2018

"Tell Us about Your Digital Archives Workstation": A Survey and Case Study

Elvia Arroyo-Ramírez
University of California, Irvine, elvia.ar@uci.edu

Kelly Bolding
Princeton University, kbolding@princeton.edu

Faith Charlton
Princeton University, faithc@princeton.edu

Allison Hughes
New York Public Library, allisonhughes@nypl.org

Follow this and additional works at: https://elischolar.library.yale.edu/jcas

Part of the Archival Science Commons

"Tell Us about Your Digital Archives Workstation": A Survey and Case Study

"Tell Us about Your Digital Archives Workstation": A Survey and Case Study

Introduction

The decision to assemble a digital processing workstation came out of a growing priority for Princeton Rare Books and Special Collections staff in the Manuscripts Division: to gain intellectual and physical control of our born-digital and audiovisual content.

Surveying Born-Digital Content

Manuscripts Division processing staff began to proactively manage born-digital content in 2015 when we participated in the Society of American Archivists' Jump In 3 initiative[1] to survey and identify legacy digital media. Staff conducted an XQuery search over our published finding aids, followed by a physical survey, and found that the division had at least 394 items, primarily floppy disks and optical media, in forty-two of approximately 1,600 collections. We determined that over 4 terabytes (TB) of storage space would be needed to preserve and provide access to these materials.

Participating in the Jump In 3 survey came at an opportune time as the Manuscripts Division began to acquire an increasing number of hybrid collections. One such collection, the Toni Morrison Papers,[2] was the first hybrid paper and digital collection we processed. Processing work began on another collection, the Juan Gelman Papers,[3] shortly thereafter. These collections combined caused the estimated total number of floppy disks present in the division's collections to quadruple. The division has since acquired or processed an additional 500 gigabytes (GB) of born-digital media, including media with larger storage capacities such as internal and external hard drives.

Surveying and Digitizing Audiovisual Content

Around the same time, processing staff in the Manuscripts Division began to identify and prioritize legacy audiovisual media for reformatting and digitization. Based on an initial survey of finding aids, staff identified at least 1,984 audiovisual items present in ninety-five collections and have, as of November 2018, physically surveyed 2,493 items from forty-eight different collections. While conducting this survey, archival staff also undertook several digitization projects: a collaborative effort with the PEN American Center to digitize audio and video recordings of their events; a Latin American Collections audio pilot program; and two requests from researchers to digitize audiovisual media. As a result of these reformatting projects, the division received approximately 26 TB of preservation masters, intermediate files, and access copies of digitized files. Staff quickly realized that, in addition to needing a workstation to prepare born-digital content for long-term preservation and access, the division was in need of a staging area to effectively manage incoming large batches of digitized audiovisual content.

---

[1] "Jump in 3 2015 Results," Society of American Archivists, https://www2.archivists.org/groups/manuscript-repositories-section/jump-in-3-2015-results.

[2] Toni Morrison Papers (C1491). Princeton University Library, https://findingaids.princeton.edu/collections/C1491.

[3] Juan Gelman Papers (C1511). Princeton University Library, https://findingaids.princeton.edu/collections/C1511.

First Digital Archives "Workstation"

The division's first semblance of a digital processing workstation came in the form of a dual booting Windows 7/BitCurator (Linux) laptop, acquired in 2015. Because the bulk of digital material staff were processing was stored on floppy disks, this laptop, used in conjunction with an array of external disk drives, initially met our needs.[4]

While the laptop provided us with a portable means to begin capturing and processing digital content, it proved limiting in several respects. First, it was unable to serve as an end-to-end processing station as it lacked a secure connection to transfer born-digital files to the library's preservation storage environment. Second, as staff began to process content from carriers with higher storage capacities like compact discs (CDs) and digital video discs (DVDs), it soon became apparent that the processing power and local storage of the Dell laptop was insufficient. In order to save files to the library's secure server and to process larger capacity carriers, the Manuscripts Division staff had to rely on the University Archives' digital processing workstation, a Forensic Recovery of Evidence Device (FRED) machine.[5] However, the two repositories are physically located in separate buildings. In order to complete processing responsibilities, Manuscripts Division staff had to carry media across campus, compromising both staff time and the security of the materials.

