

Developing an Access Strategy for Born Digital Archival Material

Brian Dietz and Shira Peltzman



**DPC Technology Watch
Guidance Note**

September 2021



Digital Preservation Coalition

© Digital Preservation Coalition 2021 and B Dietz & S Peltzman

ISSN: 2048-7916

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, without prior permission in writing from the publisher. The moral rights of the author have been asserted.

First published in Great Britain in 2021 by the Digital Preservation Coalition.

This report is made available under the [CC BY-NC-SA Licence](https://creativecommons.org/licenses/by-nc-sa/4.0/).

1 Introduction

Access is the primary means by which an archive meets its responsibility to its user community, and as such it represents the axis around which many of the core values and ethical principles that guide archival practice rotate (Society of American Archivists' Technical Subcommittee on Describing Archives: A Content Standard, 2021). Simply put, archives exist to be used, and it is our responsibility as stewards of archival material to make this a reality.

While most institutions have established access solutions for their physical collections, for many institutions, making their digital material accessible to the public remains a challenge. This Guidance Note is written for practitioners who do not currently have any access strategy in place at their organizations for digital materials, and are looking to implement a basic mediated access solution that can accommodate a range of budgetary considerations and technical skill sets. These recommendations will cover all fundamental aspects of access, including physical considerations, policies, and tools that institutions can consider implementing according to their needs, resources, and abilities. While this Guidance Note is written with born digital materials in mind, the recommendations presented here are also suitable for digitized material that cannot be made available online for wide distribution.

Digital access is an underexplored area of archival praxis. In spite of the central role that it plays, there is little agreement about what access to digital material should look like, or how this should be implemented from one institutional context to the next. Every institution has its own challenges and needs. There is no single 'correct' setup or a 'one size fits all' approach. It follows, then, that the suggestions presented here do not constitute the only way (or even the best way, necessarily) for practitioners to accomplish their goals. Our aim is to furnish readers with a set of high level considerations for starting an access program where none existed previously. This resource is not intended to provide a detailed blueprint for achieving this, but rather to serve as a jumping off point for further research. To that end, these recommendations represent a starting point rather than a destination.

2 What makes born digital access challenging?

The inherent properties of digital archival material offer certain possibilities in the context of developing an access strategy. But alongside these opportunities, there are also challenges that must be addressed.

2.1 Security

Unlike the materials that comprise physical collections, digital information can be infinitely reproduced without any loss of quality. The ease with which files can be reproduced, edited, or even distributed from one location to another represents a fundamental difference that any access strategy must take into account. This is an essential consideration in light of the presence of copyright protected, privacy protected, sensitive, or personally identifying information.

2.2 Technical Dependencies

Digital archival materials are not independently understandable - all files have a set of technical interdependencies. Some may be self-evident, such as a password for a social media or an email account, or the hardware necessary to transfer files from the external storage device on which they were received. Others may be less obvious, such as the operating system and software requirements necessary to render a specific file format. For instance, software created to work properly on one

operating system may not work on an earlier or later version, much less one that is designed for another platform altogether (for example, Windows 95 vs. Macintosh OS9).

Understanding and documenting these dependencies is important in all cases, but it is especially vital for outdated file formats that are no longer in use. This is because legacy files created on an older computer environment may be incompatible with modern systems. In this scenario, additional digital curation may be required in order to render the material. This may include strategies such as providing users with a file viewer such as Quick View Plus ([Avantstar](#), 2020) that enables access to older file formats and versions, migrating the file to a newer format that can be easily rendered on contemporary machines, or emulating the older computer environment that produced the file so that it can be accessed virtually through a project such as EaaS ([Software Preservation Network](#), 2021).

2.3 Additional Considerations

The challenge of rendering legacy files can be compounded by the sheer range of formats in existence. A single hard drive alone could contain hundreds of different file formats, each of which may have distinct software requirements.

3 Getting Started

These issues are materially different from the concerns associated with providing access to physical materials. Making files accessible requires a unique skill set as well as a dedicated set of policies and procedures. The considerations in this Guidance Note will help your organization develop an access strategy that meets your particular set of needs.

3.1 Access Policy

Updating the existing access policy (or developing such a policy, if one is not already in place) to include information specific to digital collections is an important step for both users and staff alike. For staff, an access policy enables programmatic decision making so that you are not approaching access on a case-by-case basis; for users, an access policy establishes a clear set of expectations regarding how collections may be accessed and used. Some key features of an access policy for digital materials include communicating what types of file formats can be made accessible and whether there are duplication restrictions in place. The policy should also alert users to the potential presence of privacy protected, sensitive, or personally identifying information, including what to do if they encounter these materials. Depending on your locality and institutional context, this may be accompanied by indemnification language stating that the user assumes all responsibility for harm stemming from their use and/or reproduction of the materials -- for example, infringing on copyright or the right to privacy. However, it is advisable to familiarise yourself with the relevant national legislation and related internal policies and to confer with legal counsel before putting indemnification language into practice.

