---------- Forwarded message ----------
From: Gabriela Redwine <gredwine@austin.utexas.edu
To: Peter Wosh <pw1@nyu.edu, Teresa Brinati <tbrinati@archivists.org
Date: Fri, Dec 14, 2012 at 3:26 PM
Subject: Born Digital: Guidance for Donors, Dealers, and Archival Repositories

Dear Peter and Teresa,

I'm writing to pick up our correspondence (see below) about the born-digital acquisitions report, which now has an official title: Born Digital: Guidance for Donors, Dealers, and Archival Repositories.

I am in the process of finalizing our content on the MediaCommons Press website and preparing marketing materials. The report should go live (fingers crossed) at the end of December or beginning of January.

I attach a final draft of our manuscript. Please consider this a formal submission to SAA's Publications Board. The content is the same as what will appear on MediaCommons, although some of the formatting is slightly different.

Our initial proposal mentioned "satellite documents," or appendices, that would provide more specific guidance for certain scenarios (e.g., acquisition done via FTP). After some discussion, we decided to postpone writing these. MediaCommons is open to the idea of an evolving draft, so the tentative plan is to add a limited number of short appendices (approx. 1 page each) once we have a better sense of which topics readers would like to see addressed in more detail. If there doesn't seem to be any interest in additional appendices, then we won't include them in the MediaCommons version of the report.

In a previous email you suggested that SAA would not shy away from considering a report that had been "pre-published." I am in the process of exploring Creative Commons licensing options for the MediaCommons version of the report (http://creativecommons.org/licenses/). The Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) license seems like the most appropriate choice. Does the Publications Board have a policy about publishing a work previously released under a CC license?

Thanks for your help.

Sincerely,
Gabby

Gabriela Redwine, Archivist
Harry Ransom Center
The University of Texas at Austin

On 8/20/2012 11:32 AM, Peter Wosh wrote:

Sounds good, Gabby, and appreciate your keeping us in the loop!

Peter

On Mon, Aug 20, 2012 at 12:16 PM, Gabriela Redwine
<gredwine@austin.utexas.edu> wrote:

 Dear Peter,

 Thank you for your message. I asked our group's representative, Susan Thomas, to err on the side of caution in her talk for the Manuscripts Section and to mention publication possibilities only in very general terms (e.g., a hosted draft version, a more formally published e-version).

 Since then, I've heard back from MediaCommons Press. We will be working with them to publish a draft online and invite public review. I'll have a better sense of the time frame once I've talked more with them. (My goal had been 1 October, but that might need to change.)

 I am also very excited to hear about SAA's interest in the project. Thank you to both you and Teresa for your encouragement and for circulating the proposal to the Publications Board. I will keep in touch about our progress.

 Sincerely,
 Gabby

 On 8/13/2012 1:35 PM, Peter Wosh wrote:

 Dear Gabby:

Thanks for your e-mail. Our San Diego SAA meeting just ended on Sunday, so it is time to deal with the mountain of e-mail! I am happy to say that our Publications Board was very intrigued by the concept and idea of creating an online publication concerning "Born-Digital Acquisitions." We would, of course, need to review the actual product before committing to publication and format, but would encourage you to move ahead at this stage and submit the manuscript when you are ready to go. Do you have a time frame in mind at this point?

Concerning your other question, since it is preliminary at this point, and we have not seen the manuscript, probably not better to mention SAA as your publishing option (though reading your e-mail, I am guessing that your colleague probably did so at the SAA Manuscripts Section, and that no harm was done). Anyway, look forward to working with you on this, and best of luck moving things forward!

 Peter

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

Authors

Beinecke Library, Yale University	Michael Forstrom, Archivist Nancy Kuhl, Curator of Poetry, Yale Collection of American Literature
Bodleian Library, University of Oxford	Susan Thomas, Digital Archivist and Project Manager
British Library	Jeremy Leighton John, Curator of eMANUSCRIPTS
Harry Ransom Center, The University of Texas at Austin	Megan Barnard, Assistant Director for Acquisitions and Administration Gabriela Redwine, Digital Archivist
Manuscript, Archives, and Rare Book Library (MARBL), Emory University	Kate Donovan, University Archivist Erika Farr, Coordinator for Digital Archives
Rubenstein Library, Duke University	Will Hansen, Assistant Curator of Collections Seth Shaw, Electronic Records Archivist, University Archives

External reviewers
Erin O'Meara, Archivist, Gates Archive
Kari R. Smith, Digital Archivist, MIT Libraries

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.