The surveys that staff conducted and the digitization projects that we oversaw provided us with a base of knowledge to help us determine what other equipment would be necessary for us to capture, transfer, and store born-digital and digitized materials. In order to build on this experiential knowledge, we decided to conduct further research on current best practices in digital curation and born-digital processing; in particular, about the types of hardware our peers were using to conduct born-digital processing work.

Literature Review

While there are a number of reports available about how to establish a born-digital processing program broadly, for this project we relied particularly on a smaller body of literature focused on the development of digital processing workstations and their component parts. Ricky Erway's 2012 Online Computer Library Center (OCLC) Research report broke down the technical process for acquiring born-digital content from physical media into discrete steps.[6] Erway and Julianna Barrera-Gomez followed this up in 2013 with a more detailed set of guidelines for specific workflow steps, with each step accompanied by a list of potential hardware and software

---

[4] These drives included a 3.5-inch floppy disk USB drive, a FC5025 USB 5.25-inch floppy controller, a TEAC FD-55GFR 5.25-inch floppy disk drive, and a KryoFlux. For more information, see Kelly Bolding, "Tooling Up: Building a Digital Processing Workstation," *This Side of Metadata* (blog), Princeton University Library, July 19, 2016, https://blogs.princeton.edu/techsvs/2016/07/19/tooling-up-building-a-digital-processing-workstation/.

[5] A computer designed for forensic analysis made by Digital Intelligence, https://digitalintelligence.com/products/fred/.

[6] Ricky Erway, *You've Got to Walk Before You Can Run: First Steps for Managing Born-Digital Content Received on Physical Media* (Dublin, OH: OCLC Research, 2012), http://www.oclc.org/research/publications/library/2012/2012-06.pdf .

tools available to facilitate that step, as well as a list of related reports and case studies, "experience-based advice," and sample workflows.[7] Though our research was primarily focused on various hardware specifications, as well as hardware and software combinations, we also found useful information about particular tools in COPTR (Community Owned digital Preservation Tool Registry), a collaborative digital preservation software tool registry.[8]

Concurrently, archivists at several institutions began publishing blog posts and articles about assembling custom digital archives workstations to facilitate born-digital processing workflows in their own institutions. Porter Olsen's blog posts from 2012 and 2013 document the development of a digital curation workstation at the Maryland Institute for Technology in the Humanities (MITH) for the purpose of testing the BitCurator environment.[9] With the goal of building a relatively affordable custom personal computer (PC) workstation that could handle as many media formats as possible, Olsen highlighted considerations about multiple central processing unit (CPU) cores, random access memory (RAM), and hard drive storage space for the main computer, and discussed options for memory card readers, disk drives, and write protectors. Ben Goldman of Penn State University published an article in 2014 providing a similar set of baseline hardware specifications for equipping a Windows-based PC workstation.[10] Goldman expands the recommended amount of RAM specified by Olsen the year before, an update that supports his closing point that archivists should "consider the tools and technologies we employ to be transient and in constant need of review and revision."[11] In 2016, the Harvard Library Digital Forensics Working Group conducted an internal survey on digital forensics practices and priorities across sixteen of Harvard's thirty-seven repositories, as well as interviews with representatives from thirteen other universities regarding their processing tools, workflows, and staffing models.[12]

Archivists have also adapted the Forensic Recovery of Evidence Device (FRED) from Digital Intelligence as a core part of their digital archives workstations. In 2015, Alice Prael and Amy Wickner wrote about selecting and testing a flexible set of hardware and software tools in conjunction with the FRED acquired by the Special Collections and University Archives (SCUA) at the University of Maryland Libraries in 2013.[13] Prael and Wickner's approach initially involved partitioning the FRED to operate both Windows and BitCurator (Linux) to

---

[7] Julianna Barrera-Gomez and Ricky Erway, *Walk This Way: Detailed Steps for Transferring Born-Digital Content from Media You Can Read In-house* (Dublin, OH: OCLC Research, 2013), http://www.oclc.org/content/dam/research/publications/library/2013/2013-02.pdf.

[8] "Community Owned digital Preservation Tool Registry (COPTR)," http://coptr.digipres.org/.

[9] Porter Olsen, "Digital Curation Workstation," *MITH* (blog), November 26, 2012, http://mith.umd.edu/digital-curation-workstation/; Porter Olsen, "Building a Digital Curation Workstation with BitCurator (update)," *BitCurator* (blog), August 2, 2013, https://www.bitcurator.net/building-a-digital-curation-workstation-with-bitcurator-update/.