The policy should also inform users of what accommodations are available for users with disabilities, such as what assistive technology is allowed and/or pre-installed for use, and whether users may bring a companion with them to assist with their research. For examples of how some institutions have handled these concerns, the [University of California, Santa Cruz](#) (2021) includes privacy and copyright-protected indemnification statements in their publicly available access policy, and the [University of California, Los Angeles Library](#) (2018) addresses accessibility concerns in theirs.

3.2 Providing access to digital materials

3.2.1 Onsite workstation

The most common approach to providing onsite, offline access to digital archival material (and one that supports the unique challenges of digital materials with restrictions) is to place the files on a dedicated access workstation that users can sign up to use in advance. This strategy represents the most secure approach to access, provides an effective method for supporting technical dependencies, and has the benefit of following the familiar template of providing onsite access to physical materials. The primary expense in this scenario is the cost of the computer itself. Although there can be other costs associated with equipping the workstation such as purchasing software licenses, there are free, open-source alternatives that can be readily found.

3.2.2 Security

To help guarantee the security and integrity of the files being accessed, create an administrator account and a separate generic user account. The administrator account will be used to load files on the computer and configure the user account so that the user is not able to make permanent changes to the files. This can be done at no cost by setting file permissions to read-only or by purchasing a reboot-to-restore application like Deep Freeze ([Faronics, 2021](#)). If the files have (or may have) restrictions, then disable wifi and access to external ports for an additional layer of protection. This will help ensure that the files remain on the machine and are not copied to an external device or uploaded to the internet. Disabling wifi on Windows and Mac OSX is easy and straightforward. However, disabling USB ports, while possible, should be done with extreme caution, and, preferably with the support of information technology (IT) professionals. Third party applications, such as [Endpoint Protector](#) (2021), may also be used to perform this task. You can also enable security settings using the 'parental controls' feature available on both OSX and Windows. If none of these technical solutions are feasible, then directly supervising the user for the duration of their visit could offer a low tech (if less than ideal) alternative.

3.2.3 Technical Dependencies

The workstation should include software that enables the user to view the range of files present in your collections. This typically includes applications that can open and render widely used formats, such as image files, PDFs, office documents, audio/visual materials, email, and text and tabular data. Your organization may already have licenses for many of these applications; if not, free versions are readily available. You may need to install additional software to support less widely used files present in the collection. For instance, if you are providing access to digital architectural blueprints, you will likely need to install specialized software that can render Computer-Aided Design (CAD) files. Additionally, if your organization provides access to disk images, you will need software that allows users to interact with them (for example, by mounting them and extracting or rendering their content).

For organizations with collections that include an array of file formats and content types that may have been created in software that has since become outdated, consider purchasing a file viewer. This allows users to view a range of different file formats without the need for the specific application in which they were originally created. For unknown files and files that will not open, or just for deeper analysis, a [hex editor](#) (2021) can provide helpful clues as to what a file may contain by allowing a user to look at the ones and zeroes that comprise it. However, these programs may not be self-explanatory to all users, and staff should be prepared to instruct and assist users as to their navigation and use.

To support a broad range of users, there are assistive settings that can be customized, as well as applications that can be installed, to improve the accessibility of material for users with disabilities. Customizations include ease-of-use settings like magnification and screen contrast. [UC Berkeley](#) (2021) has compiled a list of other assistive technologies, along with their purposes, that could be incorporated into an access workstation.

3.2.4 Alternative onsite solution

If you do not have an onsite workstation or the necessary customizations present a barrier (financial or otherwise), another lower cost onsite option is to lend an external hard drive that is pre-loaded with copies of the requested files directly to the user, who can then use their own computer to access the files and download any software necessary to view them. However, the user's computer may not include the software necessary to view the range of file formats present, requiring them to download the appropriate applications on their own. This could represent a considerable barrier to use for collections featuring less common and/or legacy file formats. Consequently, this access strategy is best suited to unrestricted materials (for example, those without privacy, sensitivity, and copyright concerns) and consisting of common file formats that are easily renderable on most modern machines. Something to keep in mind is that you may not be able to keep the researcher from copying files over to their computer. This option may be especially helpful for users with disabilities who have necessary accessibility software already installed and customized on their own computer. Before lending a hard drive, verify that the files on it do not contain malware by running a virus check. After getting the drive back from the researcher, reformat it so that all previously loaded content is fully removed.

3.2.5 Remote solutions

Providing remote access can also be considered. This can be especially helpful in situations where onsite resources like staffing or space are limited. Examples of this approach include sending files to a user through email, or sharing files through cloud storage or via a File Transfer Protocol (FTP) site. These options provide faster and more democratic access to materials because they do not require the user to be physically present. However, remote access comes with its own set of challenges. As in the 'lending a hard drive' approach, this option is less suitable for collections featuring less common and/or legacy file formats and/or for restricted materials with concerns related to privacy, sensitivity, and copyright.

