Born Digital: Guidance for Donors, Dealers, and Archival Repositories

by Gabriela Redwine, Megan Barnard, Kate Donovan, Erika Farr, Michael Forstrom, Will Hansen, Jeremy Leighton John, Nancy Kuhl, Seth Shaw, and Susan Thomas

October 2013

Born Digital: Guidance for Donors, Dealers, and Archival Repositories

by Gabriela Redwine, Megan Barnard, Kate Donovan, Erika Farr, Michael Forstrom, Will Hansen, Jeremy Leighton John, Nancy Kuhl, Seth Shaw, and Susan Thomas

October 2013

Cover photo: © vectomart/bigstockphoto.com

Contents

About the Authors

Gabriela Redwine is digital archivist at the Beinecke Rare Book and Manuscript Library at Yale University. Previously, she was archivist and electronic records/metadata specialist at the Harry Ransom Center. In 2010, she co-authored, with Matthew Kirschenbaum and Richard Ovenden, *Digital Forensics and Born-Digital Content in Cultural Heritage Collections* (CLIR).

Megan Barnard is associate director for acquisitions and administration at the Harry Ransom Center, where she has worked to acquire the archives of contemporary authors such as David Foster Wallace, Julia Alvarez, J. M. Coetzee, and Norman Mailer, among others. She served as curator of the exhibitions *Literature and Sport* (2013) and *Culture Unbound: Collecting in the Twenty-First Century* (2011) and co-curator of *The Mystique of the Archive* (2008). She is editor and co-author of *Collecting the Imagination: The First Fifty Years of the Ransom Center* (University of Texas Press, 2007).

Kate Donovan is public services and instruction librarian at the Tamiment Library & Robert F. Wagner Labor Archives at New York University (NYU). Prior to working at NYU, Kate served as the university archivist and coordinator of records management in the Manuscript, Archives, and Rare Book Library at Emory University, where she led long-range business planning, development, and management of the University Archives. Kate holds an M.S. in information studies from the University of Michigan, an M.A. in history from the University of Iowa, and a B.A. in history from Vassar College.

Erika Farr is head of digital archives in the Manuscript, Archives, and Rare Book Library (MARBL) at Emory University. She leads the born-digital archives program, assists with MARBL digital projects, and supports other digital initiatives involving MARBL materials. She received her M.S. in library science from the University of North Texas at Denton and her Ph.D. in English literature from Emory University. Her current research interests include human information behavior in archival settings and digital humanities research methodologies.

Michael Forstrom is an archivist at the Beinecke Rare Book and Manuscript Library at Yale University, where he is responsible for processing and cataloging modern literary collections. From 2003 to 2012, he was also responsible for the stewardship of born-digital archival material.

Will Hansen is assistant curator of collections at Duke University's David M. Rubenstein Rare Book & Manuscript Library. His duties there include the acquisition and curation of printed and archival materials in the fields of literature, Southern history, and economists' papers. Prior to coming to Duke, he worked at the Newberry Library in Chicago. He holds an M.L.I.S. from the University of Illinois.

Jeremy Leighton John belongs to the Digital Research and Curator Team in the Department of Digital Scholarship at the British Library where he has been curator of eMANUSCRIPTS since 2003, having been specialist scientific curator for the W. D. Hamilton Archive. Previously he worked as a cataloger of bioacoustic collections. In 1996 he completed a D.Phil. in Zoology at Merton College, University of Oxford, concentrating on evolutionary and phylogenetic topics. He is a Fellow of the Linnean Society of London and the Royal Geographical Society. As principal investigator of the Digital Lives Research Project funded by the UK Arts & Humanities Research Council, he

promoted the use of digital forensics with an article in the journal *Nature*: "The future of saving our past." In 2012, as project lead for the Personal Digital Manuscripts Project at the British Library, he published a technology watch paper for the Digital Preservation Coalition entitled *Digital forensics and preservation*. He is presently serving on the Committee of the Section for Archives and Technology of the Archives and Records Association of UK & Ireland. Memberships have included the Library Committee of the Royal Society and the Advisory Committee of the National Cataloguing Unit for the Archives of Contemporary Scientists.

Nancy Kuhl is curator of poetry of the Yale Collection of American Literature at the Beinecke Rare Book and Manuscript Library at Yale University (http://beinecke.library.yale.edu/blog/YCAL). She is the author of exhibition catalogs and collections of poetry, and is co-editor of Phylum Press, a small poetry publisher (http://www.phylumpress.com).

Seth Shaw is assistant professor of archival studies at Clayton State University. Previously, he served as the electronic records archivist for Duke University Archives and was responsible for everything born-digital in both the University Archives and special collections. He received his B.S. in information systems from Brigham Young University–Idaho in 2005 and his M.S. in information, archives and records management from the University of Michigan's School of Information in 2007.

Susan Thomas is digital archivist for the Bodleian Library's Western Manuscripts department, where she has responsibility for developing and managing the Library's capacity to curate, preserve, and provide access to born-digital archives. Susan has been involved in digital archives work—projects, day-to-day internal and collaborative work—since 2005. Until recently, she co-tutored the Digital Archiving module on the Masters programs of the Centre for Archive and Information Studies at the University of Dundee. She currently serves as an external examiner to the Archives and Records Management program at University College London.

Acknowledgments

Preface

Until recently, digital media and files have been included in archival acquisitions largely as an afterthought. Rare book and manuscript dealers may be uncertain about how to handle born-digital materials or how to assign monetary and cultural value to them in relation to analog materials. Repository staff may not be prepared to manage incoming digital acquisitions, much less serve as advocates for the physical and digital well-being of born-digital materials throughout the acquisition process. And donors, nervous about the contents of disks and e-mail messages from long ago, may be torn between wanting to preserve and promote their lives and work on the one hand, while also maintaining privacy for themselves, their families, and their friends on the other hand. The idea for a report offering guidelines for born-digital acquisitions arose from direct experience with these and other challenges.

In August 2011, a group of 10 archivists and curators—then at the Beinecke Rare Book and Manuscript Library, Yale University; the Bodleian Library, University of Oxford; the British Library; the David M. Rubenstein Rare Book & Manuscript Library, Duke University; the Manuscript, Archives, and Rare Book Library (MARBL), Emory University; and the Harry Ransom Center, The University of Texas at Austin—began working together to develop recommendations for the acquisition and transfer of born-digital materials to archival repositories. Our objective was to offer guidance to help ensure that digital materials are handled and documented appropriately and arrive at repositories in good condition. The target audience we had in mind included rare book and manuscript dealers, donors, repository staff, and other custodians in possession of digital media and files.

Four of the six institutions were each represented by a digital archivist who worked directly with digital media and an acquisitions specialist who was directly involved with collection development and acquisitions. The remaining two institutions, the Bodleian Library and the British Library, were each represented by one person whose position encompassed both types of expertise. We hoped that including people with both kinds of experience would expand the perspective of the group and, in particular, encourage archivists and curators to talk with each other about born-digital acquisitions.

Because the project had no budget and included participants from the United Kingdom and from different regions in the United States, we used tools such as Vidyo (videoconferencing), Google Docs, and Dropbox to collaborate remotely. We also took advantage of separately funded travel to conferences and other events to meet in person. We shared existing documentation from our own institutions as we began framing guidelines for born-digital acquisitions. Each cohort also contributed a short case study illustrating the current practices and challenges related to born-digital acquisitions

at our respective institutions. This exercise resulted in a confidential yet common set of experiences that we drew from and referred back to as we developed the structure of the report and drafted preliminary content. Our working model hints at the benefits of continuing intra-institutional collaborative work in the future and opening similar channels of communication between repository staff and dealers.

By September 2012 we had a complete draft of the report and enlisted the help of two additional colleagues, Erin O'Meara (Archivist, Gates Archive) and Kari R. Smith (Digital Archivist, Institute Archives and Special Collections, MIT Libraries), to provide an external editorial review. Their suggestions guided the next stage of our revisions, and by January 2013 a final draft of the report, now called *Born Digital: Guidance for Donors, Dealers, and Archival Repositories*, was ready for public comment.