[10] Ben Goldman, "Outfitting a Born-Digital Archives Program," *Practical Technology for Archives*, no. 2 (2014), https://practicaltechnologyforarchives.org/issue2_goldman/.

[11] Ibid.

[12] "At Home with Digital Forensics: A Conversation about Harvard Workflows" (slides), Harvard Library Digital Forensics Working Group, 2016, https://docs.google.com/presentation/d/1sttFttVdSCjZK05YMQWm9FsxlQLmIF4CN5CKf2ma0UY; "Bibliography on Digital Forensics," Harvard Library Digital Forensics Working Group, Harvard Wiki, https://wiki.harvard.edu/confluence/pages/viewpage.action?pageId=213390414.

[13] Alice Prael and Amy Wickner, "Getting to Know FRED: Introducing Workflows for Born-Digital Content," *Practical Technology for Archives*, no. 4 (2015), https://practicaltechnologyforarchives.org/issue4_prael_wickner/.

allow for experimentation with multiple disk imaging tools. Due to the resulting need to regularly switch partitions and wipe the BitCurator partition in order to update the operating system, they ultimately recommended installing BitCurator on a separate dedicated computer and using a local area network (LAN) to access working files between workstations. Echoing Goldman, Prael and Wickner also stressed the importance of flexibility—in their case, to account for anticipated shifts in workflows following the the library's implementation of a digital repository system.

Beyond formal literature and blog posts, the team also consulted a previous digital archives workstation survey and proposal drafted by our colleague Jarrett Drake, formerly the digital archivist for Princeton's University Archives, to acquire the aforementioned FRED for the University Archives in 2014. Our approach was largely inspired by his example.

Survey/Research

Methodology

Given the importance of updating recommendations for digital processing tools regularly, as well as our practical objective of selecting tools for a workstation best suited to the needs of the Manuscripts Division, we decided to conduct our own survey of current digital archives workstations at other institutions. Since the survey originated from this pragmatic objective, as opposed to being conceptualized from the outset as a formal research study, our methodology was largely informal.

We targeted our survey, where possible, toward other manuscript repositories and special collections whose needs we expected to mirror ours. In choosing repositories and individuals to survey, we selected representatives from parallel departments at peer institutions; we also mined the attendee list from the 2016 Born Digital Archiving & eXchange (BDAX) unconference,[14] as well as our existing contacts, including those with whom we had had prior informal conversations regarding workstation development. We reached out to archivists and digital preservation specialists at twenty-two peer institutions, and we received responses from representatives at twenty of those institutions, one of which included information about two of their archival repositories.[15]

The initial survey was conducted in October 2016 via an email sent individually to contacts at each of the twenty-two institutions.[16] The email contained a brief overview of our current

---

[14] Notes from BDAX 2016 can be accessed through this shared Google Drive folder: "BDAX2016," https://drive.google.com/drive/folders/0By4GpQlPYswAbDJHLUsxcXZPOUk.

[15] Some of the responses we received were from different individuals than those we had initially contacted at a particular institution, perhaps because of the variety in job titles held by those who are responsible for managing born-digital processing. The authors sometimes looked for the word "digital" in a job title as a clue in locating a contact person at each institution we surveyed; ironically, none of archivists who worked on this proposal at Princeton had "digital" in their job titles.

[16] In retrospect, our initial data analysis workflow would have been more efficient had we employed Google Forms to collect more discrete data points from the outset. However, our method suited our purposes at the time.

workstation and our objective in conducting the survey, accompanied by a set of three open-ended questions:

1. What combination of hardware, software, peripheral devices do you use to process and prepare born-digital materials for long-term storage?
    a. What are the technological specifications of your digital archives workstation? (processor, RAM, hard drive size, operating systems, etc.)
2. Is your current digital archives workstation meeting the needs of your repository? Your needs as an digital archivist? Where does it fall short?
3. Are there tools out there that you are interested in trying out?

The responses we received were, like our questions, largely narrative; in many cases, we followed up by email or phone to get additional clarification on specific tools or workflows mentioned by respondents.

As we began receiving responses, we started compiling data into a Google Sheet, parsing the responses into more discrete data points. In addition to the specific topics we asked about in our email, we created columns for other useful contextual information respondents provided, such as links to workflows and other online documentation. The data in this spreadsheet informed our selection of a workstation, and the formal proposal we eventually drafted to acquire a FRED included the spreadsheet as an appendix.