Below are definitions for some of the terms used most frequently in the report. These and other definitions also 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/or makes available archival materials transferred from donors.

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 archival repositories.

Each of the following sections provides an overview of the key issues and concludes with two lists of recommendations: one for donors and dealers, and a second for repository staff.

- 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 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.

Appendices provide more specific information about how to prepare for the unexpected and possible staffing costs, as well as 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 co-authors' institutions. Rather, they offer broad, useful guidance for donors, dealers, and repository staff involved in the acquisition of born-digital materials.

In order to ensure that born-digital materials arrive at repositories in good condition and accompanied by appropriate documentation, it is vital to convince donors, dealers, repository staff, and others 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.

Initial Collection Review

This section outlines the considerations and approaches that inform the collection review undertaken by donors, dealers, and repositories prior to a formal acquisition agreement or contract.

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 will understand better the scope and nature of the born-digital content and repositories can take preliminary steps to assure the integrity and usability of digital materials over time.

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 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 individuals use technology have had a significant impact on the types of materials that end up in archives. Digital archival materials have created complications that further underscore the benefits of an initial collection review. It is important for all parties—donor, dealer, and repository staff—to discuss early on whether the donor has digital media and files that form a significant part of the materials being offered for gift or sale. These conversations might involve repository staff members who hold accessioning, acquisition, curatorial, legal, processing, and technical responsibilities. Donors might also want to consider involving technology 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 the donor transfers only the media and files she intends. Any preliminary collection inventory should include general information about digital media and files, as well as relevant information about physical condition, such as mold, water, or other types of physical damage to hardware. This information will help repository staff evaluate the desirability of born-digital content, predict the storage requirements, estimate the associated staff and equipment costs over time, and decide whether to make an acquisition. A repository's assessment criteria might include:

- General technical characteristics of the media or files
- Volume of digital materials (including size range of files)
- Nature of the relationship between born-digital and paper materials within a collection
- 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.

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, what processes staff members will undertake, and what the repository needs from the donor or dealer and vice versa. Examples of information repositories may share with donors include:

- Collection development policy for born-digital materials
- Policies and procedures related to the acquisition, transfer, copying, embargo/restriction, user access, long-term preservation, and secure disposal of digital media and files
- Guidance in how to handle digital media and files and document the process
- Guidelines for preparing acquisition agreements or contracts

One area of particular concern involves the condition of the digital files included in collections offered to repositories. In order to protect the integrity of digital files, or 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.

In some cases, a donor might want to preview the contents of old disks or files to search for private files that she does not want to transfer. When a donor or dealer decides to access a disk or look at files on a computer before transferring ownership, it is important to 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 establish the provenance of the born-digital materials. Donors, dealers, and repository staff should also take care to handle digital media (e.g., physical objects such as disks, flash drives, etc.) carefully, in a way that does not damage them or compromise their value. Even turning on a computer may risk altering files, and opening files in applications can change dates and times, possibly affecting the future value of the materials. Digital specialists at archival repositories may be able to provide useful advice to donors and dealers.

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 on-site viewing, conducted in person, and a remote preview conducted over the Internet, via phone, or by some other method.

On-site assessment has several advantages, particularly when the proposed acquisition comprises a considerable amount of material. On-site surveys usually take place wherever a donor works or collection material is stored and 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
- Copying the donor's files or directory structures for further evaluation upon return to the repository, where staff will have more time and possibly additional tools at their disposal

Before a repository considers copying a donor's files for the purpose of later assessment, staff will want to work with the donor to put in place an agreement that specifies 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 re-capture copies of the original materials.