We decided to pursue two different approaches to publication. First, we published a draft version of the report online with MediaCommons Press, an innovative online publisher committed to open public review. The MediaCommons website provides a commenting interface that facilitates reader feedback. We hoped this functionality would encourage a variety of readers—dealers, archivists, curators, donors, researchers, technologists, students, and other interested parties—to become part of a public conversation about the issues addressed in the report. We also wanted to avoid affiliating this intensely collaborative project with a single institution and to make the review and revision process as transparent as possible. The MediaCommons version of the report will remain open for comment at http://mcpress.mediacommons.org/borndigital.

As soon as the MediaCommons draft went live, we began talking with representatives from more traditional presses about the possibility of publishing a revised version of the report as a no-cost e-publication. Publishing with the Council on Library and Information Resources (CLIR) will help us reach a broad audience that includes librarians and publishers, as well as archivists, scholars, dealers, donors, and students. CLIR's emphasis on the necessity of collaborative work; its status as an independent, nonprofit organization; and its commitment to encouraging discussion about digital curation align closely with the objectives of the report.

Since the project began, several of us have had changes in job title or moved to different institutions. The title page of the MediaCommons draft of *Born Digital* indicates our respective affiliations at the beginning of the project; our current positions and affiliations are listed on pages iv–v of the CLIR report.

October 2013

A Word on Terminology

The authors of *Born Digital: Guidance for Donors, Dealers, and Archival Repositories* fall into two general groups: archivists and special collections curators. We have tried to use terminology that will be clear to all segments of our intended audience, but there remains an unavoidable professional bias. Certain terms used throughout the report will mean different things to archivists, dealers, curators, and donors.

In some instances, we have chosen to sacrifice accuracy for consistency. For example, throughout the report we use the term *donor* to refer to the person, family, organization, estate, or other entity that sells, gives, deposits, loans, or otherwise transfers born-digital materials to an archival repository. In reality, sometimes this person or entity will also be the creator of the digital records or will be selling the materials rather than donating them. As you read the following pages, we invite you to substitute whatever language is necessary to make the report's recommendations relevant within your particular context.

The following are definitions for some of the terms used most frequently in *Born Digital*. These and other definitions have been incorporated into the main body of the report when necessary.

Acquisition. The process by which a repository assumes ownership or responsibility for a body of materials; or, a body of materials recently acquired by a repository.

Dealer. A person or company that sells rare books and manuscripts and brokers deals between donors and archival repositories.

Donor. A person, family, organization, estate, or other entity that sells, gives, deposits, loans, or otherwise transfers born-digital materials to an archival repository.

Repository. A place that acquires, houses, and makes available archival materials transferred from donors.

Technical specialist. A person with expertise in the technical characteristics and functions of some of the following: networks, computers, digital media, file systems, operating systems, and digital files. This person may also have expert knowledge about databases and networked systems.

1. Introduction

Born Digital: Guidance for Donors, Dealers, and Archival Repositories offers recommendations to help ensure the physical and intellectual well-being of born-digital materials transferred from donors to archival repositories. The main body of the report surveys the primary issues and concerns related to born-digital acquisitions and is intended for a broad audience with varying levels of interest and expertise, including donors, dealers, and repository staff.

Each of the following sections of the report provides an overview of a key area of concern:

- **Initial Collection Review** outlines the considerations and approaches that inform interactions among repository staff, donors, and dealers prior to acquisition.
- **Privacy and Intellectual Property** addresses ethical and practical concerns related to intellectual property rights as well as those associated with private and sensitive information.
- **Key Stages in Acquiring Digital Materials** addresses acquisition agreements and contracts, the transfer process, and initial handling once the digital materials arrive at a repository.
- **Post-Acquisition Review by the Repository** focuses on staff assessment of the condition and contents of digital media and files after their arrival at a repository, as well as issues related to retention, disposal, and neglect.

Each of these sections concludes with two lists of recommendations: one for donors and dealers, and a second for repository staff.

Appendixes provide more specific information about how to prepare for the unexpected and possible staffing activities, as well as a list of resources and ready-to-use checklists that incorporate recommendations from throughout the report. These recommendations are not meant to be universal and do not necessarily reflect the official policies of the authors' institutions. Rather, they offer broad, useful guidance for donors, dealers, and repository staff involved in the acquisition of born-digital materials.

To ensure that born-digital materials arrive at repositories in good condition and that appropriate documentation accompanies them, it is vital to convince all parties to be mindful of how they handle, document, ship, and receive digital media and files. The larger benefit and concern, as always, is the preservation of important cultural resources. The following recommendations will help archival repositories, donors, and dealers implement practical improvements that will ultimately lead to richer acquisitions.

2. Initial Collection Review

Born-digital materials present preservation and access challenges that place new demands on the people involved in the acquisition process. It is good practice to share as much information as possible, prior to acquisition, about archives containing digital media or files so that all parties understand better the scope and nature of the born-digital content. Early communication also helps repository staff take preliminary steps to ensure the archival and file integrity, as well as the usability of digital materials over time.

2.1 Preliminary Assessment of Digital Media and Files

Donors and dealers sometimes assess the scope and condition of a body of materials before offering it to a repository. Likewise, repositories have traditionally evaluated archival materials prior to acquisition in order to

- determine whether the content aligns with their collecting interests
- determine potential use or access restrictions
- obtain as much contextual information about the collection and materials as possible
- assess whether the items hold cultural and research value
- decide whether and what to acquire

In recent years, the relentless pace of technological change and the variety of ways in which people use technology have had a significant impact on the types of materials that end up in archives. The complications associated with digital archival materials have further underscored the benefits of an initial collection review. It is important for all parties—donor, dealer, and repository staff—to discuss early on whether digital media and files form a significant part of any materials being offered for gift or sale. These conversations may involve repository staff members who hold accessioning, acquisition, curatorial, legal, processing, and technical responsibilities. Donors may also want to consider involving technical specialists of their own to work closely with repository staff to locate and assess the digital media and files being offered as part of the acquisition.

Preliminary assessment of the digital materials can help ensure that donors transfer only what they intend. Any preliminary collection inventory should include general information about digital media and files; details about the content and context of the materials; and relevant information about the materials' physical condition, such as descriptions of any damage caused by moisture, extreme conditions, or rough handling. Other considerations may include the extent to which the born-digital content is replicated in paper form elsewhere in the collection or its availability to the public online. This information will help repository staff evaluate the desirability of acquiring the born-digital content, predict the storage requirements, estimate the associated staff and equipment costs over time, and

decide whether to make the acquisition (see Appendix A for information about possible staffing activities). A repository's assessment criteria may include

- general technical characteristics of the media or files (e.g., media format, file type, extent)
- volume of digital materials (including size range of files)
- nature of the relationship between born-digital and paper materials within a collection
- information about context and content
- possible transfer options
- particular preservation challenges

Basic repository strategies for assessing born-digital materials include sharing relevant information and documentation, conducting collection surveys, and communicating directly with donors about the history of their digital media and files (see Appendix B for links to sample surveys and other documentation).

2.1.1 Information and Documentation Sharing

Sharing information and documentation can help donors, dealers, and other parties understand a repository's concerns about born-digital materials and will help set expectations about the scope of the acquisition, the processes that staff members will undertake, and the repository's requirements of the donor or dealer and vice versa. Examples of information that repositories may share with donors include

- collection development policy for born-digital materials
- policies and procedures related to the acquisition, transfer, copying, embargo or restriction, user access, long-term preservation, and secure disposal of digital media and files
- guidance in how to handle digital media and files and how to document the process
- guidelines for preparing acquisition agreements or contracts
- sample scenarios illustrating how researchers and repository staff can use born-digital materials

One area of particular concern involves the condition of the digital files in collections offered to repositories. To protect the integrity of digital files—that is, to ensure that the files arrive at a repository with their original content, dates, and other information unchanged—donors and dealers should not manipulate, rearrange, extract, or copy files from their original sources in anticipation of offering the material for gift or purchase.