Following conversations that grew out of the survey process, it became clear that many individuals we contacted were interested in using the data we gathered to help with their own workstation development. This interest led to the team's decision to update the survey results for the publication of this article. In November 2017, we sent respondents a follow-up email linking to a Google Form that provided several options for various levels of consent, information-sharing, and attribution. We also inquired about any workstation updates in the intervening year and allowed respondents to review and update their information in the shared spreadsheet. Statistics and tables provided in this article are based on updated information respondents provided.

Findings

While the data presented in this article are not intended to provide a comprehensive picture of digital processing workstations in the field, they can provide a glimpse into what tools and specifications are in use at a number of institutions. It should be noted that the authors originally did not intend for this information to be used for computational or statistical purposes. Thus, readers should be aware that since this data is only semi-structured, some of the figures that are provided are approximations. The twenty-one repositories that responded provided information relating to about sixty computers in total—several repositories included information for multiple workstations. Because of this, our spreadsheet often does not provide a one-to-one correlation between the computers that repositories are employing and their hardware and software specifications.
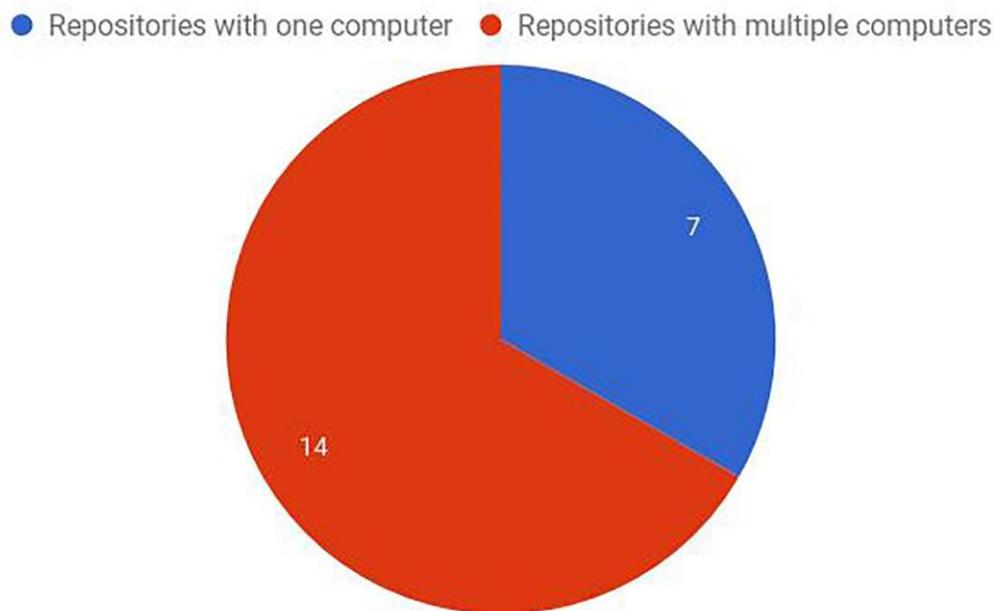
Figure 1. Number of Computers per Repository

