone in their lifetime. In this type of situation, the contents may be taken to a local historical society, library, archives, or other cultural heritage institution to be assessed for fit with that organization's collecting focus, or for general advice about how to proceed.

4.2. Why do personal digital archives matter?

Personal digital archives matter because individual lives matter. The records of people's lives are intrinsically important and worth preserving. If individuals, archivists, librarians, curators, and others do not do more to safeguard personal digital archives, historical documentation of how people lived in the late 20th and early 21st centuries will be lost. The implications of this potential gap in digital primary source materials extend beyond the records of everyday life, and also determine what digital information will survive to document the early life and work of future writers, artists, musicians, architects and others. On a more practical level, it is important to provide easy access to information and resources to help people create, manage, and preserve their own personal digital archives for the long term so they will continue to have access to the digital files necessary to conduct their daily lives.

Digital materials that have personal value to individuals and families may be shared online or not at all; they may be of interest only to immediate family, or contain information of great research value to current or future scholars. Regardless of any perceived historical or cultural value, personal digital archives represent an important chance for individuals to give voice to their own perspectives, to make sure the cultural record reflects their lived experiences, and to share their stories with other people if they choose.

Well-managed personal archives will also make it easier for individuals to maintain control over their personal digital files and make well-informed decisions about how their files are used in daily life, and to plan for the preservation or disposition of digital files after their death.

5. Key Issues and Concerns Related to Personal Digital Archives

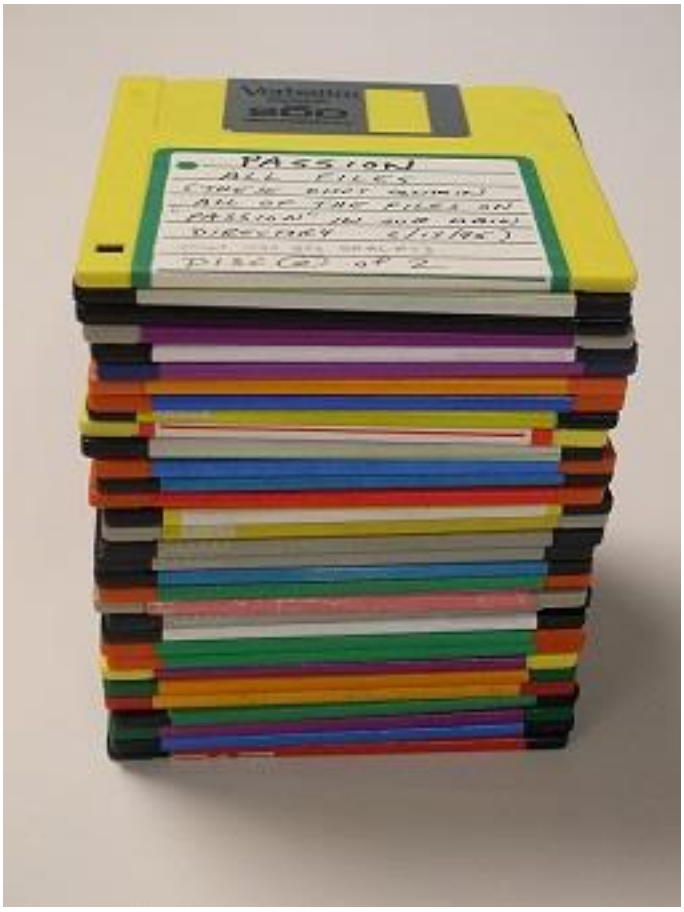

Figure 3: Personal files can become trapped on older media, like these floppy disks, once the drives needed to read them become obsolete and the internal storage medium begins to degrade.

James Lapine Papers, Yale Collection of American Literature, Beinecke Rare Book and Manuscript Library. Photograph © Gabriela Redwine 2015

A number of issues present challenges for individuals trying to manage their personal digital files, whether those files reside on vintage disks, a modern computer, or are stored in the cloud. Awareness of the following key issues is important, regardless of whether an individual manages their content for the purpose of keeping and preserving it for the long term, or focuses their efforts on gaining control over their digital files with the ultimate goal of irrevocable destruction.

5.1. What are the key threats to a personal digital archive?

Technological obsolescence and lack of planning are two of the most obvious threats to the wellbeing of personal digital files, but less visible threats can be equally harmful. Most risks to digital files fall into the following categories:

- ageing hardware and software
- lack of secure storage and backup
- natural and manmade disasters
- neglect
- loss of cloud-based host or service provider
- lack of planning
- death of an individual

It is easy to amass personal digital content but more time-consuming and challenging to manage it. Unless an individual pays attention to the size and contents of their own digital archive, they run the risk of being overwhelmed by the sheer volume of files, unable to determine what or where their files are. The average person can accumulate hundreds of thousands of emails, digital photographs, text messages, digital audiovisual recordings, word processing documents, and other digital files in a lifetime; if these files are not well managed, it will be much more difficult to locate, identify, and take good care of the items that are truly meaningful. Or the entire corpus might disappear altogether, without warning, due to any of the risks enumerated above.

5.1.1. Ageing software and hardware

Technological obsolescence presents challenges for vintage hardware and software, as well as for more modern machines and programs. Simply put: no single storage medium or file format will remain perpetually accessible and reliable (Rothenberg, 1999). Hardware and software can fail in a variety of ways. The mechanical components of a computer may start to age and cause the hard disk to spin up less reliably. Files may become corrupted. Disks, computers, flash drives and other physical media can be broken, scratched, dirtied or bent.

Even when there is no physical damage to the hardware or virtual damage, or corruption, of files, an individual may be unable to access a piece of media or a digital file because the computer, disk drive or software used to create and read those files has become obsolete. These days, we have become accustomed to automatic software updates, sometimes multiple times in a month. New versions of mobile phones, laptops and cameras come on the market with startling regularity. Given the high rate of technological innovation, even recent software and hardware are likely to be rendered obsolete relatively quickly. The reality of ageing hardware and software requires us to be actively attuned to the age and condition of the digital items in our care.

5.1.2. Lack of secure storage and backup

Secure storage and reliable, up-to-date backup copies are necessary elements of a well-managed digital archive. Geographical distribution is a key component of a secure storage and back-up plan; when the back-up copies are stored in different geographical locations they are more likely to remain safe in the event that some harm comes to the original files. Because every storage medium is prone to obsolescence, actively monitoring the backup storage media will ensure that (1) the physical media remains accessible, and (2) the files on it are in good working condition. If relying on cloud storage, use a service that regularly checks for corrupted files and also automatically creates a back-up copy, but also plan to run your own periodic checks to verify that all of your files are present and unaltered. In addition, periodically test your procedure for downloading files from the cloud and storing them on removable media like an external hard drive or USB stick. If files are stored or backed up on physical media, such as an external hard drive, plan to move them onto a different storage medium every two to four years. Most importantly, do not rely on a single storage solution or service. Diversifying your back-up strategy to include more than one cloud-based service and more than one back-up copy, all located in different geographical locations, will increase the likelihood that your files will remain accessible to you over the long term.

Encryption is a technique that can be used to implement an additional layer of protection for files, but which can also increase risks for long-term preservation and access. As with any strategy it is important to weight the benefits and drawbacks (Gordon, 2015). Key considerations include whether the individual encrypting the files is likely to remember the encryption key, whether the encryption software is proprietary and/or well supported, and whether the encryption will be applied to files stored on physical media in an individual's possession (e.g., computer, flash drive) or to files stored in the cloud.

5.1.3. Natural and man-made disasters

Natural and man-made disasters can be equally destructive to personal digital files, which is one reason it is so important to maintain multiple, geographically distributed backups. An event such as a fire or flood can damage a computer, remote servers ('the cloud'), or external storage media beyond repair. Theft of equipment is another physical threat, with potential security implications if the stolen devices contain sensitive personal information. Another category of disaster includes a breach of virtual defences in the form of malware, viruses or individuals who circumvent security measures to steal personal information such as account numbers, private photographs and even an entire identity. Robust anti-virus and anti-malware software will go a long way toward preventing basic attacks, but many people will remain vulnerable to identity theft because of the extent to which their personal and financial lives are conducted online. Some key points to keep in mind when preparing for both natural and man-made disasters include:

- having a reliable back-up plan which includes geographically distributed copies of your files.
- keeping each back-up copy up to date so that the same information is reliably stored in each location.
- being prepared, in the event of fire or flood, to evacuate with a flash drive containing copies of your most important digital files.
- implementing security measures that guard against both theft of equipment and virtual intrusions.
- being able to rely on a back-up copy to recreate the contents of your drive should your original personal files become lost or compromised.