Even turning on a computer can alter files, and opening files in applications can change date and time stamps, possibly affecting the monetary and cultural value of the materials. In some cases, a donor may want to preview the contents of old disks or files to make certain that he or she does not transfer private material. In the event that a donor or dealer decides to access a disk or look at files on a computer before transferring ownership, it is important that he or she

work with repository staff to document what, if anything, has been done to the digital media and files during the assessment process, when, and by whom. This documentation will help the repository staff establish the provenance of born-digital materials. Strategies for previewing files without altering them include using write-blockers, flipping the write-protect tab on the back of 3.5-inch floppy disks, or using a CD-ROM (Read-Only Memory) drive. Technical specialists at archival repositories may be able to provide useful advice to donors and dealers about how to access files safely. Donors, dealers, and repository staff should also take care to handle digital media (e.g., physical objects such as disks, CDs, flash drives) carefully, in a way that does not damage them or compromise their value.

2.1.2 Collection Survey

A collection survey is a process by which repository staff members gather information about a collection, including the quantity, forms, condition, and location of digital materials. There are two main strategies for conducting surveys of born-digital materials prior to acquisition: an onsite viewing conducted in person and a remote preview conducted over the Internet, via phone, or by some other method.

Onsite assessment has several advantages, particularly when the proposed acquisition comprises a considerable amount of material. Onsite surveys usually take place wherever a donor works or wherever the collection material is stored; it may involve

- preliminary evaluation and discussion of a donor's born-digital materials
- an integrated survey of the digital and paper portions of the proposed acquisition or independent assessments of the two components at different times
- safely copying the donor's files or capturing directory structures for further evaluation upon return to the repository, where staff will have more time and possibly additional tools at their disposal

Before copying a donor's files for the purpose of later assessment, staff should have an agreement with the donor in place, specifying how the files will be copied, stored, and securely deleted in the event that either party decides not to proceed with the acquisition or the repository needs to recapture copies of the original materials.

When a donor and the materials are located some distance from a repository, a remote survey may make more sense than a site visit. Similarly, repositories in geographically isolated locations or with small travel budgets may need to evaluate materials remotely. Remote surveys can be conducted via e-mail or other means of communication. One strategy is to ask the donor to complete a formal survey tool (e.g., a list of initial questions developed by a repository) and then continue the conversation through a further exchange of letters, faxes, e-mail messages, or telephone calls (see Appendix B for links to sample surveys). A remote survey may also include a document created by the dealer or donor that lists the number and type of digital media and provides a general characterization of the contents.

Alternatively, a donor may provide a repository with access over the Internet via a file transfer protocol (FTP) or secure peer-to-peer communication that gives the repository remote access to the donor's desktop. This approach makes it possible for staff or a dealer to preview the files and generate a survey, discuss the logistics of a possible transfer, and decide whether to make the acquisition. Care must be taken not to alter files when accessing a desktop or directory remotely.

2.1.3 Communication

Surveys and other pre-acquisition strategies can be supplemented by conversations with donors about their past and present computing environments. Any information that repository staff can learn about a donor's computing habits, including the environmental conditions under which computer media have been stored in the past, will help them contextualize and preserve the digital materials.

Donors and dealers should ask about a repository's readiness to receive and store digital media and files; strategies for long-term stewardship of digital material; and policies regarding privacy, capture and storage methods, and security. If a donor relies on a technical specialist to manage digital files, it may be helpful to involve that person or team in the collection review process. Guidelines for describing and handling digital items prior to offering the materials to a repository can help all parties anticipate problems and formulate solutions well in advance of the actual transfer of materials. Communication between repository staff and donors about the technical nature and extent of the digital materials in a prospective acquisition, as well as about issues of privacy and confidentiality, access and restrictions, and expectations regarding capture and transfer, can forestall potential processing difficulties and ultimately enable repositories to serve as better stewards of born-digital content.

2.2 Recommendations for Donors and Dealers

- Be aware that simply viewing files can alter them.
- Avoid manipulating, rearranging, extracting, copying, or otherwise altering data residing in the original source media in anticipation of offering the materials for gift or purchase, or preview files in accordance with established guidelines provided by the repository.
- Ask repositories for guidance and documentation on
  - determining the most appropriate repository for the digital materials
  - negotiating the terms of an acquisition agreement or contract as it relates to born-digital materials
  - describing the context and history of the files and media being transferred
  - handling digital media
  - documenting storage, access attempts, copies, and transport of digital media and files
- Clarify expectations about the extent to which the digital materials on offer will be preserved and made available for use.

2.3 Recommendations for Repositories

- Assess born-digital materials prior to acquisition.
- Weigh the cultural and research value of the collection, or components of it, against the cost of capture, ongoing preservation, and access.
- Share relevant information and documentation with donors and dealers about the collection review and acquisition processes, as well as sample use scenarios.
- Clarify expectations regarding digital preservation and access.
- Conduct a survey of born-digital materials:
  - Determine the appropriate survey method by consulting with the donor or dealer.
  - Conduct onsite surveys (in person) when possible.
  - Have policies in place regarding the capture, storage, and disposal of files copied for the purpose of preliminary assessment.
  - Safely capture the donor's files or directory structures when closer analysis or more time is needed than is possible on a site visit or when an onsite survey is not possible.

3. Privacy and Intellectual Property

There is a wide range of opinion and understanding about digital privacy among all groups involved with the acquisition, transfer, cataloging, and description of archival materials. Issues related to the protection of private information and intellectual property in born-digital materials include copyright and other intellectual property considerations, the management of sensitive content in large bodies of e-mail correspondence, legally protected private files, and more technically complicated issues such as password decryption and disk imaging. Whenever possible, donors, dealers, and repositories should discuss and reach consensus on these issues as they relate to the acquisition of born-digital materials.

3.1 Copyright and Intellectual Property

In some ways, dealing with copyright and intellectual property in digital files is relatively straightforward: Donors or other copyright holders may either retain or transfer their intellectual property rights in digital formats just as they may in physical formats. Matters are complicated, however, by the ease with which digital materials can be collaboratively created and shared, and by the desire of archival repositories to provide online or other access to patrons who cannot travel to the repository to use the materials. Furthermore, acquisitions that cross national boundaries can be challenging because of the different copyright and intellectual property laws and practices around the world.

Donors and dealers should know if the digital files that they offer to a repository include other people's intellectual property. For instance, a computer may be shared by coworkers or by an entire family; in the latter case, it may even contain files created by children

and spouses. When possible, donors and dealers should provide repositories with information about the likely primary creators of born-digital materials. As with all recorded information, the intellectual property rights of contractors, contributors, or collaborators must be respected in digital formats, but repository staff may be unaware or unable to determine that a digital file was created by a third party unless told. Agreements transferring digital files to a repository should include provisions governing how the repository will handle third-party digital materials created and accumulated by computer users other than the donor.

Repositories continue to work toward providing access to digital materials while balancing privacy and intellectual property concerns. Acquisition agreements may include capture and access restrictions specific to born-digital materials. Even if a copyright holder retains copyright to all materials in a collection, special considerations such as licensing terms or online access via a limited number of Internet protocol (IP) addresses may enable a repository to publish certain materials online.

3.2 E-mail and Other Digital Correspondence

The sheer volume of sent, received, and saved e-mail messages, as well as the presence of attachments in many different file formats, can complicate e-mail acquisition. The ease with which a message in an e-mail thread can be edited and manipulated without its original author's knowledge or obvious indications to later researchers also differentiates e-mail messages from physical correspondence. In addition, born-digital correspondence may include letters written with word-processing software and sent either as attachments to e-mail messages or printed and sent in physical form. These documents can present complications similar to those encountered in e-mail archives.