The primary question we hoped to answer with our survey was: What hardware and software specifications are archival institutions currently employing? We were especially interested in learning about the use of FREDs, which come prepackaged for digital forensics work, versus PCs adapted to meet digital processing requirements. Survey results indicate that fourteen of the twenty-one repositories, or 66 percent, use more than one dedicated computer or workstation to conduct digital processing (see figure 1). Nine of the twenty-one repositories, or 40 percent of respondents, have at least one FRED. Of the nine repositories using a FRED, seven are using at least one FRED in addition to one or more computers. Two repositories indicated an interest in acquiring a FRED in the future. Seventeen of the repositories are using at least one PC. Three repositories reported outfitting their workstations with Macs; two repositories reported using vintage computers (see figure 2).
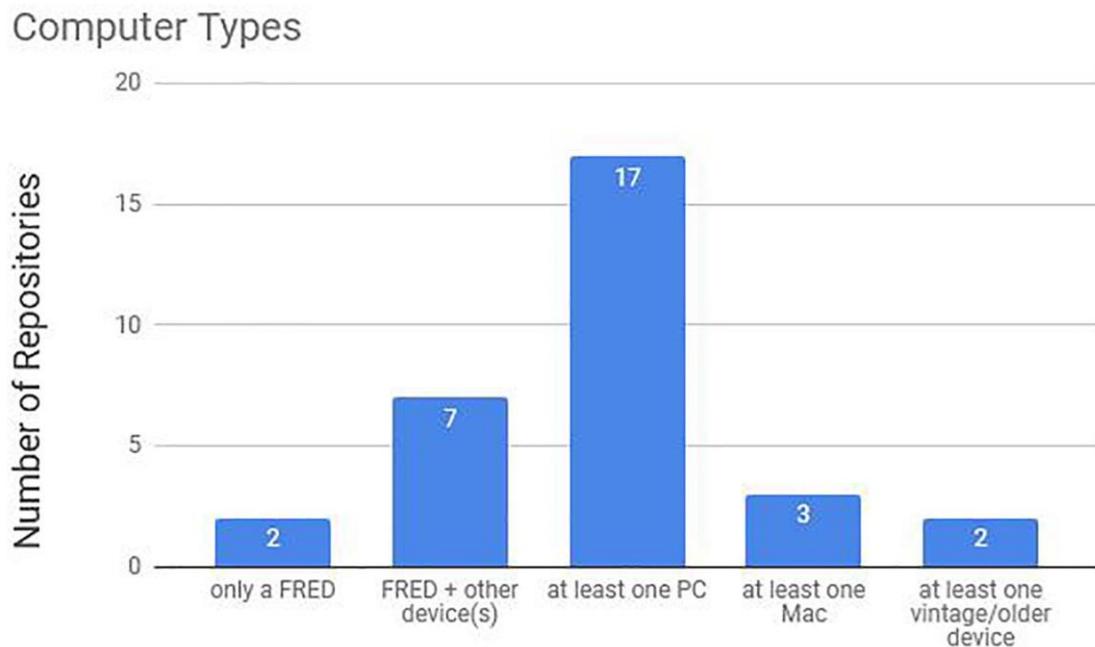
Figure 2. Computer Types

Respondents reported having the following hardware specifications: between 4 and 128 GB of RAM, with most between 8 and 32 GB; an average drive for operating systems of 512 GB; and an average size of additional hard disk drives and solid state drives of 4 TB. Participants also indicated that their storage for working files is between 8 TB and 30 TB.

We were also interested to know whether institutions were partitioning devices between Linux/Ubuntu (specifically BitCurator) and Windows. We were already aware of some of the resulting complications of this practice based on our own experiences using a partitioned laptop and the experiences shared by other BitCurator Consortium members on the BitCurator Users Google Group.[17] About half of the computers for which respondents provided information run Linux and the remaining half run Windows. A dozen computers have BitCurator; ten of these use it as a native install, while two run the virtual machine. Windows and Linux are most often run on separate machines; though twelve devices, two of which are FREDs, are partitioned.

In terms of software, the most common tools are BitCurator and AccessData's Forensic Toolkit (FTK), both of which include a suite of digital forensics tools for analysis. Other tools being used include Iso(buster), ClamAV, and Brunnhilde.[18]

---

[17] "Bitcurator 1.8.0 Booting to Terminal" (thread), BitCurator Users Google Group, https://groups.google.com/forum/#!searchin/bitcurator-users/partition%7Csort:date/bitcurator-users/jzEeZMNz10E/sUmQD1sFAwAJ.

[18] For more on Brunnhilde, see: Tim Walsh, "Data-Driven Reporting and Processing of Digital Archives with Brunnhilde," *Practical Technology for Archives*, no. 8 (July 2017), https://practicaltechnologyforarchives.org/issue8_walsh/.