Companies offering cloud-based management and storage services are also vulnerable to data breaches and natural disasters – remember, files stored in 'the cloud' are physically located on a remote server somewhere – so make sure to research service-providers' back-up or disaster plans before selecting a service.

5.1.4. Neglect

One of the most insidious and pervasive threats to personal digital files is benign neglect at the hands of an individual. Researcher Cathy Marshall has written extensively about this phenomenon, pointing out that people often accumulate a large quantity of digital files but do not actively organize them or consider secure storage until a crisis presents itself or a major life change occurs (Marshall, 2008). With decent and relatively stable environmental conditions, papers can be neglected without suffering too much harm. Digital media and files, on the other hand, must be regularly monitored for hardware failure, software obsolescence and file corruption; they must be well and consistently named so that you can find and use particular files easily; and they must be securely stored and backed up to prevent against data loss.

A key moment of neglect – but one which also presents an opportunity for responsible management – occurs at the point when digital files are created or transferred from a camera to a computer or other storage. For example, a folder filled with 2,000 digital photographs sequentially numbered with a camera-generated root (e.g., IMG_0001, IMG_0002) will require a considerable amount of work to manage. File names need to be meaningful in order for an individual to take advantage of digital affordances—in other words, actions that are made possible because of the digital nature of the files, such as built-in search functionality or sorting capabilities—to navigate a digital archive. Notice if the date/time stamp of a camera is accurate; if it is not, and accurate date and time information is important to you, change the camera's battery and update the clock setting. Regularly transfer and back up photos, videos and other important digital files from mobile device(s) to a more secure form of storage, either cloud-based or physical media such as a computer's hard drive.

5.1.5. Loss of cloud-based host or service provider

Although cloud-based services and storage reduce the amount of work required of an individual to access and save their digital files, the reality is that 'the cloud' consists of remote physical storage managed by a company or other organization, often for profit. By agreeing to the Terms of Service of one of these hosts or service providers, you often relinquish a certain amount of control over your files. For example, if a company providing an email service (Gmail), cloud-based storage (Dropbox), or social media services (Facebook) goes bankrupt,

the files hosted or stored by those organizations are suddenly at risk and may be difficult to retrieve from the service provider. Other risks include the loss of all digital files, a data breach that results in the exposure of private files, or having one's identity stolen and used for nefarious purposes (Palmer, 2014). It is important to keep these risks in mind when searching for a cloud-based host or service provider and to compare services to find the best fit (Casserly, 2015).

If relying on a hosting service or cloud storage for personal digital files, do research to determine which service providers have a digital preservation plan in place for the files they host. Whenever possible, choose a host or service provider that makes it easy for you to export or transfer your digital files to a different service, and consider maintaining a separate back-up copy just in case the company turns out to be unreliable. For email, this could mean installing an email client (e.g., Thunderbird, Mac Mail) that downloads messages onto a local machine, and then maintaining both an online and local version of an email account. With a cloud-based storage service, it is a good idea to put a back-up copy of those same files onto removable media that can be stored in a safe location. In short, a company that provides a helpful service is not necessarily reliable; the well-being of your digital files should always be a concern.

Our tendency to use cloud-based email as the primary mailbox, as a backup for email files downloaded locally, as an archive for older emails, and also as a backup for documents sent as attachments means that email has become much more than an electronic means of communication. In the event that you need to switch from one email client or service provider to another, make sure to devise and then test a strategy for migrating any saved email messages and attachments to the new service, and then verify that you can read the older messages and open the attachments. Your strategy can be as basic as selecting a single message or a small folder from your mailbox, researching how to export content from your current email client and import it into the new email service you're considering, and then testing the procedure. If it is not easy to move saved messages to the new vendor you have chosen, consider selecting a different email client or service.

5.1.6. Lack of planning

Planning for software and hardware obsolescence, or even the transfer of files after one's death, is not as difficult as it seems. Individuals would do well to start by establishing practical (human-readable) and consistent file naming conventions and organizing their files in a way that makes it easy to search and locate files. Clear documentation – making notes about what you have done and why, the location of back-up copies, and any passwords that you wish to share – is key to ensuring consistent practices and giving other people the information they need to access your digital files if necessary.

Cost is one of the most important factors to incorporate into planning, and also one of the most difficult to estimate. The number of digital files we receive, share, and save via email, social media, and online purchasing (e.g., iTunes) grows exponentially every year, which means factoring in that growth when making cost estimates. Key expenditures may include the cost of purchasing a new external hard drive (or other storage media) every two to four years, as well as the cost of cloud storage for a variety of digital files.

5.1.7. Death of an individual

The sudden or impending death of a family member, partner or friend presents several challenges to the management of personal digital archives. The question of a person's intentions becomes particularly important: for example, what does the individual want to happen to their computer, social media accounts and cloud-based email? Do they want anyone to inherit their digital legacy, or would they rather that their digital files be destroyed after their death? In a 2014 article in the *Guardian*, Jack Schofield explored some of the different tools and strategies available to people who want to plan for their digital legacy while they are still living (Schofield, 2014). He wrote in particular about the options Facebook provides for deceased users; the comments section includes experiences shared by *Guardian* readers, as well as information about the rising number of services dedicated to helping people make clear their intentions for the disposition of their

digital files after death. Although different services providers have different policies about what happens to digital content after a person's death, the discussion and comments on Schofield's article provide a good introduction to the complicated personal, ethical and legal concerns surrounding digital legacies (*see also section 5.3.*).

Recommendations: addressing key threats to personal digital files

Quick Wins	Choose software that is well supported and creates files that can be read by a variety of different programs.
	Develop file naming conventions that are easy to remember and apply these consistently.
	Create multiple back-up copies and store them in different geographical locations.
	Test your back-up copies to make sure they are accessible and contain what you intend them to.
	Transfer files to new media every 2 to 4 years.
More Effort	Access storage media regularly to confirm that it is in good working condition.
	Regularly transfer and back up photos, videos, and other important digital files from mobile device(s) to a more secure form of storage, either cloud-based or physical media such as a computer's hard drive.
	Be conscientious about migrating files to newer file formats if the software necessary to open them is becoming obsolete.
	Compare cloud storage providers and do the legwork to find a secure service.
	Maintain more secure, geographically distributed back-up copies of your files both in the cloud and on physical media.
Maximum Effort	Survey your digital files and create an active plan that includes selecting files for safekeeping, assessing the need for file format migrations, and utilising secure and backed up storage.
	Identify and locate hardware and software necessary to access obsolete formats in your care.
	Reorganize and rename unidentified files in accordance with file naming conventions.

5.2. How can you figure out what is in your archive?

One of the initial difficulties in gaining control over a body of personal digital files is figuring out how many, and what type, of files exist. This task is made more complex by the fact that most personal archives include both paper and digital files. Would a personal archive be easier to manage if all of the files were digital and stored in one place? What if a person decides to consolidate storage media, transferring music files from CDs and digital photographs from storage cards, and storing everything on a single hard drive? Does it make sense to discard the original physical media (e.g., CDs) once the files have been copied?

The task of ascertaining the scope of a digital archive becomes more difficult when the archive includes obsolete media, such as floppy disks, as well as a large body of email files, perhaps spread across multiple accounts, in addition to files stored on a primary computer. It is vital to transfer files from older media as soon as possible. If the disk drive or software necessary to read the files no longer exists, it is possible to buy an

external floppy drive online that can connect to a modern computer via USB. If files from several previous computers now reside on a newer machine, make sure that you can still open and read the older files. Open-source equivalents of proprietary word-processing software are sometimes able to open and view files in older formats that would otherwise be inaccessible.

Once you are able to access the majority of your digital files, it will be possible to use the search functionality and utilities built into a modern computer to perform basic searches to learn more about the digital files on a computer or external hard drive. Tools designed to manage and clean up disk space on a hard drive can also be used to locate different types of files (e.g. photographs, videos, documents) on a computer, identify duplicate copies, and save a back-up copy to external media.