Donors may want to screen e-mail files for sensitive or extraneous messages prior to transfer. Repositories and donors will need to make clear who will do the screening and what the process will be, but screening e-mail messages need not be an overly cumbersome or time-consuming process. The easiest strategies may be for the donor to search for keywords and sort by correspondents and time periods that most likely indicate the presence of sensitive messages. If a donor is unable or unwilling to screen for sensitive messages, the repository must decide, in accordance with policy, whether to devote staff time to such a search. If a repository decides to undertake detailed screening for more specific sensitive materials as defined by a particular acquisition agreement, staff resources may necessarily limit strategies for screening.

In some situations, access restrictions or an embargo period on the use of e-mail messages can be implemented as a means of lowering the risk of inappropriately exposing sensitive messages when it is not feasible for staff to undertake screening. As with paper materials, it helps when donors can flag potential areas of concern in digital

materials. In all cases, screening and redaction of files should be undertaken only after digital files have been copied from their original source media.

Preservation for access in the longer term is also a reasonable option and may be the most realistic alternative to screening for many repositories, especially if the nature of the sensitive materials in the e-mail messages is unknown or unspecified. A repository may choose to limit its work on a body of e-mail correspondence to basic maintenance of the digital files, including transfer to secure, stable servers or media. These initial preservation actions can ensure the continuing viability of the files for a future date when better tools or additional staff are available to screen the messages or when all parties involved are deceased.

Donors, dealers, and repositories need to collaborate and discuss how to handle sensitive messages or message threads. If a repository discovers obviously sensitive messages that the donor has not identified and the donor or intellectual property holder is no longer available, the repository must decide whether to provide access to the messages, remove them, or retain them with an access restriction. In many repositories, the amount of labor required to screen e-mail messages may prompt repository staff to explore alternatives such as an embargo when staff resources are limited.

3.3 Legally Protected Private Files

Repositories need to know when a collection offered contains legally protected private files, such as confidential government files; medical records; legal case files; or other kinds of sensitive information, such as Social Security and credit card numbers, whenever possible. Some of these files may need to be removed in their entirety; others require targeted search and redaction. Asking donors or dealers whether they know of such materials in the digital files being offered is an important step, but repository staff also must be vigilant in identifying and handling legally protected files, using both screening software and human analysis.

3.4 Hidden Content: Disk Images, Password Protection, Firewalls

At the most comprehensive level, a repository may wish to create a disk image (a file or files that contain an exact copy of the contents and structure) of a hard drive or other piece of storage media. The potential benefits of capturing an image of a whole system disk include the possibility of future interactive access, which could increase the value of the materials. Digital files and disk images, in particular, can also provide a unique opportunity to recover or discover content previously hidden or thought lost. In many cases, repositories may be able to recover deleted files or automatically saved materials (such as "Auto Save" word processor drafts). Agreements should specify whether disk images are to be included in a

collection of digital files and, if so, whether the disk image may need to be amended to address privacy concerns. Further, donors may wish to transfer disk images to a repository with the understanding that although individual files may be extracted from them and made public, the disk images themselves will serve as a master copy that remains untouched and is restricted to staff use for preservation purposes.

In some cases, especially if digital files come to a repository through an estate, materials may be hidden behind an unknown login or password. Content may also reside behind subscriber-only paywalls or in a private intranet (especially for organizational records). Any acquisition agreement should designate whether a donor allows repository staff to decrypt passwords and logins, or to recover deleted files, and whether the donor or dealer grants ownership of files recovered by these methods. Decryption and file recovery may be the only way to gain access to some digital content. All parties should be aware that repositories may discover materials not intended to be included in a collection (as indeed has long happened with paper archives), such as files created by a user other than the donor or items that fall outside the scope of the acquisition.

3.5 Recommendations for Donors and Dealers

- Consider screening e-mail files for sensitive or extraneous messages. When this is not possible, consider appropriate embargo periods, and discuss restrictions on access to e-mail before that date with the repository.
- Inform repository staff if there is a possibility that the digital records include the intellectual property of people besides the creator or donor of the materials.
- Inform repository staff if there is a possibility that the digital records include legally protected private files, such as confidential government files, medical records, and legal case files, or other kinds of sensitive information, such as Social Security and credit card numbers.
- Consider whether a repository should be allowed to decrypt passwords and logins.
- Discuss with repository staff the variety of deleted information that may be present in digital files and media, and come to an agreement about how such information will be handled and made available to researchers (e.g., perhaps after an embargo period).

3.6 Recommendations for Repositories

- Make clear who will screen e-mail messages for sensitive information and what the process will be.
- Carefully review e-mail files to identify and quarantine entire sensitive threads of sent and received messages and their attachments. Ensure that the preservation of and access to these messages are handled in accordance with long-term research value, all applicable laws, and provisions in the acquisition agreement.

- When staff, time, or technology constraints do not allow for review of sensitive e-mail messages or other files, discuss restrictions on researcher access for a defined period of time with the donor or dealer.
- Before acquisition, ask if the collection includes legally protected files, such as confidential government files, medical records, and legal case files, or especially sensitive types of information, such as Social Security and credit card numbers.
- Anticipate the presence of other people's intellectual property in a donor's files and establish relevant policies.
- Balance the desire to capture information about the donor's working environment and organizational strategies (via authentication and retention of original file structure or other means) with respect for the donor's wishes regarding privacy and restrictions.
- Consider potential future use and access mechanisms when contemplating the value of disk images, deleted files, and automatically saved files.
- Make sure donors and dealers are aware of the different kinds of deleted information that may be present in their digital materials.
- Be realistic about restrictions, redaction, and the potential for sensitive material to be missed in an initial review, considering
  - legal restrictions
  - donor requests for restrictions
  - third-party restrictions (e.g., oral histories without permission forms)
  - technical constraints (e.g., obsolete formats, corrupted files, other access problems)
  - use restrictions (e.g., need to ensure authenticity and appropriate use by patrons)

4. Key Stages in Acquiring Digital Materials

Once all parties have decided to proceed with the acquisition of a collection that includes born-digital materials, an agreement or contract should be established between the repository and the donor or dealer that defines the materials to be transferred and specifies other details related to the acquisition. Because digital media may contain files or fragments not immediately apparent to the donor and because repositories must invest substantial resources to capture, maintain, and make accessible born-digital materials, both parties may consider limiting the scope and types of files included in a transfer. Likewise, some archival repositories may want to take a more comprehensive approach, retaining digital media not only for their potential research and iconic value but also in anticipation of the recovery of more information as technology advances in the future. Furthermore, one especially noteworthy difference between physical and digital property lies in the ease with which exact copies of a digital file may be created and disseminated. A donor may wish to retain a copy of the digital files for his or her own use. A repository may require that it be the sole authorized entity for the ownership,

preservation, and management of a body of digital files. Donors and repositories should negotiate and clarify the specific terms of ownership of all digital files included in the acquisition.