The most common peripheral devices currently being employed are various floppy disk drives, including the KryoFlux. Other devices employed to a lesser degree include Tableau write blockers, Universal Serial Bus (USB) drives, Zip disk drives, and optical media drives. Two repositories reported using memory card readers.
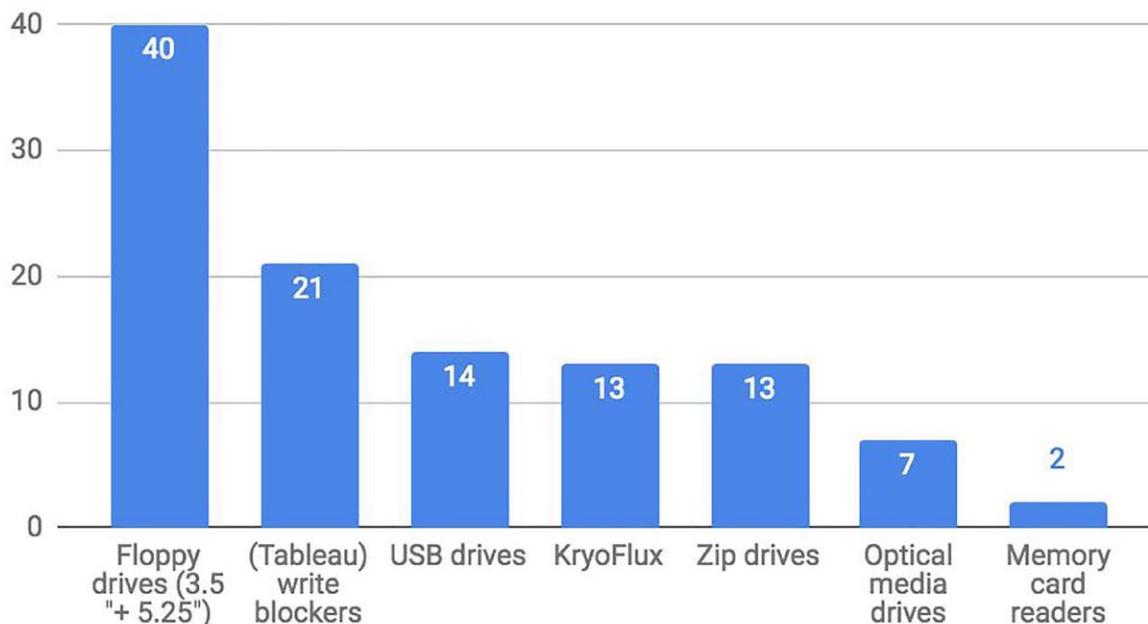
Peripheral Devices

Figure 3. Peripheral Devices

One additional finding from the information participants provided is that at least eight of the respondents have or are in the process of developing dedicated labs for accessioning and processing born-digital materials. Currently, all of our equipment is stored in a cubicle adjacent to those of processing staff. This cubicle also doubles as a processing area for student workers. As the volume of born-digital materials and the amount of time spent processing these records increases, the team acknowledges the need for additional space and is interested in exploring the lab model as a potential next step.

Moreover, because the survey was initially intended to gather information from institutions similar to Princeton, the sample is heavily skewed toward large, relatively well-funded academic repositories in the United States,[19] and the responses reflect this biased sample. In an effort to move beyond these limitations, the authors encourage archivists at other repositories to add their data to an open-access version of the data-collection spreadsheet. By opening up this resource to future collaborative data collection, the authors hope to support continued dialog about digital processing tools broadly.

---

[19] There was one Canadian institution in the sample. Otherwise, all institutions surveyed were located in the United States.

To FRED or Not to FRED? Choosing a Tool and Making Our Case

Our findings show that archival institutions are employing varying combinations of hardware, software, and peripheral devices to manage their born-digital content. The workstations that archivists assemble are dependent on a variety of available resources, including money, infrastructure (physical and technical), and staff. Based on how respondents reported their current setups, most institutions, regardless of configuration, indicated that they were able to achieve desired outcomes with their existing workstations. Most identified local storage space and processing power as overall limiting factors.

As an added measure before making our decision, we decided to compare the specifications of one of the FRED models with a comparable Dell computer assembled to match or exceed the FRED's specifications (see appendix A). We concluded that either option had the potential to meet our needs at a similar cost.

Beside the technological and cost comparisons between purchasing a FRED and Dell computer, we took into account information that survey participants shared regarding their level of satisfaction with their workstations. Survey respondents who opted to build their own workstations prefer their modularity, which allows archivists to easily replace portions of their workstation as needed. Proponents of the FRED noted that Digital Intelligence's technical support was reliable, and highlighted the fact that users can replace internal drives with those of higher capacities in order to scale up. Others voiced concerns about the longevity of FREDs as well as their inability to switch back and forth between dual booting operating systems without restarting the machine.

Among other issues, members of the team also considered the fact that the company that manufactures FRED machines, Digital Intelligence, has a primary customer base not of cultural heritage institutions but law enforcement agencies. The sobering fact that our purchase would indirectly help support the tools of the criminal justice system, and by extension, the prison industrial complex, caused some unease.