Recommendations: determining the contents of a personal digital archive

Quick Wins	Conduct a simple survey by counting the number of disks and other physical media that might contain your personal files.
	Using the information on disk labels, note which physical items might be worth a closer look.
More Effort	Use viewing software or no-cost, open-source alternatives to popular proprietary software to try to read files created with older software programs.
	Make notes about which files are most significant and why, which ones are inaccessible, and which ones can be discarded.
	Use built-in utilities or no-cost hard disk management software to count the number and different type of files on a computer.
Maximum Effort	Cull your digital files based on the priorities you have determined above.
	Transfer files from older media (e.g., floppy disks) to more modern and reliable storage.

5.3. What happens to a digital archive during a person's life and after death?

The choices a person makes during their lifetime about the management and care of their digital files can have a significant impact on what happens to their digital archive after death. As mentioned earlier, personal digital files may include medical records, legal documents, financial information, photographs, social media posts, email, and many other types of material. It is important to leave family and friends clear instructions about the disposition of your digital files, either by creating a separate 'digital will' or devoting a section of a standard will to describing how to access digital assets like social media accounts, email and computers, and explaining how you want them dealt with after your death. The UK-based Saga company has published two helpful guides with step-by-step instructions for how family and friends can move forward with managing a person's digital legacy, as well as a detailed list and instructions for how to deactivate a person's online accounts once they have died (Saga, 2015a; b).

As Jack Schofield's 2014 *Guardian* article makes clear, the increasing awareness of digital files as personal belongings that need to be dealt with after a person's death has dove-tailed with a rise in the number of services available to help people control the posthumous disposition of their digital files. Not surprisingly, people have very different ideas about how to manage their digital afterlives. In a 2014 article, also in the *Guardian*, Jenny Kleeman wrote about two such services: DeadSocial, whose tagline is 'Prepare for a Digital

Death & build your Digital Legacy', and If I Die, a service that allows people to write digital notes to be delivered only upon their death (Kleeman, 2014). These services help people do more than organize their digital files and share passwords. They encourage people to think in terms of legacy, sometimes by helping a person assiduously document what should be kept or discarded, but also by giving people the opportunity to use digital technologies to communicate posthumously with loved ones – in essence creating a digital file that is shared after death and becomes part of someone else's digital archive. Taking the time to plan for how you want your digital archive to be handled will limit the amount of work other people need to do on your behalf, and it may also give you an opportunity to share aspects of your digital life even after death.

Recommendations: making long-range plans for your personal digital archive

Quick Wins	Determine what you want to happen to your digital files in the event that you are incapacitated or die.
	Document your wishes in whatever way makes the most sense to you.
	Share that documentation with people you trust.
More Effort	Create a list of usernames and passwords, and either share it with someone you trust or store it with a hosted service with instructions for what is to become of it after your death.
	If you do not want anyone to access your digital files after death, create documentation that makes this clearly known and develop a plan for a friend or family member to carry out on your behalf.
	Create a legally binding will that includes a section addressing the disposition of your digital assets, or create a separate digital will.
Maximum Effort	Work with a lawyer to document your wishes for what happens to your digital files after death and inform your family and friends.
	Inventory and cull your digital files on a regular basis so that you leave behind only what you intend other people to have.
	Enroll with a service provider that, when notified of your death, will carry out your pre-determined instructions for the disposition of your digital files.

5.4. How can you take good care of your personal digital archive?

Taking care of a personal digital archive means understanding the conditions and practices necessary for digital files to remain safe, secure and in good working order.

File naming conventions and documentation are important, particularly if the archive will be used frequently. The Bentley Historical Library at the University of Michigan has prepared a series of videos and resources illustrating file naming best practices (Bentley Historical Library, undated). Strong and consistent file naming can also make it easier to manage file storage and backup. One of the common misconceptions about digital files is that moving them onto storage guarantees that they will be readily available well into the future. The reality is considerably bleaker. Personal digital archives require careful, ongoing management and the pro-active selection of files that are the most important to preserve.

Secure storage is one of the most important factors in caring for personal digital archives. Choosing high-quality storage media and refreshing it regularly will help avoid hardware obsolescence. In the event that a

storage environment is compromised and files are lost, a reliable back-up copy will be a vital part of helping you to regain control over your digital archive. Although file recovery may be a possibility in certain circumstances, the recovered files may be older or fragmented versions of the originals, and therefore unreliable.

In previous sections I have mentioned the threat of hardware and software obsolescence, as well as different options for retrieving content from older media and files. Making informed choices about hardware and software before file creation can mitigate the effects of technological obsolescence in the future. For example, using popular and robust software programs (including open-source options) can help ensure that the file formats in a personal archive will be well supported throughout their lifetime and more easily migrated when the time comes.

There are a few key things you can do to ensure that your files are stored securely and remain accessible to you and others over the long term.

Recommendations: taking good care of a personal digital archive

Quick Wins	Choose high-quality storage media and refresh it regularly.
	Be proactive about refreshing storage media, replacing outdated equipment before it fails, and not relying exclusively on one service provider or storage solution.
	Follow best practice when naming files.
More Effort	It takes time to organize digital files; at the very least, pick a date once a year to sit down and sort through the digital files you have accumulated over the past year.
	When copying files from one location to another, use software to confirm they have all transferred correctly.
Maximum Effort	Actively monitor the health of file formats and types of material that are particularly fragile.
	Prioritize the most fragile or most important files and focus your energy and time on preserving those.

5.5. How much will it cost to maintain your digital archive?

Although you need resources to take good care of your personal digital archive, it is not difficult to make significant progress with a limited budget as long as you have time to devote to the work. Because most people have both paper and digital files, building a digital archive might mean investing in a scanner or digital camera that can be used to digitize non-digital items. Once digital, these files need to be securely stored and backed up. Key areas of expense include:

- external storage (e.g., hard drive) or cloud-based storage
- software to manage files (e.g., digital photos)
- software to read older formats
- scanner to digitize analog materials
- time

The two key questions to ask about cost are: what resources are necessary to manage and securely store digital files now, and what additional resources will be necessary in the future as the number of digital files

continues to grow? In order to keep costs down, you might evaluate the files in your archive periodically, make carefully considered choices about which files are necessary to keep, and discard the rest.

Fortunately, many of the software products that can help with reading older file formats and analysing the contents of a digital archive are low or no cost, or offer a free trial period that can be put to good use. The recommendations in the following chart illustrate what you can do on limited, moderate and generous budgets.

Recommendations: costs of maintaining a personal digital archive

£	No-cost, open-source software that can read proprietary file formats
	Free 30-day trial of file viewing software that can render a variety of formats, both vintage and recent, including exported email files, databases, word processing documents, and others
	Free 30-day trial of software that can analyse the contents of your hard drive, making it easier to identify duplicate files and decide which files to keep
	USB flash drive (low-cost local storage; refresh every year)
	Cloud-based storage service (small amount of free storage)
££	Standard editions of open-source or other software that can read proprietary file formats or render a variety of vintage and recent file formats
	Standard edition of hard drive analysis software
	External hard drive (larger storage capacity, local storage; refresh every 2 to 4 years)
	Cloud-based storage service (additional storage capacity)
	Purchase an external USB 3.5-inch floppy drive online (£10–15) and use it to access 3.5-inch floppy disks
£££	Full, professional editions of useful software
	Cloud-based storage service (larger storage capacity)
	Scanner
	Outsource the transfer of your files from obsolete media and computers

6. Case Studies

The following three case studies represent different points along the personal digital archiving spectrum. Charlotte Abney Salomon walks the reader through her personal digital archive, explaining how she uses different types of software, storage devices and hosted services to keep track of her files. Cathy Marshall's case study is a thought piece that explores the question of whether individuals or institutions should assume the responsibility for archiving social media content. She shares research findings that indicate a misplaced trust in the permanence of social media sites on the part of users, as well as a lack of conviction in the importance of content shared through platforms like Facebook and Instagram which is often duplicated elsewhere in a person's digital archive – as Charlotte Abney Salomon also notes. Would an institutional archive of Facebook or Instagram serve as an accessible backup for what Salomon describes as her 'photo diary', were the original service to go away? The questions Marshall raises highlight tensions between private and public personal digital archives, users' expectations of service providers, and the changing role of cultural heritage institutions involved in both the creation and preservation of digital archives. In the final case study, Paul Wilson describes his remarkable efforts to identify, and then adhere to, a preservation planning process suitable for his personal digital collections. His exploration of different resources in the US and UK, and the pilot projects he conducts with his own personal archives, include important lessons for individuals embarking on either small- or large-scale personal archiving or digital preservation projects.