4.1 Acquisition Agreement or Contract

Formal, written agreements or contracts transfer title of the materials to the repository and address a number of important matters related to the acquisition. Key factors to clarify in a written agreement or contract related to the transfer of born-digital materials include

- what digital materials will be transferred by a donor to a repository (e.g., hard drives, disks, e-mail archives, websites)
- what limitations, if any, will be set regarding the type and quantity of materials to be acquired
- whether all or only specifically identified files (e.g., word-processing files, e-mail messages) will be captured, preserved, and made accessible
- what will be done with files that are transferred but do not fall within the scope of the acquisition agreement
- whether digital media should be returned to the donor or kept by the repository after files have been captured
- whether a donor or dealer will be allowed to retain a copy of the digital files for reference use or provide another copy of the digital files to a different repository
- how sensitive materials not protected by law (e.g., Social Security numbers, passwords, financial information) will be handled
- whether any files or information should be captured but redacted or restricted from access and when those restrictions will expire
- how the parties will change or modify the terms of the agreement, if necessary

4.2 Communication and Interaction

Direct communication between a repository's staff and the donor is invaluable in ensuring the preservation of born-digital materials and access to them. Often the donor is the only person able to answer questions about the hardware or software used to create particular files, the specific materials that he or she intends to transfer, and issues related to privacy and sensitive information. The exact nature of these conversations may vary according to the technical knowledge of the parties involved, but it is important to clarify and discuss issues such as

- how sensitive materials will be reviewed and screened
- what the process will be for restriction or redaction of private content
- how materials will be made accessible to researchers (including whether or when they will be available online)
- how digital materials will be stored and preserved

The donor is also encouraged to provide repository staff with a personal computing history that details any hardware and software used and the methods of creating, storing, and maintaining digital files. The more detailed information the donor can provide, the more success the repository will have in preserving born-digital materials. It may also be useful for the donor to involve a technical specialist in conversations with the repository, particularly if the donor is an organization or other entity that relies on technical staff to manage its files.

If a donor is unavailable to discuss collection materials, a repository may consider establishing communication with others who can provide information about the born-digital materials in an acquisition; the donor's estate, family, or associates are potential sources of assistance. Sometimes, particularly in the case of media and files created long ago by a donor who is now unreachable, it is not possible for repositories to learn more about the digital materials in a collection.

If the repository expects to acquire additional materials from a donor in the future, early communication can have the residual effect of better educating the donor about preserving subsequent born-digital files. While discussing the initial transfer of digital media or files, repositories may establish and discuss protocols for future or ongoing digital acquisitions. Similarly, repositories should consider involving suitable technical specialists in the initial stages of acquisition, including for site visits and before the finalization of acquisition agreements.

4.3 Transfer of Materials

Once acquisition terms have been established, the repository and the donor should determine how the born-digital materials will be transferred. Transfer strategies differ for files stored on removable media, files that a donor wishes to transfer electronically, or files that an archivist may capture on a site visit. The best method for transferring digital materials from a donor to a repository depends on the specific materials in question. Each repository may have its own preferred methods of transfer, but any strategy involving the copying and electronic transfer of data can be time-consuming for both donor and repository. Donors should also be aware that some acquisitions present novel scenarios that may require repositories to research, develop, and test new methodologies to capture the collection materials (see Appendix C for recommendations about how to prepare for the unexpected).

It is strongly recommended that donors and dealers seek the guidance of archival repositories before any transfer takes place. Often, repository staff prefer to retrieve the media and files in person. A donor who is to send disks, computer hard drives, or other hardware directly to a repository should do so in accordance with the repository's guidelines for safely and securely handling, packing, and shipping digital media. It is crucial that media are well protected
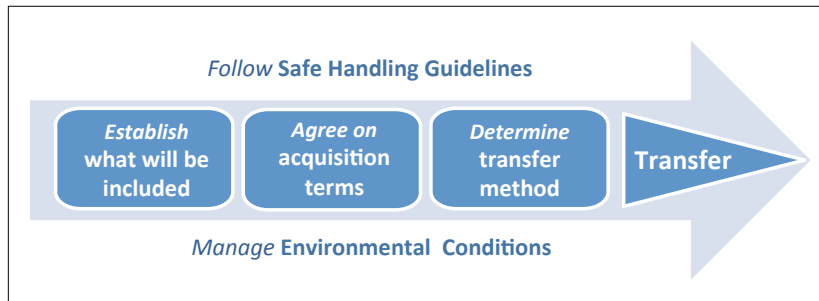
Fig.1: Components of the transfer process

from moisture, extreme temperatures, strong electromagnetic fields, and rough handling when being transferred to a repository. Any equipment should be cushioned during transport, and moving parts should be stabilized (e.g., inserting a dummy floppy disk into a floppy disk drive may be advisable). Alternatively, files may be transferred electronically by secure means. All of these scenarios will benefit from established repository protocols, easy-to-follow instructions, and tested documentation strategies to ensure that the correct files are acquired in a way that is well authenticated.

Throughout this process, it is important for donors and dealers to document who has handled the digital media and what actions have been taken prior to the files' transfer to the repository. For example, donors or dealers should indicate whether the files were copied from a computer hard drive onto a disk, noting the computer's make, model, and operating system, if known. It is also important to consider early on what security measures should be in place regarding the storage, handling, and secure deletion of electronically transferred files.

4.4 Initial Handling of Materials at the Repository

Digital media may pass through the hands of numerous staff members during the acquisition process. Repositories should minimize the number of people who handle digital media and files, and staff should carefully document each transfer of digital materials. After born-digital materials arrive on site, the repository should ensure that they are transferred to the appropriate department safely, quickly, and in their original condition. For example, items that are shipped physically, particularly if they are included in a larger shipment of paper-based materials, may require an inspection by conservators before they are integrated into the repository's collection storage areas. Curators or other staff may need to examine the materials in preparation for announcements about the acquisition, exhibitions, or other purposes. Collections that have been in storage for an extended period may contain digital media that are not discovered until archival processing is under way.

Because attempts to access born-digital materials can change the content, formatting, and metadata associated with the files, repositories must establish clear protocols for the staff's handling of these materials. Such protocols should also include strategies to document

which repository staff members have handled the media as the acquisition moves through different departments and what actions they have taken.

4.5 Recommendations for Donors and Dealers

- Discuss whether the repository will be the exclusive owner of the digital files transferred, whether copies of the digital files will be allowed, and if so, how they are to be made.
- Clearly identify which born-digital materials are to be offered to the repository.
- If offering digital media to a repository, determine whether all files on the media can be captured and made accessible or whether only certain types of files are accessible.
- Be prepared for the repository to retain the original digital media unless the acquisition agreement indicates otherwise.
- Determine whether limited, specific files or information need to be restricted or redacted and when those restrictions will expire.
- Consider seeking advice from a legal professional when working with a repository on a contract or agreement.
- Consider involving a technical specialist in conversations with repository staff about what guidelines to follow when copying and transferring media and files.
- Be prepared to have ongoing communications with the repository during the capture and processing of digital materials.
- Determine whether computer equipment no longer used by the donor can be offered as part of the acquisition.
- Consider writing a personal computing history to provide context for the digital media and computers in the acquisition.
- Seek guidelines from the repository for the shipping of hardware or transfer of digital files.
- Document the ways in which digital media and files have been stored, accessed, and transported prior to their arrival at, or collection by, the repository.
- Be prepared for electronic copying and transfer to take a significant amount of time and for repository staff to develop new capture techniques to accommodate novel acquisition scenarios.

4.6 Recommendations for Repositories

- Clarify whether the repository will be the exclusive, long-term owner of a unique or master set of digital files that will not be made available to other repositories or purchasers.
- Clearly identify which born-digital materials are to be included in the acquisition.
- Consider limiting the scope of the files to be acquired to ensure that the materials transferred are of research value to the repository, but do not overlook the potential benefits of a comprehensive acquisition, including entire disks.

- Consider what will be done with files that are transferred but do not fit the scope of the acquisition agreement.
- Document the details of the acquisition with a written agreement or contract.
- Determine how digital materials will be transferred to the repository.
- If possible, establish direct communication with the donor or estate.
- Establish protocols for how digital materials should be handled and documented from their arrival at the repository until they reach the digital archivist (or other appropriate person).
- Be prepared to research, develop, and test new capture methods for novel acquisition scenarios.

5. Post-Acquisition Review by the Repository

Born-digital materials may require multiple levels of assessment. Ideally, the initial assessment occurs prior to acquisition, when the donor, dealer, and repository staff work together to determine what materials the repository will collect. Staff need to undertake another assessment after the materials have arrived at the repository. The post-acquisition review process can provide a repository with an opportunity to supply the donor with information about the digital materials, such as an inventory of disk titles or a list of directories and files received.