In situations where a donor and collection materials are located some distance from a repository, conducting a survey remotely might make more sense than doing a site visit. Likewise, repositories in geographically isolated locations or with small travel budgets may need to evaluate materials remotely. Remote surveys can be conducted via email or other means of communication. One strategy is to send a formal survey tool (e.g., a list of initial questions developed by a repository) to a donor via email or postal mail, ask the donor to respond, and then continue the conversation through a further exchange of letters, emails, or telephone calls. A remote survey might also include a document created by the dealer or donor that lists number and type of digital media and provides a general characterization of the contents. Alternatively, a donor might provide a repository with access over the Internet, via FTP (File Transfer Protocol) or secure peer-to-peer communication providing remote access to the donor's desktop. This approach could allow staff to preview the files in order to generate a survey, discuss the logistics of a possible transfer, and decide whether to make the acquisition.

Communication
Surveys and other pre-acquisition strategies can be supplemented by conversation 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 technical specialists to manage digital files, it might be helpful to involve that person or team in the collection review process. Guidelines describing how to handle and describe digital items prior to offering the materials for gift or purchase can help all parties anticipate problems and formulate solutions well in advance of the actual transfer of materials. Similarly, communication between repository staff and donors about the technical nature and extent of the digital materials in a prospective acquisition, as well as issues of privacy and confidentiality, access and restrictions, and expectations regarding

capture and transfer, can forestall potential processing difficulties, ultimately enabling repositories to serve as better stewards of born-digital content.

Recommendations for Donors and Dealers

- 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, do so in accordance with established guidelines provided by the repository.
- Ask repositories for guidance and/or documentation on:
  - Determining the most appropriate repository for your 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.

Recommendations for Repositories

- Assess born-digital materials prior to acquisition.
- Weigh the cultural and research value of the collection, or components of it, with the cost of capture and ongoing preservation and access.
- Share relevant information and documentation with donors and dealers about the collection review and acquisition processes.
- Clarify expectations regarding digital preservation and access.
- Conduct a survey of born-digital materials:
  - Determine the method by which to do this through consultation with the donor or dealer.
  - Conduct in-person/on-site surveys when possible.
  - Have policies in place regarding the capture, storage, and disposal of files copied for the purpose of preliminary assessment.
  - 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 on-site survey is not possible.

Privacy and Intellectual Property

This section addresses concerns related to the protection of private information and intellectual property in born-digital materials. Issues may include copyright and other intellectual property considerations; managing sensitive content in large bodies of email correspondence; legally protected private files; and more technically complicated issues such as password decryption and disk imaging.

It is important to note the wide range of opinion and understanding about digital privacy amongst all groups involved with the acquisition, transfer, cataloging, and description of archival materials. Whenever possible, donors, dealers, and repositories should discuss and reach consensus on the types of born-digital materials in an acquisition and the strategies for capturing and providing access to those materials.

Copyright and intellectual property
In some ways, copyright and intellectual property in digital files is relatively straightforward: donors and/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 use the materials on-site. Furthermore, acquisitions that cross national boundaries can be challenging due to different copyright and intellectual property laws and practices around the world.

Donors and dealers should be aware of other people's intellectual property in the digital files that they offer to a repository; for instance, a computer may be shared by coworkers or by an entire family and 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 will need to be respected in digital formats, but repositories may not be aware that the creations of others are present in the files, nor able 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 the repository's handling of 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 also 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.

Email and other digital correspondence
The sheer volume of sent, received, and saved email messages, as well as the presence of attachments in many different file formats, can complicate email acquisition. In addition, born-digital correspondence may include letters written with word-processing software and sent either

as attachments to email messages or printed and sent in physical form. These documents can present complications similar to those encountered in email archives.

Donors may want to screen email files for sensitive and/or extraneous messages prior to transfer. Repositories and donors will need to make clear by whom, and the process by which, this screening will be done. If a donor is not able or willing to screen for sensitive messages, the repository will need to make a decision, in accordance with policy, regarding whether and to what extent to devote staff time to searching for information above and beyond what a repository is required by law to restrict. If a repository decides to undertake detailed screening for sensitive materials as defined by a particular acquisition agreement, strategies for screening may be necessarily limited by staff resources. In some situations, access restrictions or an embargo period on the use of email can be implemented as a means of lowering risk 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 so that staff can lower the risk around those more easily. Preservation for access in the longer-term is also a reasonable option and for many repositories may be the most realistic alternative to screening.