6.1. Case study: Personal Digital Archiving

Author: Charlotte Abney Salomon, a PhD student in the History of Science and an assistant at the Medical Historical Library at Yale. Before returning to graduate school, she taught middle and high school science for six years.

'The vast majority of my digital life generates, and consists of, a few primary types of file. I conduct my daily and professional work as a graduate student in email, PDF, and .doc or .docx files, while my personal digital life exists largely in photos, videos, social media streams, and links, as well as email. My digital archive is not a constant concern for me, but rather something that I actively manage at semi-regular intervals, usually occasioned by one or another storage device or service notifying me that I have run out of space.

My materials are archived informally in a few digital spaces, some physical to me and some cloud-based. The first and most important is Google: I store all of my emails and attachments in three Gmail accounts, one of which also contains the transferred contents of two other former webmail services. The vast majority of the professional documents I create are eventually sent as email attachments; I also store some of these documents in Google Drive. My broader Google account retains preferences, searches, and other information in its other products. Similarly, since my personal devices are all Apple, so much of my digital self is continually saved in iCloud that I have replaced entire devices, including my phone and laptop, without losing much data and without having intentionally backed it up. iCloud also tends to hold the links, passwords, and apps that mediate my consumption of online material. Although I used to rely on Dropbox and Box for services such as file transfer and external storage, I haven't actively used either in over a year, and have replaced those functions with the expanded capabilities of Google as well as physical storage. I also have active social media accounts that hold an informal chronological record of my digital presence on them, such as my updates, profile changes, and photos on Facebook. Of these, I use Instagram most heavily; it currently represents a photo diary of sorts that documents the past 26 months.

My physical storage devices feel more secure to me than my cloud storage. I most actively use my MacBook Air, iPad2, and iPhone: my laptop holds the text and other documents I have created as well as research and the programs I use to read, create, and edit; my iPad2, which has the most limited storage of the three, holds primarily PDF readings for my coursework; and my iPhone holds my photos. All three of these I regularly offload onto a 1TB external USB hard drive that I share with my husband. I also have about seven years' worth

of camera photos from a series of Canon pocket cameras on four SD cards; as each one filled up, I bought a new one rather than deleting photos, thus using the cards as long-term storage.

With the exception of my external hard drive, none of my storage feels particularly “safe” or “secure”, but at the same time it is all reliable enough that I rarely actively worry about it. I back up my files to the hard drive more to make space on my devices than out of concern for their permanency. Over the past few years, the changing nature of the Internet has led me to feel much less concerned with saving files. The files I read, I can always find again by searching again; the files I create, I usually send to someone, so they are in my email archive if nowhere else.’

6.2. Case study: Institutional Archives of Social Media

Author: Catherine C. Marshall, an independent researcher and writer and Adjunct Professor at the Center for the Study of Digital Libraries (CSDL) at Texas A&M University. For many years she was a principal researcher at Microsoft Research, Silicon Valley. Before joining Microsoft, Cathy was a hypertext researcher at Xerox PARC at the dawn of the Internet era.

‘When researchers ask users of social media platforms like Facebook whether they would want a personal archive of the content they’ve contributed, they generally say no. In fact, they seem perplexed. “Why would I need that?” they ask.

There are at least three good reasons why they remain unconvinced of the need to archive this content. First, they perceive the existing service as a long-term store: everything users have ever put there seems to be available on an on-going basis and there’s little reason to be concerned that it will go away anytime soon. Second, most discrete items – for example, photos and videos – are stored other places too (for example, on users’ phones or in cloud backups). Finally, users question the persistent worth of much social media content; its value was in the moment, as part of communication and self-presentation.

Yet even the most popular of commercial services may fall victim to changing fortunes. Myspace, once a popular social media site, has turned into a niche service. Other services like Geocities are gone for good, leaving content contributors and people who relied on the content stunned, unable to replace what they have lost.

This specter of loss raises the notion of an archive, maintained by a non-commercial institution, which would preserve progressive snapshots of a service’s data for future access under specified conditions. Although there are users who see the value of this sort of archive (as a legacy and a resource), to others, this type of archive seems useless, if benign. Still others find it to be a highly controversial idea.

Many of us remember the flap associated with Twitter’s donation of its public feed to the Library of Congress: it is easy to believe that even more people will question the virtue of, say, a Facebook archive. When I’ve posited the need for such an archive (particularly a Facebook archive), sometimes even close colleagues turn hostile. “You’re going to ruin Facebook for me”, they say. “I’ll never use it again.” Others don’t object, partly because they sense they are powerless in the face of complex license agreements, and partly because they believe the content they post to Facebook is so anodyne and uncontroversial that it is unimportant what happens to it.

Why do people harbor such strong feelings about the creation of institutional archives of social media? Even if we limit what’s in the archive (e.g. public content only), who is able to access it (researchers only), and when it may be accessed (embargo the data for 75 years), the idea remains controversial. A large-scale set of surveys

my colleague at Texas A&M University, Professor Frank Shipman and I have been performing over the past five years has revealed four of the thorniest objections. These user objections can be summarized as follows:

1. *Institutional social media archives violate current social norms surrounding privacy, ownership, and identity control*.

   In spite of the legal stipulations made by license agreements, users adhere to social norms when it comes to their personal data. For example, they feel that they have a right to privacy through obscurity and a right to control how their own data is used. Reciprocity isn't necessarily part of these social norms either: I can reuse a stranger's photos, but she can't reuse mine unless I give her permission. The guarantee to be forgotten, a priority in on-going European policy discussions, is a good illustration of how social norms are changing in a way that argues against the creation of an institutional social media archive.

2. *Bad applications of this type of personal data outweigh good ones*.

   We have gradually entangled anodyne aspects of personal data (say, someone's date-of-birth or their first pet's name) with security. This leads people to harbor expectations that their personal data will be used by bad actors for nefarious applications.

3. *Content veracity and timeliness cannot be guaranteed*.

   The original owners of the archived content frequently raise the issue of content accuracy. "What good is an archive", they seem to say, "if you can't trust what's in it?". Furthermore, in services like Facebook, practices like identity curation and content removal ensure that nothing is fixed, and unlike past resources, the rhythm of fixity and fluidity varies from account to account.

4. *There is no particular societal benefit to using public resources this way*.

   The absolute banality of what's in services like Facebook and Instagram makes people doubt their long-term value. Less than 1% of those we surveyed thought there was any benefit to saving user-contributed social media content.

On the face of it, these objections may seem to us both obvious and silly. They're true of many of today's non-digital personal archives as well. In fact, the content of a Facebook archive might seem less personal than the type of letters people used to write and receive and less revealing than the information that was published in some small town newspapers and city directories. And certainly many of these point-to-point or local sources of personal information have been replaced by Facebook and other social media services. Yet these are real concerns that we have encountered in many studies (both our own and those performed by others), and they'll need to be addressed before there is broad public support for institutional social media archives.

They also suggest design decisions that would need to be made before an institution reflexively embarks on such an expensive and potentially intrusive endeavor. Because social media archiving necessarily involves a considerable investment of resources, we need to pick which social media service(s) we want to archive. Agreements must be struck, and decisions must be made about which parts of the selected service to collect. For example, it would be very different to archive only the public content on Facebook, or only the social network (say, profiles and their interconnections) than it would be to collect the entirety of the content (comments, messages, likes, photo albums, and other user-contributed content).

The unpredictable revision of social media content and profiles suggests that it is important to determine how frequently snapshots of the service would be collected. Furthermore, because removal is as important (if not more important) than any other kind of user interaction, decisions about how often to collect social media and

what to do about deleted accounts and material are crucial to balancing between the acceptability and utility of the archive.

Just as collection policies are important, so too are access policies and embargo periods. Frank Shipman and I have explored two mitigating conditions: limiting access to the social media archive, and creating an embargo period. These are roughly the same types of conditions placed on resources like the US Census. But they are complicated by the fact that this material is digital: how can we apply the usual digital safeguards to such an extensive resource so that it is both safe (from a technical perspective) and secure (from a social/privacy perspective)?