We ultimately decided that the FRED made the most sense for us based on our practical needs and institutional context. As a repository that manages literary and personal papers of contemporary authors and other cultural figures, we took into account our need to image high-capacity carriers like internal hard drives, which we have begun to encounter much more frequently in new acquisitions. The FRED's specifications, connectors, and ports would enable us to image drives in order to preserve the context of an author's working computer with more ease than building out this capability with a custom computer. Moreover, processing staff were already relatively familiar with the FRED due to our regular use of the FRED that University Archives had acquired in 2014.

We assumed that our administration's familiarity with the FRED would lead to the smooth approval of our purchase request. However, initial discussions with administrative staff revealed a lack of understanding about processing staff's shifting responsibilities in terms of managing born-digital records. Staff were met with the assumption that the digital archivist for University

Archives should be responsible for processing all digital content and making decisions about equipment purchases. Elsewhere, we have discussed the problems inherent in the "lone digital archivist model," noting how it overburdens individuals in digital archivist roles, while also limiting the growth of "traditional" processing archivists by making their digital labor invisible.[20] While archivists are functionally shifting into more technical roles, institutional structures seem slow to acknowledge this change. Moreover, in our collective experience as an all-female team, we faced initial challenges due to perceived gendered divisions of labor. Even though the library and archives profession is heavily female, men tend to dominate library information technology (IT) positions.[21] When these gender-coded realities are not acknowledged or challenged, undue and sometimes stubborn expectations can be placed on those who are expected to do digital work and those who are not.

In light of the questions raised by the administrative staff, the processing team decided to present our request in a formal proposal that would demonstrate the research and reasoning behind it. The proposal summarized much of the information presented above, including our survey findings, and laid out our recommendations for a FRED with a secure connection to the preservation storage environment. The proposal was submitted in December 2016 and the FRED purchase was approved the following month.

Implementation and Moving Forward

The Manuscripts Division now uses the FRED in a native BitCurator environment as our primary workstation, along with the KryoFlux and a growing collection of external drives. This streamlined setup has vastly improved our born-digital processing workflow and meets our operational needs. Since the new workstation has been up and running, the Manuscripts Division's 2017 summer fellow, Kathryn Antonelli, has used it to successfully image additional floppy disks from the Toni Morrison Papers, and other staff have processed and transferred large quantities of digitized audiovisual files to preservation storage. Archivists can now process born-digital materials at scale without traveling to a separate building. As a result, processing of paper-based and born-digital materials can more easily happen in parallel for the hybrid collections that are increasingly frequent within the Manuscripts Division's contemporary holdings.

We anticipated the technical support provided by Digital Intelligence would make the installation process easier for our already overburdened colleagues in IT. However, the journey between receiving the FRED and having a fully operational workstation took several months of back-and-forth discussions with IT colleagues, troubleshooting with the assistance of Digital Intelligence staff, and searching for answers in the BitCurator Users Google Group.[22] The FRED arrived at

---

[20] See: Elvia Arroyo-Ramirez, Kelly Bolding, and Faith Charlton, "Moving Beyond the Lone Digital Archivist Model through Collaboration and Living Documentation," presentation at 2017 Code4Lib, UCLA, March 9, 2017: https://blogs.princeton.edu/techsvs/2017/03/28/moving-beyond-the-lone-digital-archivist-model-through-collaboration-and-living-documentation/.

[21] Eleta Exline, "Gender Composition and Salary Gaps in Association of Research Libraries (ARL) Institutions" (masters' thesis, University of New Hampshire, 2014), https://scholars.unh.edu/cgi/viewcontent.cgi?article=1002&context=thesis.

[22] BitCurator Users Google Group, https://groups.google.com/forum/#!forum/bitcurator-users.

the beginning of May 2017, but it was not fully operational until the end of July. The issue that slowed the process down the most involved a problem with installing BitCurator version 1.8 wherein BitCurator was booting to the terminal rather than as an environment.[23] Due to the fact that our problem did not fall neatly under the umbrella of Digital Intelligence and their out-of-the-box FRED setup or the BitCurator software development community, we were not able to rely entirely on either resource for a solution.