Donors, dealers, and repositories need to collaborate with each other and discuss how to handle sensitive messages or message threads. If a repository discovers obviously sensitive messages not identified as such or removed by the donor, and the donor or intellectual property holder is no longer available, a decision will need to be made regarding whether to make the messages available, remove them, or retain them with a restriction on their access. In many repositories, the amount of labor required to screen email messages may prompt repository staff to explore alternatives such as embargo, or restricting access to materials for a certain period of time, when staff resources are limited.

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

Hidden content: password protection, firewalls, disk images
In some cases, especially if digital files come to a repository through an estate, materials may be hidden behind an unknown login or password. Digital files can provide the 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). 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/or logins, or to recover deleted files, and whether the donor or dealer grants ownership of files recovered by these methods. In some cases decryption and file recovery may be the only way to gain access to 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 of the scope of the acquisition.

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 a whole system disk include future interactive access, which could increase the value of the materials. Agreements should specify whether disk images are intended or allowed to be included in a collection of digital files, and, if so, whether the disk image might 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 carved out from them and made public, the disk images themselves will serve as a master copy, or original copy that remains untouched, that is restricted to staff use for preservation purposes.

Recommendations for Donors and Dealers

- Consider screening email files for sensitive and/or extraneous messages. When this is not possible, consider appropriate embargo periods and discuss restrictions on access to email before that date with the repository.
- Inform repository staff if there is a possibility that your 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 your digital records include legally protected private files, such as confidential government files, medical records, and legal case files, and other kinds of sensitive information, such as Social Security and credit card numbers.
- Consider whether a repository should be allowed to decrypt passwords and/or 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).

Recommendations for Repositories

- Make clear by whom, and the process by which, screening of email for sensitive messages will be done.
- Carefully review email 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 your acquisition agreement.
- When staff, time, or technology constraints do not allow for review of sensitive email or other files, discuss restrictions on researcher access for a defined period of time with the donor or dealer.
- Before acquisition, ask about the presence of legally protected private files, such as confidential government files, medical records, and legal case files, and 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 and/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:
    o Legal restrictions
    o Donor requests for restrictions
    o Third-party restrictions (e.g., oral histories without permission forms)
    o Technical constraints (e.g., obsolete formats, corrupted files, access problems)
    o Use restrictions (e.g., need to ensure authenticity and appropriate use by patrons)

Key Stages in Acquiring Digital Materials

This section identifies key stages in the acquisition of born-digital materials and offers recommendations to help repositories, donors, and dealers navigate this process.

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 for their potential research and iconic value and in order to recover more information as future technology advances. 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. Donors may wish to retain a copy of the digital files for their use. Repositories will want to clarify whether they will be the sole authorized entity for the ownership, preservation, and management of a body of digital files.

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, email archives, websites, etc.)
- 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, emails) will be captured, preserved, and made accessible
- What will be done with files that are transferred but do not fit 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, etc.) 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

Communication and interaction
Direct communication between a repository's staff and the donor is invaluable in ensuring the preservation and access of born-digital materials. Often the donor is the only person able to answer questions about the hardware or software used to create particular files, which materials

the donor intends to transfer, and issues related to privacy and sensitive information. The exact nature of these conversations may vary depending on 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
- The process 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 her methods of creating, storing, and maintaining digital files. The more detailed information the donor can provide the repository, the more success the repository will have in preserving born-digital materials. Similarly, repositories should consider involving suitable technical specialists from the outset, including for site visits, and before acquisition agreements are finalized.

Furthermore, if the repository intends to acquire future materials from a donor, early communication could 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. It might also be useful for the donor to involve her own 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 will not be possible for repositories to learn more about the digital materials in a collection.

Transfer of materials
Transfer strategies will 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. Determining the best method to transfer digital materials from a donor to a repository will depend 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 repository and donor. 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.

Once acquisition terms have been established, the repository and donor should determine how the born-digital materials will be transferred. It is strongly recommended that donors and dealers seek the guidance of archival repositories before any transfer takes place. Often, repository staff will prefer to retrieve the media and files in person. If a donor is to send disks, computer hard

drives, or other hardware directly to a repository, she 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 from moisture, extreme temperatures, strong electro-magnetic fields, and rough handling when being transferred to a repository. Any equipment should be cushioned during transport, and moving parts stabilized (e.g., inserting a dummy floppy disk into a floppy disk drive may be advised). Alternatively, files may be transferred electronically by secure means. All of these scenarios would 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 transfer to the repository. For example, donors or dealers should note whether files were copied from a computer hard drive onto a disk and identify the computer's make, model, and operating system, if known. It is also important to consider early on what security measures need to be in place regarding the storage, handling, and secure deletion of electronically transferred files.