Yet, in spite of the complexity, cost, and potential unpopularity of creating institutional social media archives, the benefit of doing so is clear and the need to begin (or to some, continue) down this road is urgent and vital.'

6.3. Case Study: Digital Preservation Planning

Author: Paul Wilson's collection of work documents was initially established in 1981 to gain practical experience of the emerging field of office automation. Once established, the collection became, de facto, an integral and essential part of his working life. Most of his work documents were stored in it up until his retirement in 2012. He has also applied the filing techniques developed for his work documents to some of his domestic collections, including those of mementos and photos.

Summary

Paul Wilson's case study summarizes a fuller description of his findings (to be published with the accompanying Toolset by the DPC in 2016) when trying to find a suitable preservation planning process and associated documentation to apply to his personal digital collections. Since he could find no preservation planning process appropriate to individuals, he obtained a slide set detailing a simple preservation workflow from the Digital Preservation Coalition, and used that as a base on which to establish an approach to the work. This general approach and accompanying documentation was tested and refined on two of his personal digital collections (one of 800 mementos and the other of 17,000 photos). Template documents were then derived from the results.

Introduction

'My collection of work documents was initially established in 1981 to enable me to gain practical experience of the emerging field of office automation when I was working in the newly constituted Office Systems Division at The National Computing Centre in Manchester. In 1987, my index for the collection was computerized using Filemaker software; in 1996 I acquired a scanner and software and started digitizing the documents themselves. The index comprises some 17,000 entries representing about 170,000 scanned pages, 4,000 Word files, 400 Excel files, 250 PowerPoint files, 150 files of various other types and around 40 CDs containing a wide variety of material. The index and all the electronic files are stored and accessed on a laptop computer running MS Windows. Some 450 special documents have been retained in their original physical form and are stored in three archive boxes.

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, and that the cost of purchasing and implementing upgrades to the (expensive) Index and Document Management software is becoming prohibitive. While investigating how I could address these problems, I came across the field of digital preservation and was heartened by the huge amount of work that has already been done on this subject. However, a closer inspection revealed that most digital preservation models and approaches are, unsurprisingly, designed specifically for the professional curation of digital materials owned by organizations. As such, their detailed processes assume far greater levels of data

volumes, time, budget, facilities, commitment and organizational constraints than individuals are likely to have.

The two sources of materials I found that provide guidance specifically for individuals are the PARADIGM (Personal Archives Accessible in Digital Media) project conducted by the research libraries of the Universities of Oxford and Manchester, and the content on the Personal Archiving section of the website of the Library of Congress. PARADIGM is aimed at taking the papers of individuals into the digital repositories of large institutions, and has an extensive website of publicly available supporting material, including an Appendix of 'Guidelines for Creators of Personal Archives' which includes 'Eleven Top Tips for Preserving Your Personal Data' (PARADIGM, 2005). The Library of Congress's site provides general guidance, including a document advising '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). Although generally helpful, neither set of resources provided a detailed preservation planning description of the sort I was looking for. In fact, none of the models, approaches and guidance that I came across provided a process description of preservation planning which I felt would meet my needs as an individual.

In addition to searching online to find a suitable preservation planning workflow, I approached a number of individuals and organizations working in the field. One of these was the Digital Preservation Coalition (DPC), which kindly provided me with a 12-step preservation planning exercise sheet used in a DPC Training Programme called 'Preservation Planning: from theory to practice' (Kilbride, 2013). This seemed to be the closest I could get to what I was looking for. The DPC also advised me to take a closer look at the Plato tool.

Developing an approach

To explore what specific approach I was going to take, I tried out both the Plato tool (version 4) and the DPC 12-step preservation planning sheet template on a small, uncomplicated collection of mementos and other personal electronic documents (PERS).

Plato is a free comprehensive web-based tool which you have to register to use and which you can then apply to any number of collections (PLATO, 2015). It takes you through a workflow which includes inputting sample documents and defining experiments to test preservation options. However, it doesn't provide any guidance on what those options could be. Although it is clearly aimed at institutions with large complex collections, as an owner of relatively small collections I found it useful to experience these preservation steps.

The DPC 12-step questionnaire is a much more flexible tool, which, while also being designed for large institutional collections, is much more flexible and is easily adapted for use for small personal collections. Since neither tool fully met my needs, I derived a target set of three preservation planning documents (a Scoping document, a Project Plan, and a Maintenance Plan). I then developed and refined the documentation in two separate trials: first, with my PERS collection (approximately 800 files of various uncomplicated file types with an Excel index), and then with my PHOTOS collection (approximately 17,000 files, mostly, but not exclusively, JPGs, with an Excel index).

The trials

I started the first trial by using the DPC 12-step preservation planning exercise to create a scoping document which I then used to create a Preservation Project Plan for my PERS collection. I researched file formats as best I could on the Internet and concluded that the best preservation approach for my files would be to convert them to the PDF/A format using a more recent version of the PDF editing software that I use – eCopy PDF Pro Office. 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, including PDF/A-1a, PDF/A-1b, PDF/A-2a, PDF/A-2b, and PDF/A-2u – all of which were supported by my updated e-Copy software. The process of selecting a particular PDF/A version, converting the documents to PDF from other formats, and then troubleshooting errors turned out to be a very

tortuous process, so I decided that I was only going to check and ensure PDF/A-1b compliance for the files that, at the start of this exercise, were not in PDF format.

I mention some of 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 happy with), but to highlight 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* you create the Preservation Project Plan.

I undertook the second trial on my PHOTOS collection, using the insights gained from the first trial to create the following pre-planning task list in the scoping document:

- Decide what different types of backup arrangements are to be put in place.
- For each file type:
    a) open up a few example files;
    b) decide what application is preferred to open the file type in;
    c) define what conversion action, if any, is to be taken;
    d) 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 forward.
- Decide what cross referencing should be included in each of the different types of components (e.g., electronic files, index, physical albums, suitcase containing negatives).
- Decide whether to discuss the collection with the potential future recipients.

Completing this pre-planning work enabled me to create a more realistic and reliable Preservation Project Plan for my PHOTOS collection.

I am confident that the process steps and associated documentation that have emerged in these trials will scale up to enable me to undertake preservation planning on my somewhat larger and more complex collection of work documents.'

Future plans

Paul Wilson's template documents for personal digital archives preservation, derived from the results of these trials, as well as a more complete description of his research and trials, will be published by the DPC in 2016. While developed specifically for personal collections, these may be useful for undertaking a digital preservation exercise on any relatively small collection when unconstrained by organizational policies.

7. Conclusions

7.1. Reasonable steps forward

The most difficult step in creating and maintaining a personal digital archive is often figuring out where and how to begin. A good first step is to make a back-up copy of your files and store it in a different location from where your primary computer or other storage resides. That way, even if you never get around to organizing your files, at least you will have a back-up copy in case of disaster. A reasonable next step would be to spend some time figuring out what your digital archive contains and which items hold the most significance. The following questions can help guide this process:

- Which files would you miss most if they suddenly disappeared?
- What qualities about those files are most important – for example, does it matter if the formatting of a word-processing document changes if the text is still readable?
- Do your digital photos include important descriptive or contextual information that you need to use a particular program to see?

The answers to these and related questions will help you prioritize the files in your archive. Another way to decide which files need more active management is to consider which items are most at risk. These may be the oldest materials, or the ones that have been stored in harmful physical conditions. Once you have identified which files are the most valuable – either because of their content or because they fall into a high-risk category – the next step is to begin more actively organizing, managing, storing, and preserving them.

7.2. Managing expectations

Caring for your digital files will take deliberate effort, every step of the way. It is important to be realistic about the amount of time you are willing to devote to organizing and describing your files. Once you have identified which files are most at-risk or hold the most personal significance, prioritize the items that need immediate care and also identify which items can wait until you have more time. Accept that some file loss is inevitable – for example, files that have been corrupted, reside on damaged media, or were created using software that you no longer have access to. Do the best you can with the resources you have available.