5.1 Physical Condition

In order for the repository staff to capture and preserve their contents, digital media must be reasonably clean and physically intact. The staff may provide donors and dealers with storing and packing recommendations to help protect digital media from harm prior to or during transit. In some cases, computer media will have sustained damage long ago. Examples of damage to computers, disks, and tapes include a bent computer chassis or disk drive, a cracked cartridge case, an exposed internal magnetic disk, a scratched optical disk, and a floppy disk that is covered in dust. Pre-transfer documentation about physical condition, as well as information about who has handled the media and where they have been stored, may help pinpoint when damage occurred. Dated digital photographs of computer equipment and media may be very helpful in this context. Donors and dealers are advised not to attempt any restoration or repair without first consulting an archival repository.

Physical damage to digital media may prevent access to important content and compromise the item's value as a material artifact. In addition, inserting bent, dirty, or broken media into a functional disk drive in an attempt to access the contents could irreparably harm a repository's processing workstation. When digital media arrive at the repository in a condition different from that expected, repository staff must decide whether to accept the materials or revisit

Fig.2: Digital media with physical damage: a CD, a USB drive, a 3.5" floppy disk.
photos: © shutterstock.com

the repository's agreement with the donor or dealer. Physical damage to media may also change the level of preservation and access that a repository can provide. Documenting the physical condition of the files in relation to other factors and making that information easily accessible will facilitate preservation planning. Information about physical condition may also be of interest to future researchers. Donors, dealers, and repositories should not assume that nothing of value can be recovered from digital media just because they may be decades old and battered looking.

Repository staff should also pay close attention to the physical labels on digital media. Although labels are not always accurate, sometimes they offer richly detailed and useful information that may help repository staff conduct an assessment of the media and verify inventory. Even when information on a label seems unrelated to the digital content, it may provide insight into the provenance of a piece of media.

5.2 Digital Condition

It may be preferable for repository staff to be the first to preview and capture digital media, using write-protection technologies that protect the original media and files from inadvertent change. When copying or imaging media, it is good practice to create hash values (sometimes known as checksums), which serve as unique "digital fingerprints" for each and every digital object or file. If a single bit is changed in a file, the altered file will produce a different hash value; conversely, if in 100 years a file is subjected to the same test and it yields the same hash value that it did when it was first received, a repository can be confident that the file has remained unchanged over the century.

In some cases, donors and dealers may want to generate hash values themselves, perhaps with the assistance of a technical specialist. If so, the hash information can help repository staff verify that the files arriving at the repository match those transferred by the donor. Mismatched hash values would indicate that the incoming materials had been altered or corrupted in transit. Staff would need to decide whether to accept these damaged files, decline them, or work with the donor to determine the source of the problem and, if possible, fix it. Comparing and confirming hash values generated prior to transfer with those generated once the materials have arrived at the repository requires a certain level of technical comfort and expertise on the part of both the donor and the receiving repository.

If internal policies and donor agreements allow, repository staff may create a disk image, or an exact copy of the entire contents, of a piece of media or working copies of individual files and use available tools to preview the contents shortly after the materials arrive at the repository. Sometimes an archival repository may offer or prefer to undertake a preview or even an actual capture of files at the site of the donor or dealer. If the repository staff determine that the content of the transferred media or set of files does not match that of the original inventory or does not fit the repository's collection development policy, further discussion with the donor or dealer may be in order.

5.3 Retention and Disposal

Depending on a repository's internal policies and agreements with donors, disk images, system files, inaccessible files, unlawful content, damaged or blank media, and content that falls outside a repository's collecting policy may be returned to the donor. Any action should be well documented and supported by the acquisition agreement and a repository's policies.

For example, repositories may consider developing a retention policy for damaged media whose contents are ultimately inaccessible. An acquisition agreement may state how and under what circumstances materials may be culled from a collection, and a policy may outline how these materials will be disposed of securely. Some repositories may be reluctant to delete the last known copy of a file and may insist that the donor take responsibility for its destruction.

A blank disk or one containing inaccessible files may still hold cultural value as a physical object or even as a teaching tool. Even if some files are currently unavailable because of an esoteric file format or disk format, if the bits represented by the file or disk have been captured, future technologies (or a dedicated research project) may make it possible to interpret these bits in a meaningful way. Thus, a repository may come to regret not retaining apparently inaccessible material. It is also important to note that in a digital context, neglect or a decision not to devote resources to active preservation is often tantamount to gradual, but inexorable, destruction.

5.4 Recommendations for Donors and Dealers

- Include information about physical damage to digital media in the initial collection inventory or survey. Noting damaged digital media is as important as disclosing damage to paper materials. Digital photographs of media and equipment may be helpful.
- Remain open to ongoing communication with repository staff about the acquisition.
- Take great care when packing and shipping physical media and transferring digital files. Media damaged in shipping or files corrupted in transfer lose significant cultural value.

5.5 Recommendations for Repositories

- Determine whether digital media have been damaged in transit.
- Determine whether files may have been affected by transit or the transfer process.
- Use hash values or checksums (unique digital fingerprints) and preview tools to check the digital condition and authenticity of materials.
- Consider the physical condition of digital media before trying to access them.
- Determine whether inaccessible media retain value as physical artifacts.
- Develop policies regarding the retention and disposal of certain types of files and digital media. Make sure decisions are supported by policy.

6. Conclusions

The stewardship of born-digital archival collections promises nothing if not routine encounters with the unexpected. Unfamiliar or unannounced file formats, hardware, and additions to collections seem to be reliable constants as repositories increasingly collect born-digital archival content.

This report describes good practices that can help reduce archival surprises. Conducting thorough and clear surveys, interviews, and other types of assessment prior to acquisition can help reduce the occurrence of unexpected large additions to collections, unfamiliar media and hardware, and unanticipated expansions to the scope of an acquisition. Documented and well-formed acquisition policy and practices may alleviate ambiguity about the details of transferring born-digital materials, such as timing, packing and shipping standards, frequency of accruals, and "rogue data" transferred unintentionally or not fully addressed in the acquisition agreement or contract. Furthermore, to enhance the quality of digital transfers and accessions, dealers and repositories should work to improve understanding and documentation of donor computing habits. Early archival intervention in records and information management will help shape the impact on archives of user and donor idiosyncrasies around file management and data backup.

The unexpected will continue to challenge and surprise repositories acquiring and managing born-digital materials, despite reasonable efforts at creating clear and actionable policies. Tactics such as opening the lines of communication among donors, dealers, and repository staff, and establishing transparent, efficient archival practices promise to minimize unpleasant surprises and improve the quality of born-digital acquisitions.

APPENDIX A

Potential Staffing Activities for the Repository

Activities undertaken by staff represent the largest cost element to a repository acquiring born-digital archives. There can be many steps from initial contact to first capture, and the intensity of activity in each may vary from collection to collection.

WHAT?	HOW?	WHO?
Initial contact	Remote	Subject specialist learns of collection, either through own research, contact initiated by donor, or via third party (e.g., dealer).
Initial survey (content)	Onsite	Subject specialist evaluates potential research value of the material offered.
Initial survey (technical)	Possibly a mixture of remote information gathering and onsite analysis	Process and technical specialists establish technical characteristics to inform capture and preservation techniques.
Research new capture scenarios (e.g., special scoping requirement of donor; new data format)	Remote	Process and technical specialists research and determine capture techniques for new scenarios. Input from donor's technical support and repository technical staff may be required.
Interaction with donor to define scope and explain processes	Onsite/remote	Subject specialist likely leads on content-driven matters. Process staff may need to explain possibilities (e.g., recovery of deleted material).
Drafting of terms of agreement	Remote	Most repositories begin discussion with a template terms of agreement. Subject specialists and process specialists may make collection-specific provisions if required. Repository's legal team supports this work.
Actual capture of material (especially for onsite capture)	Onsite/remote	Process specialist executes capture. This process can be time-consuming, depending on the scale of material involved.
Creation of accession record	Repository	Curatorial and process specialists create an accession record at the repository.
Ingest of material to digital repository	Repository	Process specialist manages ingest of new material to the repository's digital preservation system.
Ongoing support/dialogue with donor (for ongoing arrangements)	Onsite/remote	Curatorial and process specialists remain available for future consultation with regular donors as new scenarios dictate re-working of capture procedures.