Initial handling of materials at the repository
Digital media may pass through the hands of numerous staff members during the acquisition process. It is good practice for repositories to minimize the number of people who handle digital media and files, and for each transfer of digital materials to be well documented. After born-digital materials arrive on site, the repository will want to ensure that they are transferred to the appropriate department safely, quickly, and in 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 inspect the materials in preparation for announcements about the acquisition, exhibitions, or other purposes. Digital media in collections that have been in storage for an extended period may not be discovered until archival processing is underway.

Because attempts to access born-digital materials can change the content, formatting, and metadata associated with the files, it is important for repositories to establish clear protocols for how these materials should be handled by staff. 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.

Recommendations for Donors and Dealers

- Discuss whether the repository will be the exclusive owner of the digital files you transfer, and whether and how copies of the digital files will be allowed.
- 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 only certain types of files.
- Be prepared for the repository to retain the original digital media unless the donor or repository 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.
- If a donor has a technical specialist, consider involving her 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 earlier computer equipment is still in the donor's possession and 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 how 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.

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 don't 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 between the donor or estate and repository staff.
- 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

This section 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.

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 will also need to undertake subsequent assessment after materials have arrived on-site. The post-acquisition review process can also 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.

Physical condition
In order for staff to capture and preserve their contents, digital media need to be reasonably clean and physically intact. Providing donors and dealers with storing and packing recommendations can help protect digital media from harm prior to or during transit. In some cases computer media will have long ago sustained damage. 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 arrives at the repository in a different condition than expected, repository staff will need to decide whether to accept the materials or revisit the repository's agreement with the donor or dealer. Physical damage to media may also change the level of preservation and access a repository can provide. Yet donors, dealers, and repositories should not assume just because digital media may be decades old and battered looking that nothing of value can be recovered from them.

Furthermore, repository staff should pay close attention to the physical labels on digital media. Although physical labels are not always accurate, sometimes they are richly detailed and useful, and therefore may help staff verify inventory and conduct an assessment of the media at the repository. Sometimes information on a label may be additional to any digital content of the media item (e.g., the name of the original owner of the disk).

Digital condition
It may be preferable for repository staff to be the first to preview and capture disk media, using write protection technologies that protect the original media from inadvertent change. In doing so, 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, a different hash value will be produced by the altered file; conversely, if in 100 years' time a file is subjected to the same test and it yields the same hash value as 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 undertake this task themselves, perhaps with the assistance of a technical specialist. If so, hash information can help staff verify that the files arriving at a repository match those transferred by the donor. Comparing and confirming hash values generated prior to transfer with those generated once the materials have arrived on-site requires a certain level of technical comfort and expertise on the part of both the donor and the receiving repository. Mismatched hash values would indicate that the materials that arrived at the repository had been altered or corrupted in transit. Staff would need to determine 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.

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

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 all be candidates for return to the donor followed by secure destruction. Any action should be well documented and supported by the acquisition agreement and a repository's policies.

For example, repositories might consider developing a retention policy for damaged media whose contents are ultimately inaccessible. An acquisition agreement might include language stating how and under what circumstances materials may be culled from a collection, and be supported by a policy outlining how these materials will be disposed of securely. Equally, 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 might still hold cultural value as a physical object or even as a teaching tool. Even if some files are currently unavailable due to 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.

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.
- Consider the likelihood that media damaged in shipping or files corrupted in transfer have lost their cultural value. This possibility underscores the importance of agreed-upon transfer methods.

Recommendations for Repositories

- Assess whether digital media have been damaged in transit.
- Assess whether files may have been affected by transit or 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 still 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.

Conclusions

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

The sections of this report describe 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, dealers and repositories should work to better understand and document donor computing habits to improve the quality of digital transfers and accessions. Earlier archival intervention in records and information management will help shape the archival impact 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 between donors, dealers, and repository staff and establishing transparent, efficient archival practice promise to minimize unpleasant surprises and improve the quality of born-digital acquisitions.