7.3. Resource implications

Although resources are necessary for you to take good care of your personal digital archive, it is not difficult to make significant progress with a limited budget as long as you have time to devote to the work. Creating an inventory of your physical media (e.g., disks) and digital files will help you identify the most important files in the archive. A free 30-day trial of a tool developed for analysing a hard drive can be repurposed to provide a bird's-eye view of the contents of a laptop or large-capacity external hard drive. A few cloud-based storage providers offer a certain amount of storage for free; select one of these services and use it to store a back-up copy of your most important files for no or little cost. A slightly more expensive alternative would be to store a back-up copy of your most important files on a USB flash drive, store the drive in a different geographical location from your original files, and plan to transfer your files to fresh media every two to four years. Individuals who want to explore the contents of files in obsolete formats can install the free trial version of software designed to view the contents of older files (see Section 6.3 for resources that give specific software recommendations).

A moderate budget will allow you to pay for more storage and multiple back-up copies, purchase full versions of key software rather than relying on free trials, and also begin the work of accessing files on older media. External USB 3.5-inch floppy drives are readily available online for a moderate price and work well with modern computers. Devoting the time to copying files from older media to a more modern machine, or from an older to a newer computer, and migrating them into a different format will allow you to identify and

preserve older files as well as those in more current formats. When copying files from one medium or software environment to another, try not to alter the date/time stamp or file name if that information is integral to the file's significance to you.

With a generous budget, you could pay a vendor to copy files from all of your obsolete media and machines and store them securely. Additional resources would pay for a storage service that closely monitors the health and security of files stored in the cloud and maintains a secondary backup in case of data loss. It would also be possible to scan or digitally photograph all of the paper documents in an archive and store them alongside the born-digital files to create a comprehensive digital archive which could be managed as a unified digital corpus.

7.4. A recap of steps you can take to preserve your own materials

Personal digital archives are unwieldy and slightly mysterious. They are often stored in several different locations, both local and cloud-based; they contain both vintage and modern file formats; they consist of a range of files that represents the incredible technological developments of the late 20th and early 21st centuries; and they include many files that people have forgotten ever existed. What can you do in the face of such challenges? The basic steps are straightforward and require minimum effort:

- figure out what your archive contains so that you can prioritize certain files over others and focus your efforts.
- develop a file naming convention that makes sense to you (and others), and apply it consistently.
- back up your files periodically to multiple storage devices kept in different locations from the primary files.
- choose high-quality storage media if storing your files locally; if using cloud-based storage, research the different options carefully before deciding which services best meet your needs.
- periodically refresh the media on which your files are stored.
- make sure that your back-up files are up to date across all storage media and services.

Beyond these basic steps, a higher level of effort and commitment would require you to actively monitor the health of your files once they are in storage. This might include assessing the fragility of particular file formats and developing a preservation strategy that anticipates and actively mitigates file corruption and format obsolescence.

7.5. Timescales

It can be a challenge to find the time to manage your own personal digital archives. Because digital media and files will continue to degrade, it is vital for you to check the health of your digital media and the accessibility of your files at least once per year, if not more frequently. The estimated life expectancy for physical audiovisual media has been estimated at 10 to 15 years; by 2028, it is expected that audiovisual tape media such as VHS tapes will have reached the end of their lifespan, which means that personal collections containing these items, as well as magnetic media such as floppy disks, are at high risk for loss (Lacinak, 2013; IASA, 2014). Keep this dire prediction in mind when prioritizing which files to copy from older media. It is also important to refresh back-up storage media periodically to combat hardware failure or obsolescence. If your digital files are stored locally on a flash drive, transfer them to a new flash drive every one to two years; if the storage medium is sturdier, e.g., an external hard drive, consider moving your files to new media every three to four years.

8. Current Activities and Resources Devoted to Personal Digital Archiving

The following resources provide guidance and information on tools, research, organizations and activities devoted to personal digital archiving. Some of these resources are written with an audience of professional librarians and archivists in mind, while others offer advice specifically for members of the general public.

8.1. Guidance

All URLs were current at the date of publication

Library of Congress (US)

http://digitalpreservation.gov/personalarchiving/

Extensive network of resources designed to help individuals and professionals capture and preserve personal digital archives of many types.

The National Archives (UK)

Information about digital preservation, including guidance and case studies specific to cloud storage and a helpful list of more general digital preservation resources: http://www.nationalarchives.gov.uk/archives-sector/digital-collections.htm

Among other resources, includes an introduction to archives for non-archivists: http://www.nationalarchives.gov.uk/archives-sector/archives-records.htm

Digital Preservation Coalition (DPC)

A non-profit membership organization that serves as an advocate for digital preservation. The DPC offers advice to help you get started with digital preservation, as well as information about publications and training opportunities.

http://www.dpconline.org/advice

PARADIGM workbook on personal digital papers (Bodleian Library and John Rylands Library)

http://www.paradigm.ac.uk/workbook/

Best-practice guidelines for archiving personal digital papers, including advice for surveying collections, issues to consider when deciding which files to keep, legal concerns, and an appendix containing useful recommendations for creators of personal archives. Many of the workbook sections are geared towards a professional audience.

AVPreserve (formerly AVPS)

A US-based consulting and software development firm that develops tools, conducts research, and publishes papers and presentations that advance the preservation of digital and analog audiovisual materials.

http://www.avpreserve.com/

Columbia University Libraries

An extensive list of resources, aggregated from different sources, that addresses best practice, media types, backup and recovery, file formats, file naming, and online storage, among other topics.

http://library.columbia.edu/locations/dhc/personal-digital-archiving/online-resources.html

Bentley Historical Library (University of Michigan)

Best practices for directory and file naming, digitizing photographs, and selecting file formats.

http://bentley.umich.edu/giving/donate-your-archives/

National and State Libraries Australasia (NSLA)

Guidelines for library staff assisting donors to prepare their personal digital archives for transfer to NSLA libraries. Provides a helpful overview of personal digital archives and basic advice for staff tasked with making digital acquisitions.

http://www.nsla.org.au/sites/www.nsla.org.au/files/publications/NSLA.Guidelines_donor_digital_archives_20111111.pdf

Demystifying Born Digital Reports (OCLC Research)

A series of accessible, informative reports that lay out first steps, tools, software, and other resources for surveying and gaining control over natively digital files in archives.

http://www.oclc.org/research/publications/library/born-digital-reports.html

North Carolina Department of Cultural Resources

A series of YouTube videos, some of which are closed captioned, providing an overview and practical advice on topics such as saving your Facebook data, file naming, and managing and preserving digital photos.

http://digitalpreservation.ncdcr.gov/tutorials.html

8.2. Grassroots groups and activities

Home Movie Day (Center for Home Movies)

Grassroots effort to unite amateur filmmakers and film archivists in a common goal of preserving personal home movies.

http://www.centerforhomemovies.org/hmd/

StoryCorps

An independent, non-profit project that encourages people to record, share, and listen to each other's stories.

http://storycorps.org/

Community Archives and Heritage Group (CAHG)

Cataloging guidelines for community archives.
http://www.communityarchives.org.uk/content/resource/cataloguing_guidelines

8.3. Software resources

Max Eddy, 'The Best Digital Scrapbooking Software for 2015', *PC Magazine* (UK), 22 December 2014.

Compares 10 different types of digital scrapbooking software and ranks them according to selected criteria.

http://uk.pcmag.com/photo-management-sharing-products/38461/feature/the-best-digital-scrapbooking-software-for-2015

Donald T. Hawkins, 'Software and Services for Personal Archiving', in D. Hawkins (ed.) *Personal Archiving: Preserving our Digital Heritage,* Medford, NJ: Information Today, Inc., 2013.

Reviews several software products that facilitate personal digital archiving, including products and services for organizing and archiving photographs, documents, email, and videos.

Kari Smith, 'Tools for Understanding Digital Files', MIT Libraries.