APPENDIX B

List of Resources and Related Projects

The following is a brief list of projects and publications that include samples of policies, surveys, tool evaluations, and other documentation related to born-digital collections.

AIMS Work Group. 2012. *AIMS Born-Digital Collections: An Inter-Institutional Model for Stewardship*. Available at http://www.digitalcurationservices.org/aims/white-paper/.

The AIMS project, funded by The Andrew W. Mellon Foundation from 2009 to 2011, was a collaboration between the libraries of the University of Virginia, Stanford University, Yale University, and the University of Hull. The resulting white paper includes

- case studies
- sample policies
- sample survey template
- tool evaluations
- guidelines for creating agreements

Charlesworth, Andrew. 2012. *DPC Technology Watch Report: Intellectual Property Rights for Digital Preservation*. Available at http://www.dpconline.org/advice/technology-watch-reports. (DOI http://dx.doi.org/10.7207/twr12-02).

This Technology Watch report from the Digital Preservation Coalition (DPC) explores United Kingdom and European Union copyright laws in the context of digital preservation and considers strategies for managing intellectual property rights and any related legal risks that may arise in the course of digital preservation activities. It includes

- an appendix of legislation in the United Kingdom and the European Union
- a list of references and resources

Dooley, Jackie, and Ricky Erway. 2013. *Demystifying Born Digital*. Available at http://www.oclc.org/research/activities/borndigital.html.

The goal of this ongoing OCLC research project is to produce reports that will help research libraries manage born-digital materials more effectively. The target audience ranges from people new to working with digital media to those who have some experience and need additional guidance. The first three reports are available at http://www.oclc.org/research/publications/library/2012/2012-06r.html:

- Erway, Ricky. 2012. *You've Got to Walk Before You Can Run: First Steps for Managing Born-Digital Content Received on Physical Media*. Dublin, Ohio: OCLC Research.
- Erway, Ricky. 2012. *Swatting the Long Tail of Digital Media: A Call for Collaboration*. Dublin, Ohio: OCLC Research.
- Barrera-Gomez, Julianna, and Ricky Erway. 2013. *Walk this Way: Detailed Steps for Transferring Born-Digital Content from Media You Can Read In-house*. Dublin, Ohio: OCLC Research.

Hill, Catey. 12 April 2013. 5 Steps to Creating Your Digital Estate Plan. *Next Avenue.* Available at http://www.nextavenue.org/article/2012-05/5-steps-creating-your-digital-estate-plan.

This article provides recommendations to help individuals evaluate and manage their digital assets as part of estate planning. Resources include

- sample inventory of digital assets
- sample questions to consider when writing instructions for an executor
- information about what popular social media services require before providing an executor with access to an account

John, Jeremy Leighton. 2012. *DPC Technology Watch Report: Digital Forensics and Preservation*. Available at http://www.dpconline.org/advice/technology-watch-reports (DOI http://dx.doi.org/10.7207/twr12-03).

This Technology Watch report from the Digital Preservation Coalition (DPC) explores the application of digital forensics in the preservation of cultural heritage materials with special attention to personal archives. It includes a

- detailed overview of tools and techniques
- bibliography of references and additional resources
- chapter devoted to legal and ethical concerns

Kirschenbaum, Matthew, Richard Ovenden, and Gabriela Redwine. 2010. *Digital Forensics and Born-Digital Content in Cultural Heritage Collections*. Available at http://www.clir.org/pubs/reports/pub149/pub149.pdf.

This report, funded by The Andrew W. Mellon Foundation and published by CLIR, explores the key challenges facing professionals charged with collecting and maintaining born-digital cultural heritage materials and suggests ways that tools and methodologies from the computer forensics field may help. The report includes

- tool comparisons
- case studies
- resource list

Library of Congress. 2013. Personal Archiving: Preserving Your Digital Memories. *Digital Preservation.* Available at http://www.digitalpreservation.gov/personalarchiving/.

This web page offers individuals general advice about how to preserve their own born-digital files, including photographs, audio and video recordings, e-mail, text, and websites.

Paradigm Project. 2005–2007. *Workbook on Digital Private Papers*. Available at http://www.paradigm.ac.uk/workbook/.

Funded by the Joint Information Systems Committee (JISC) in the United Kingdom (UK), the Paradigm Project was a collaboration between the John Rylands University Library at the University of Manchester and the Bodleian Library at the University of Oxford. The resulting workbook includes a wealth of information about how to manage digital personal papers, including

- best practice guidelines for donors and repository staff
- a model gift agreement
- tool evaluations
- a sample survey template
- instructions for creating screenshots and capturing directory structures
- a chapter devoted to legal issues related to digital archives (UK-specific)
- lists of publications and other resources

Prom, Christopher J. 2011. *DPC Technology Watch Report: Preserving Email*. Available at http://www.dpconline.org/advice/technology-watch-reports (DOI http://dx.doi.org/10.7207/twr11-01).

This Technology Watch report from the Digital Preservation Coalition (DPC) reviews efforts to capture and preserve e-mail messages for the long term. It offers recommendations for professionals tasked with digital preservation, as well as guidance for individuals interested in managing their own e-mail archives. The report includes

- case studies
- tool assessments
- recommendations for information professionals and individuals
- references and additional resources

Walker, Rob. 2011. Cyberspace When You're Dead. *New York Times Magazine.* Available at http://www.nytimes.com/2011/01/09/magazine/09Immortality-t.html?pagewanted=all&_r=0.

This article presents several examples of how families and friends have handled an individual's digital personal effects after he or she died. Walker provides an overview of the key concerns related to digital estate planning for oneself and the task of managing someone else's digital legacy after death. Information includes

- case studies
- examples of digital estate management services and tools

APPENDIX C

Preparing for the Unexpected: Recommendations

Donors, dealers, and repositories can anticipate and prepare for the complexities of transferring, managing, and preserving born-digital and hybrid collections. The following table organizes select recommendations from the report across the categories of policy, collaboration, preparation, and documentation.

	DONORS AND DEALERS	REPOSITORIES
POLICY	• Know whether your digital files may include the intellectual property of people other than the creator or specific donor of the materials. • Know the circumstances under which the material was created. Was it part of your work? Your research? Your personal correspondence?	• Develop policies for the retention and disposal of files and digital media. • Balance the desire to gather information about the donor's working environment with the need to respect donor privacy and wishes. • Develop strategies to manage restriction, redaction, and discovery of sensitive material that may have been overlooked in the initial review.
COLLABORATION	• Ask repositories for guidance and documentation. • If possible, establish direct communication with the repository staff to facilitate understanding. • Discuss the overall process, timetable, and likely outcomes with repository staff. • Inform the repository of the presence of legally protected private files, such as medical records, and especially sensitive types of information, such as Social Security numbers. • Involve technical or legal support people in the process, if needed or suggested by the repository staff.	• Prior to acquisition, gather information about archives containing born-digital content. • Share documentation with donors and dealers. • Clarify expectations regarding digital preservation and access.
PREPARATION	• Consider the amount of time available to understand, and contribute fully, to the process of transferring the digital material. • Decide which content you want to transfer to the repository. • Do not manipulate, extract, or copy data from their original source before offering the material for gift or purchase, or do so in consultation with repository staff. • If you wish, screen e-mail files for sensitive or extraneous messages; do so in consultation with repository staff. • Work closely with the repository to ensure that digital media are packed and shipped in a way that guarantees their safe arrival. • Be prepared for the copying and transfer of your files to take time.	• Consider their potential future use when assessing the value of disk images, deleted files, and automatically saved files. • Assess the collection's location and complexity, need for ongoing communication, and possibility of remote acquisition. • Conduct a digital records survey. • Determine how digital materials will be transferred to the repository. • Establish a process for screening e-mail correspondence for sensitive messages. • Consider embargo as an alternative when staff, time, or technology constraints do not permit intensive screening.
DOCUMENTATION	• Prior to transfer, determine which files need to be restricted or redacted and the terms. Give this information to the repository. • Flag areas of concern for repository staff. • Carefully review all agreements.	• Address born-digital materials in the acquisition agreement or contract. • Document whether a donor can provide another copy of the digital files to a different repository. • Develop and distribute guidelines for the transfer of digital files. • Clearly define which materials are to be included in the acquisition. • Document how to handle sensitive information included in the collection. • Document the handling of passwords, encryption, and deleted content. • Document the type of access that users may have to processed collections.