Appendix A: Preparing for the Unexpected

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 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 desire to gather information about donor's working environment with need to respect donor privacy and wishes. • Develop strategies to manage restriction, redaction, and discovery of sensitive material that may have been overlooked in initial review.
Collaboration	• Ask repositories for guidance and documentation. • If possible, establish direct communication between the donor and 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 (IT) or legal support people in the process, if needed and/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 its original source before offering the material for gift or purchase; or do so in consultation with repository staff. • Most donors will want to screen email files for sensitive and/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 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 email for sensitive messages. • Consider embargo as 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. • Identify areas of concern to flag 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 her 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 what type of access users may have to processed collections.

Appendix B: Staffing: Potential Costs 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. Possible staff activities are shown below with some indication of how the activity might be performed, and who might be involved in it.

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)	On-site	Subject specialist evaluates potential research value of the material offered.
Initial survey (technical)	May be a mixture of remote information gathering and on-site analysis	Process and technical specialists establish technical characteristics to inform capture and preservation techniques.
Research new capture scenarios (e.g., donor has special scoping requirement; new data format)	Remote	Process and technical specialists research and determine capture techniques for new scenarios. May require input from donor's technical support and/or repository technical staff.
Interaction with donor to define scope and explain processes	On-site/remote	Subject specialist will likely lead 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 will begin discussion with a template terms of agreement. Subject specialists and process specialists may make collection-specific provisions if required. The repository's legal team will support this work.
Actual capture of material (especially for on-site capture)	On-site/remote	Process specialist executes capture. This process can be time-consuming, depending on the scale of the material involved.
Create 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)	On-site/remote	Curatorial and process specialists remain available for future consultation with regular donors as new scenarios dictate re-working of capture procedures.

Appendix C: Checklist of Recommendations for Donors and Dealers

Initial Collection Review
- ☐ 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, do so in accordance with established guidelines provided by the repository.
- ☐ Ask repositories for guidance and/or documentation on:
    - o Determining the most appropriate repository for your digital materials
    - o Negotiating the terms of an acquisition agreement or contract as it relates to born-digital materials
    - o Describing the context and history of the files and media being transferred
    - o Handling digital media
    - o 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 email files for sensitive and/or extraneous messages. When this is not possible, consider appropriate embargo periods and discuss restrictions on access to email before that date with the repository.
- ☐ Inform repository staff if there is a possibility that your 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 your digital records include legally protected private files, such as confidential government files, medical records, and legal case files, and other kinds of sensitive information, such as Social Security and credit card numbers.
- ☐ Consider whether a repository should be allowed to decrypt passwords and/or 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 you transfer, and whether and how copies of the digital files will be allowed.
- ☐ 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 only certain types of files.
- ☐ Be prepared for the repository to retain the original digital media unless the donor or repository 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.
- ☐ If a donor has a technical specialist, consider involving her 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 earlier computer equipment is still in the donor's possession and 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 how 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.
- Consider the likelihood that media damaged in shipping or files corrupted in transfer have lost their cultural value. This possibility underscores the importance of agreed-upon transfer methods.

Appendix D: 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, with the cost of capture and ongoing preservation and access.
- Share relevant information and documentation with donors and dealers about the collection review and acquisition processes.
- Clarify expectations regarding digital preservation and access.
- Conduct a survey of born-digital materials:
  - Determine the method by which to do this through consultation with the donor or dealer.
  - Conduct in-person/on-site surveys when possible.
  - Have policies in place regarding the capture, storage, and disposal of files copied for the purpose of preliminary assessment.
  - 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 on-site survey is not possible.

Privacy and Intellectual Property
- Make clear by whom, and the process by which, screening of email for sensitive messages will be done.
- Carefully review email 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 your acquisition agreement.
- When staff, time, or technology constraints do not allow for review of sensitive email or other files, discuss restrictions on researcher access for a defined period of time with the donor or dealer.
- Before acquisition, ask about the presence of legally protected private files, such as confidential government files, medical records, and legal case files, and 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 and/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, access problems)

    o 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 don't 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 between the donor or estate and repository staff.
- 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

- Assess whether digital media have been damaged in transit.
- Assess whether files may have been affected by transit or 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 still 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.