Describes and provides links to a variety of software tools that can be used to analyse the contents of a hard drive, confirm that files have copied correctly from one location to another, view the contents of older file formats, and many other helpful functions. The summary is written with professional archivists in mind, but some of the tools will also be useful for individuals working with their own digital files.

http://libguides.mit.edu/digitalarchivestools

8.4. Professional/Academic Writing and Resources

Catherine C. Marshall

Publications related to personal digital archiving, 2006–2013.

http://www.csdl.tamu.edu/~marshall/pubs.html#personalDigitalArchiving

Catherine C. Marshall and Frank M. Shipman

Research exploring Facebook and the value of social media, as well as the idea of an institutional archive of Facebook.

http://www.csdl.tamu.edu/~marshall/DL14-marshall-and-shipman.pdf

http://www.csdl.tamu.edu/~marshall/CSCW15-final-final.pdf

Clifford Lynch, 'The Future of Personal Digital Archiving: Defining the Research Agendas', in *Personal Archiving: Preserving Our Digital Heritage*, Donald T. Hawkins (ed.). Medford, NJ: Information Today, Inc., 2013.

http://www.cni.org/wp-content/uploads/2013/09/Personal-Digital-Archiving-Cliff-Lynch-Oct-29-2013.pdf

Digital Lives Research Project (British Library)

A research project investigating both the practical and theoretical aspects of caring for and managing personal digital archives, as well as other types of digital materials, throughout their lifecycle. The project's archived website provides links to blog posts and publications relevant to personal digital archives, and the comprehensive final report shares research data and findings, provides an extensive bibliography and resource lists, and gives an in-depth analysis of many of the key challenges and opportunities presented by personal digital lives.

www.webarchive.org.uk/ukwa/target/9175069/source/subject

http://britishlibrary.typepad.co.uk/files/digital-lives-synthesis02-1.pdf

Personal Digital Archiving Conference

Annual event that brings together scholars, professionals from the cultural heritage sector, archivists, librarians, practitioners, and others to discuss challenges related to personal digital archiving.

http://personaldigitalarchiving.com/

Donald T. Hawkins (ed.), *Personal Archiving: Preserving Our Digital Heritage*, Medford, New Jersey: Information Today, Inc., 2013.

Christopher A. Lee (ed.), *I, Digital: Personal Collections in the Digital Era*, Chicago: Society of American Archivists, 2011.

Sarah Kim. *Personal Digital Archives: Preservation of Documents, Preservation of Self* (dissertation). University of Texas at Austin, 2013.

http://repositories.lib.utexas.edu/handle/2152/21134

8.5. Digital Legacy

Jack Schofield, 'What happens to your Facebook account when you die?', Ask Jack, *The Guardian*, 30 October 2014.

Explores some of the different tools and strategies people can use to plan for their digital legacy, with particular attention to Facebook's deletion and memorialization processes. At the end of the article Schofield includes a list of additional resources; the comments after the article also provide recommendations for different types of services to help individuals manage the details of their digital lives in preparation for death.

http://www.theguardian.com/technology/askjack/2014/oct/30/what-happens-to-your-facebook-account-when-you-or-a-loved-one-dies

Jacqui Shine, 'You Can Delete, but You Can't Forget', *The Atlantic*, 3 July 2014.

Explores the frailty of email and storage media through the lens of grief.

http://www.theatlantic.com/technology/archive/2014/07/you-can-delete-but-you-cant-forget/373662/

Saga, 'A guide to preserving your digital legacy', February 2015.

Saga, a British company that offers a variety of resources and services aimed at people aged 50 and over, has published a helpful guide that outlines the main issues that can arise after a person's death, as well as step-by-step instructions on how to move forward with preparing a digital legacy.

http://www.saga.co.uk/saga/media/Legal/Digital%20legacy%20guide%20Feb%2015/Digital%20Legacy%20guide.pdf

Saga, 'Digital Legacy: How to deactivate online accounts', February 2015.

This document lists several of the most commonly used email and social media services, online retailers, gambling websites, transport and airline companies, and online entertainment providers and gives instructions for how to deactivate accounts with each service provider.

http://www.saga.co.uk/saga/media/Legal/Digital%20deactivation%20document%20Feb%2015/Digital%20deactivation%20document.pdf

Jenny Kleeman, 'Web immortality: the social media sites that keep you alive in the digital world', *The Guardian,* 7 June 2014.

Exploration of services available to help people manage their digital afterlives, with a particular focus on DeadSocial, a service that allows people to upload social media content to be released after their death.

http://www.theguardian.com/lifeandstyle/2014/jun/07/web-immortality-social-media-sites-alive-die-digital

8.6. Audio and video recordings

The following resources are intended to help readers locate online or regional resources to help with the management, conservation or digitization of audiovisual media and files.

Activists' Guide to Archiving Video.

Describes itself as 'A practical resource to help you manage, authenticate, and preserve your digital video'.

http://archiveguide.witness.org/

DJM Digital

A UK-based media production company that can transfer files from video to DVD

http://www.djm.co.uk/

Digital Media Services at BBC Studios and Post Production

Offers digitization and restoration of audio and video materials in many different formats

http://www.bbcstudiosandpostproduction.com/our-services/digital-media-services/

Activist Archivists (in collaboration with Witness, an 'international organization that trains and supports people using video in their fight for human rights')

'7 Tips to Ensure Your Video Is Usable in the Long Term'

http://activist-archivists.org/wp/?page_id=328

The MediaPreserve

An audiovisual laboratory that can transfer all formats of audio, video, and film.
http://ptlp.com/en/mediapreserve/overview/about-us

Pathéscope 9.5

Focus point for all 9.5mm film collectors and silent film enthusiasts. Serves as a place to exchange information and ideas and keep alive interest in 'obsolete' film gauges.

http://www.pathescope.freeserve.co.uk/Pathe95.htm

Film Forever: Home film preservation guide.

Focuses on the 'most common types of films found in private collections and give[s] basic information on implementing a practical preservation strategy for film materials'.

http://www.filmforever.org/

Internet Archive, Moving Image Archive, and Home Movies collection

Contains a wide range of digital movies uploaded by Archive users, as well as a discrete collection of home movies.

https://archive.org/details/movies

https://archive.org/details/home_movies

8.7. Web archiving

Jill Lepore, 'The Cobweb: Can the Internet be archived?', *The New Yorker*, 26 January 2015.

An in-depth exploration of the Internet Archive and the persistence and fragility of the Web.

http://www.newyorker.com/magazine/2015/01/26/cobweb

International Internet Preservation Consortium (IIPC)

An organization devoted to improving the tools, standards, and best practices of web archiving by means of collaborative working groups, projects and initiatives.

http://netpreserve.org/web-archiving/overview

9. Summary of Key Points and Resources for Curators

Curators who are in regular contact with individuals about their personal digital archives may find themselves in a position to offer guidance about how best to care for and preserve the digital files in those individuals' possession. An important first step is to ask the individual open-ended questions designed to elicit information about which files in their archive they consider to be most important, what they perceive as the key threats to their well-being, and what types of hardware and software they use regularly. Follow-up questions aimed at ascertaining the relationship between the different information management systems used by the individual (e.g., databases, email, different types of file storage) will help the curator gain a clearer picture of the entire corpus of digital items at stake. Key issues to offer guidance about include:

- file naming best practices
- an individual's intentions – what do they intend to save, pass on to family or friends, or transfer to a repository?
- the importance of secure, backed-up storage

The 'Collection development' and 'Working with records creators' chapters in the Paradigm project workbook produced by the Bodleian and John Rylands Libraries between 2005–07 remain the best sources for information about how to approach working with individuals on personal digital archives (Paradigm, 2005). The two survey templates included in *Appendix F* of the 2012 white paper that came out of the Mellon-funded AIMS project, a partnership between the University of Virginia, the Stanford University Libraries and Academic Resources, the University of Hull Library, and Yale University Library, provide valuable guidance for curators considering working with a donor to survey their digital collection (AIMS, 2012). In addition, the 2013 CLIR report *Born Digital: Guidance for Donors, Dealers, and Archival Repositories* offers guidance to a broad audience and explores certain elements and challenges relevant to curators (Redwine *et al*., 2013). Several of the resources listed in Section 6.1 offer guidance and best practices that will be of use to curators and other professional staff involved in the selection and curation of personal digital files.

The possible uses for a digital archive are as varied as the individuals who create personal files. For many people, a personal digital archive will be a living thing – a body of files that is constantly being added to, revised, used and reorganized for personal purposes. Other people assemble a digital archive with the goal of passing it along to another caretaker or transferring it to an institution or community archives. A digital archive can be actively used, or it can be stored away and never accessed. Still other individuals assemble personal files for their own private use and intend to securely destroy them without ever sharing. The afterlife of digital files when the creator dies raises questions about what level of documentation is necessary to know someone's wishes. What *can* be done with a digital archive is not always the same as what *should* be done with it.