APPENDIX D

Checklist of Recommendations for Donors and Dealers

Initial Collection Review

❏ Be aware that simply viewing files can alter them.

❏ Avoid manipulating, rearranging, extracting, copying, or otherwise altering data residing in the original source media in anticipation of offering the materials for gift or purchase, or preview files in accordance with established guidelines provided by the repository.

❏ Ask repositories for guidance and documentation on
  • determining the most appropriate repository for the digital materials
  • negotiating the terms of an acquisition agreement or contract as it relates to born-digital materials
  • describing the context and history of the files and media being transferred
  • handling digital media
  • documenting storage, access attempts, copies, and transport of digital media and files

❏ Clarify expectations about the extent to which the digital materials on offer will be preserved and made available for use.

Privacy and Intellectual Property

❏ Consider screening e-mail files for sensitive or extraneous messages. When this is not possible, consider appropriate embargo periods and discuss restrictions on access to e-mail before that date with the repository.

❏ Inform repository staff if there is a possibility that the digital records include the intellectual property of people besides the creator or donor of the materials.

❏ Inform repository staff if there is a possibility that the digital records include legally protected private files, such as confidential government files, medical records, and legal case files, or other kinds of sensitive information, such as Social Security and credit card numbers.

❏ Consider whether a repository should be allowed to decrypt passwords and logins.

❏ Discuss with repository staff the variety of deleted information that may be present in digital files and media, and come to an agreement about how such information will be handled and made available to researchers (e.g., perhaps after an embargo period).

Key Stages in Acquiring Digital Materials

❏ Discuss whether the repository will be the exclusive owner of the digital files transferred, whether copies of the digital files will be allowed, and if so, how they are to be made.

❏ Clearly identify which born-digital materials are to be offered to the repository.

❑ If offering digital media to a repository, determine whether all files on the media can be captured and made accessible, or whether only certain types of files are accessible.

❑ Be prepared for the repository to retain the original digital media unless the acquisition agreement indicates otherwise.

❑ Determine whether limited, specific files or information need to be restricted or redacted and when those restrictions will expire.

❑ Consider seeking advice from a legal professional when working with the repository on a contract or agreement.

❑ Consider involving a technical specialist in conversations with repository staff about what guidelines to follow when copying and transferring media and files.

❑ Be prepared to have ongoing communications with the repository during the capture and processing of digital materials.

❑ Determine whether computer equipment no longer used by the donor can be offered as part of the acquisition.

❑ Consider writing a personal computing history to provide context for the digital media and computers in the acquisition.

❑ Seek guidelines from the repository for shipping of hardware or transfer of digital files.

❑ Document the ways in which digital media and files have been stored, accessed, and transported prior to their arrival at, or collection by, the repository.

❑ Be prepared for electronic copying and transfer to take a significant amount of time, and for repository staff to develop new capture techniques to accommodate novel acquisition scenarios.

Post-Acquisition Review by the Repository

❑ Include information about physical damage to digital media in the initial collection inventory or survey. Noting damaged digital media is as important as disclosing damage to paper materials. Digital photographs of media and equipment may be helpful.

❑ Remain open to ongoing communication with repository staff about the acquisition.

❑ Take great care when packing and shipping physical media and transferring digital files. Media damaged in shipping or files corrupted in transfer lose significant cultural value.

APPENDIX E

Checklist of Recommendations for Repositories

Initial Collection Review

☐ Assess born-digital materials prior to acquisition.

☐ Weigh the cultural and research value of the collection, or components of it, against the cost of capture, ongoing preservation, and access.

☐ Share relevant information and documentation with donors and dealers about the collection review and acquisition processes, as well as sample use scenarios.

☐ Clarify expectations regarding digital preservation and access.

☐ Conduct a survey of born-digital materials:
  • Determine the appropriate survey method by consulting with the donor or dealer.
  • Conduct onsite surveys (in person) when possible.
  • Have policies in place regarding the capture, storage, and disposal of files copied for the purpose of preliminary assessment.
  • Safely capture the donor's files or directory structures when closer analysis or more time is needed than is possible on a site visit or when an onsite survey is not possible.

Privacy and Intellectual Property

☐ Make clear who will screen e-mail messages for sensitive information and what the process will be.

☐ Carefully review e-mail files to identify and quarantine entire sensitive threads of sent and received messages and their attachments. Ensure that preservation and access to these messages are handled in accordance with long-term research value, all applicable laws, and provisions in the acquisition agreement.

☐ When staff, time, or technology constraints do not allow for review of sensitive e-mail messages or other files, discuss restrictions on researcher access for a defined period of time with the donor or dealer.

☐ Before acquisition, ask if the collection includes legally protected files, such as confidential government files, medical records, and legal case files, or especially sensitive types of information, such as Social Security and credit card numbers.

☐ Anticipate the presence of other people's intellectual property in a donor's files and establish relevant policies.

☐ Balance the desire to capture information about the donor's working environment and organizational strategies (via authentication and retention of original file structure or other means) with respect for the donor's wishes regarding privacy and restrictions.

❑ Consider potential future use and access mechanisms when contemplating the value of disk images, deleted files, and automatically saved files.

❑ Make sure donors and dealers are aware of the different kinds of deleted information that may be present in their digital materials.

❑ Be realistic about restrictions, redaction, and the potential for sensitive material to be missed in an initial review, considering
  - legal restrictions
  - donor requests for restrictions
  - third-party restrictions (e.g., oral histories without permission forms)
  - technical constraints (e.g., obsolete formats, corrupted files, other access problems)
  - use restrictions (e.g., need to ensure authenticity and appropriate use by patrons)

Key Stages in Acquiring Digital Materials

❑ Clarify whether the repository will be the exclusive, long-term owner of a unique or master set of digital files that will not be made available to other repositories or purchasers.

❑ Clearly identify which born-digital materials are to be included in the acquisition.

❑ Consider limiting the scope of the files to be acquired to ensure that the materials transferred are of research value to the repository, but do not overlook the potential benefits of a comprehensive acquisition, including entire disks.

❑ Consider what will be done with files that are transferred but do not fit the scope of the acquisition agreement.

❑ Document the details of the acquisition with a written agreement or contract.

❑ Determine how digital materials will be transferred to the repository.

❑ If possible, establish direct communication with the donor or estate.

❑ Establish protocols for how digital materials should be handled and documented from their arrival at the repository until they reach the digital archivist (or other appropriate person).

❑ Be prepared to research, develop, and test new capture methods for novel acquisition scenarios.

Post-Acquisition Review by the Repository

❑ Determine whether digital media have been damaged in transit.

❑ Determine whether files may have been affected by transit or the transfer process.

❑ Use hash values or checksums (unique digital fingerprints) and preview tools to check the digital condition and authenticity of materials.

❑ Consider the physical condition of digital media before trying to access them.

❑ Determine whether inaccessible media retain value as physical artifacts.

❑ Develop policies regarding the retention and disposal of certain types of files and digital media. Make sure decisions are supported by policy.