Levelling up your remote access support could include managing a virtual 'reading room.' Possible examples include an online repository with files accessed using a web browser, or a virtual computer that you manage and on which you install the appropriate software, and that a user can access from their own computer using a remote desktop client. In both of these scenarios, you can require users to register and/or authenticate in order to access materials. Arroyo-Ramirez, Berdini, Black, Cram, Gronsbell, Krabbenhoeft, Lynch, Preston, Smedberg (forthcoming) explore various challenges and affordances of virtual access models. Additionally, some file formats, such as digital images, lend themselves to online repositories and digital collections sites, which may be good options for unrestricted collections provided that there are no copyright or privacy concerns. These solutions require significant software development and support, but if you have ready access to similar systems, their use may be suitable.

3.2.6 IT Considerations

Get to know your IT department (if you have one). Regardless of your chosen access strategy, IT can provide recommendations on computer specifications like memory and processing speed, and they can also provide guidance about what software licenses your organization already has or can readily acquire. However, do not assume that because IT are computer experts, they understand the particular needs associated with providing access to digital materials. These requests are unique, and you will likely need to advocate for and explain the reasoning behind the customizations that you are requesting.

North Carolina State University Libraries makes born digital personal histories available in its 'Rare & Unique Digital Collections' site, for example "[Phil Freelon oral history, 8 May 2017](#)"; University of North Carolina-Chapel Hill makes some born digital materials available in their repository, including [Lauren Adams, "Crazy Quilt": Photos of installation, individual collage files, and Ackland Project Description documents](#); and the Schweizerisches Sozialarchiv has demonstrated the use of [IIIF standards](#) to deliver [born digital materials online](#), as has Princeton University, in [Allen W. Dulles Papers: Digital Files Series, 1939-1977](#).

Customizing the workstation requires administrator privileges. Not all IT departments allow for this, so you may need to collaborate with IT to find a solution -- for example, working with IT colleagues to request that they install relevant software and configure the computer's settings on your behalf.

4 Levels of Access

Additional guidance for making digital materials available to users can be found in the Digital Library Federation's Born-Digital Access Working Group's [Levels of Born-Digital Access](#) (2020). This resource provides benchmarks to guide policy and decision making in the form of a tiered set of format agnostic practices to facilitate and improve access to born digital materials. The document addresses accessibility, description, researcher support and discovery, security, and tools. Each of these areas is organized into three levels of progressive complexity. The document provides resources, examples, and tips for each area and level, and there is a one page table for quick visual reference.

5 References

Arroyo-Ramirez, Elvia, Berdini, Annalise, Black, Shelly, Cram, Greg, Gronsbell, Kathryn, Krabbenhoef, Nick, Lynch, Kate, Preston, Genevieve, and Smedberg, Heather (forthcoming). *Speeding Towards Remote Access: developing shared recommendations for Virtual Reading Rooms*. In Matienzo, M., ed., *Lighting the Way Handbook*.

Avantstar (2020). *Quick View Plus 2020*. Available at: <https://www.avantstar.com/>. [accessed 12 August 2021]

Faronics (2021). *Deep Freeze*. Available at: <https://www.faronics.com/products/deep-freeze>. [accessed 12 August 2021]

EndPoint Protector (2021). *EndPoint Protector*. Available at: <https://www.endpointprotector.com/>. [accessed 12 August 2021]

Digital Library Federation Born-Digital Access Working Group (2020). *Levels of Born-Digital Access*. Available at: <https://osf.io/r5f78/>. [accessed 12 August 2021]

"Hex editor" (2021). *Wikipedia*. Available at: https://en.wikipedia.org/wiki/Hex_editor. [accessed 12 August 2021]

Society of American Archivists' Technical Subcommittee on Describing Archives: A Content Standard, (2021). *Statement of Principles*. Available at: https://saa-ts-dacs.github.io/dacs/04_statement_of_principles.html [accessed 12 August 2021]

Software Preservation Network (2021). *EaaS: Emulation-as-a-Service Infrastructure*. Available at: <https://www.softwarepreservationnetwork.org/emulation-as-a-service-infrastructure/> [accessed 12 August 2021]

UC Berkeley (2021). *Types of Assistive Technology*. Available at: <https://webaccess.berkeley.edu/resources/assistive-technology> [accessed 12 August 2021]

University of California, Los Angeles Library (2018). *Access Special Collections in the Reading Room: Patrons with Disabilities*. Available at: <https://www.library.ucla.edu/location/library-special-collections/research-instruction/access-special-collections-reading-room> [accessed 12 August 2021]

University of California, Santa Cruz University Library (2021). *Our Policies*. Available at: <https://guides.library.ucsc.edu/speccoll/policies#s-lg-box-wrapper-20360309> [accessed 12 August 2021]

Special thanks to Charlie Macquarie, farrell, Danielle Moses, and Margo Padilla.