10. Glossary

Archive: A term used by different disciplines to mean, alternately, a body of materials no longer in active use, a back-up copy of important files, records that have been selected for long-term retention, and a collection of things, memories, and even feelings. In this report, the term 'archive' is used loosely to refer to a distributed body of digital files, papers, photographs, and other physical objects that were created by or about an individual and hold special significance.

Born digital: An item that began life in digital form is considered 'born digital'. A Microsoft Word document, an email, and a digital photograph are all examples of born-digital files.

Cloud storage: A service provided by an organization or company that stores files on remote servers. Although from a user's perspective files are stored 'in the cloud', as opposed to on the user's hard drive, in reality the files reside on physical servers managed by the service provider.

Digitized item/digital surrogate: A digital copy of an original item with a non-digital form. Examples include a digitally reformatted copy of an original VHS tape, a digital scan of a Polaroid photograph, and a digital photograph of a handwritten letter from the 1890s.

Digital media: A term that encompasses many different types of media that house digital files, including, but not limited to, floppy disks, Zip disks, CDs, hard drives, and flash drives.

File format: The encoding that characterizes a particular type of file and determines how the bytes are stored within it. File format can be determined in a few different ways; the most common indicator, which is not always reliable, is file extension. Microsoft Word documents (.doc), Portable Document Format documents (.pdf), and Joint Photographic Experts Group images (.jpg) are common examples of file formats.

Personal archiving: The act or process of aggregating both digital and non-digital photographs, letters, mementos, documents, and other items that one intends to keep.

11. References

AIMS Work Group 2012, *AIMS born-digital collections: An inter-institutional model for stewardship*, online at: http://dcs.library.virginia.edu/aims/white-paper/

Beagrie, N 2002, 'An update on the Digital Preservation Coalition'. *D-Lib Magazine* 8, online at: http://www.dlib.org/dlib/april02/beagrie/04beagrie.html

Beagrie, N 2005, 'Plenty of room at the bottom? Personal digital libraries and collections'. *D-Lib Magazine* 11, online at: http://www.dlib.org/dlib/june05/beagrie/06beagrie.html

Bentley Historical Library, 'Digital Recordkeeping Best Practices for Directory and File Naming', online at: http://bentley.umich.edu/giving/donate-your-archives/digital-recordkeeping-best-practices-for-directory-and-file-naming/

British Library, Digital Lives Research Project
http://www.webarchive.org.uk/ukwa/target/9175069/source/subject

Casserly, M 2015, 'Best cloud storage services 2015: Dropbox vs Google Drive, OneDrive, iCloud & more'. *PC Advisor*, online at: http://www.pcadvisor.co.uk/test-centre/internet/3506734/best-cloud-storage-dropbox-google-drive-onedrive-icloud/

Community Archives Development Group (CADG) 2007, 'The Impact of Community Archives: Summary of independent research commissioned by the Community Archives Development Group', online at: http://www.aughty.org/pdf/impact_comm_archives1.pdf

Digital Preservation (website), Library of Congress. http://www.digitalpreservation.gov/

Editorial Board 2015, 'Digital neglect at the Library of Congress'. *New York Times*, online at: www.nytimes.com/2015/04/05/opinion/sunday/digital-neglect-at-the-library-of-congress.html

Gordon, W 2015, 'A beginner's guide to encryption: What it is and how to set it up'. *lifehacker*, online at: http://lifehacker.com/a-beginners-guide-to-encryption-what-it-is-and-how-to-1508196946

International Association of Sound and Audiovisual Archives (IASA) 2014, 'Handling and Storage of Audio and Video Carriers'. IASA-TC 05, online at: http://www.iasa-web.org/handling-storage-tc05

John, JL, Rowlands, I, Williams, P and Dean, K 2010. *Digital Lives: Personal Digital Archives for the 21st Century: An Initial Synthesis*. Beta version 0.2, online at: http://britishlibrary.typepad.co.uk/files/digital-lives-synthesis02-1.pdf.

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

Kirschenbaum, M, Farr, E, Kraus, K, Nelson, N, Peters, C, Redwine, G and Reside, D 2009, *Approaches to Managing and Collecting Born-Digital Literary Resources for Scholarly Use*. White paper for NEH Office of Digital Humanities, online at: http://mith.umd.edu/wp-content/uploads/2012/03/whitepaper_HD-50346.Kirschenbaum.WP.pdf

Kleeman, J 2014, 'Web immortality: the social media sites that keep you alive in the digital world'. *The Guardian*, online at: http://www.theguardian.com/lifeandstyle/2014/jun/07/web-immortality-social-media-sites-alive-die-digital

Lacinak C 2013, 'The end of analog media: The cost of inaction and what you can do about it'. Association of Moving Image Archivists (AMIA) annual conference, 8 November 2013, online at: http://www.avpreserve.com/wp-content/uploads/2013/11/Lacinak_COI_AMIA_2013_dist.pdf

Lee, C and Capra, R 2011, 'And Now the Twain Shall Meet: Exploring the Connections between PIM and Archives'. In *I, Digital: Personal Collections in the Digital Era*. Chicago: Society of American Archivists.

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

Marshall, CC 2008, 'Rethinking personal digital archiving, part 1: Four challenges from the field'. *D-Lib Magazine* 14.

Microsoft Research, MyLifeBits. http://research.microsoft.com/en-us/projects/mylifebits/default.aspx

NDIIPP 2013, *Perspectives on Personal Digital Archiving*. National Digital Information Infrastructure and Preservation Program, online at: http://www.digitalpreservation.gov/documents/ebookpdf_march18.pdf

Osterberg, G 2013. Update on the Twitter archive at the Library of Congress. Library of Congress Blog, online at: http://blogs.loc.gov/loc/2013/01/update-on-the-twitter-archive-at-the-library-of-congress/

O'Neill, I 2013, 'The future of lifelogging – Interview with Gordon Bell'. *Autographer* blog, online at: https://web.archive.org/web/20150402102034/http://blog.autographer.com/2013/05/the-future-of-lifelogging-interview-with-gordon-bell/

Palmer, D 2014, 'iCloud celebrity photo hack: Are we too quick to trust cloud storage?' *Computing*, online at: http://www.computing.co.uk/ctg/analysis/2362822/icloud-celebrity-photo-hack-are-we-too-quick-to-trust-security-of-cloud-storage-services

Paradigm Project 2005–7, *Workbook on Digital Private Papers*, online at: http://www.paradigm.ac.uk/

PLATO (Preservation Planning Decision Support Tool) 2015, version 4, online at: http://www.ifs.tuwien.ac.at/dp/plato/intro/

Redwine, G, Barnard, M, Donovan, K, Farr, E, Forstrom, M, Hansen, W, John, JL, Kuhl, N, Shaw, S, and Thomas, S 2013, *Born Digital: Guidance for Donors, Dealers, and Archival Repositories*. CLIR, online at: http://www.clir.org/pubs/reports/pub159

Rothenberg, J 1999, *Ensuring the Longevity of Digital Information*. Council on Library and Information Resources, online at: http://www.clir.org/pubs/archives/ensuring.pdf

Saga 2015a, *Guide to Preserving Your Digital Legacy*, online at: http://www.saga.co.uk/saga/media/Legal/Digital%20legacy%20guide%20Feb%2015/Digital%20Legacy%20guide.pdf

Saga 2015b, *Digital Legacy – how to deactivate online accounts*, online at: http://www.saga.co.uk/saga/media/Legal/Digital%20deactivation%20document%20Feb%2015/Digital%20deactivation%20document.pdf

Saltzman, D 2010, 'Snooping Through Salman Rushdie's Computer'. *The Atlantic*, online at: http://www.theatlantic.com/entertainment/archive/2010/04/snooping-through-salman-rushdies-computer/39260/

Schofield, J 2014, 'What happens to your Facebook account when you die?'. Ask Jack, *The Guardian*, online at: http://www.theguardian.com/technology/askjack/2014/oct/30/what-happens-to-your-facebook-account-when-you-or-a-loved-one-dies

The Signal: Digital Preservation (blog), Library of Congress, online at: http://blogs.loc.gov/digitalpreservation/

Tuck, J 2004, 'The British Library Web Archiving Programme'. Online at: http://www.bcs.org/upload/pdf/jtuck-141204.pdf

UK Web Archiving Consortium (UKWAC), Digital Preservation Coalition.
http://www.dpconline.org/advice/web-archiving