While our proposal originally specified that we would partition the FRED in order to dual-boot Windows and Linux (BitCurator), we ultimately decided to remove the Windows partition and run the FRED solely using Linux. This was due to the issues institutions reported with their partitioned FREDs on the BitCurator Users Google Group, as well as problems we were having running both Windows and Linux on our laptop. While we had been using the KryoFlux on the Windows partition of the laptop, we decided that running the FRED solely using Linux would be feasible if we could successfully install the KryoFlux software on the FRED's Linux operating system, which we ultimately did.[24] Given the amount of time it took to set up the FRED with a single Linux operating system, we are glad that we did not attempt to partition the device.

While the current FRED workstation meets our existing use cases, staff are considering other devices and software for future needs. In February 2018 we acquired a Mac laptop, which will allow us to process acquisitions that include files created in this environment.[25] Other additions the team will likely need include more workstations and physical work space to address capacity issues and a dedicated quarantine machine for virus scanning. Additional software, such as Archivematica, that would streamline and automate many processes would also likely help with scalability. Staff are also working with archivists from University Archives and colleagues in our systems department to implement a digital repository to support long-term preservation. Currently, preservation and access files are stored together on a secure server.

Now that staff are able to process born-digital records, we have been concentrating on providing access to this content, and like many other institutions, trying to address the myriad issues that doing so entails. As the topic of providing access to born-digital and digitized materials could warrant its own article entirely, we will briefly mention our current situation. Access to much of our digital content must be restricted to the reading room due to copyright and privacy concerns. At the moment, the reading room is equipped with three non-networked machines that have born-digital and digitized content stored on their hard drives. We are currently working with our systems department to implement a local network to provide site-limited access to these files. Rather than managing duplicate copies of digital content on each computer's hard drive, this solution would involve using a host server to add and disseminate on-demand access to requested

---

[23] This issue, along with several strategies for bypassing it, is documented in two threads in the BitCurator Users Google Group: "Bitcurator 1.8.0 Booting to Terminal" (thread), BitCurator Users Google Group, https://groups.google.com/forum/#!topic/bitcurator-users/jzEeZMNz10E; "BitCurator v. 1.8.12 Still Booting to Terminal" (thread), https://groups.google.com/forum/#!topic/bitcurator-users/i9Vk6P9u1Gc.

[24] A more detailed description of this process, which was also not without hiccups, can be found on our blog: Kelly Bolding, "Digital Archives Workstation Update: KryoFlux, FRED, and BitCurator Walk into a Bar . . ." *This Side of Metadata* (blog), Princeton University Library, October 3, 2017, https://blogs.princeton.edu/techsvs/2017/10/03/digital-archives-workstation-update-kryoflux-fred-and-bitcurator-walk-into-a-bar/.

[25] A 2012 Macbook Pro with 8 GB memory, a 500 GB SATA Drive, and a 2.5 GHz Intel Core i5 processor.

files on dedicated reading room laptops. Ideally, staff would like to implement a virtual reading room to promote more user-friendly access to semi-restricted digital content, but in the meantime, having a local networked solution to allow reading room access will be a significant improvement over the current setup.

Conclusion

Assembling a digital processing workstation is not a simple and straightforward undertaking. As the Princeton University Library Manuscripts Division processing team discovered, the process necessitated a significant amount of research and consideration. Conducting a survey of digital archives workstations at twenty peer institutions enabled us to make a data-driven decision to acquire a workstation that met our particular needs as a manuscript repository and make a convincing argument to library stakeholders in favor of purchasing a FRED. The process also allowed us to build relationships with other archivists working with born-digital materials and develop our own knowledge of hardware and software specifications, which in turn benefited our hands-on digital processing work. We hope to extend the usefulness of this project by opening up the resulting spreadsheet as a communal resource.[26]

While the FRED made sense for our needs, we acknowledge that our decision to purchase such an expensive tool was based on our institutional context where infrastructural challenges are often greater than financial ones. Smaller or under-resourced institutions, including institutions with more limited budgets, have been involved in groundbreaking work to build practical and affordable workstations. The Digital POWRR (Preserving Digital Objects With Restricted Resources)[27] project has fostered inclusive communities of practice by investigating and recommending digital preservation strategies for smaller and under-resourced institutions, providing a host of openly available training resources on their website, and running professional development institutes that support archivists and digital preservation specialists from these institutions. To cite an instructor at the November–December 2017 institute who was frustrated when an attendee revealed they had been told that digital processing was unfeasible without a FRED, "For the giant majority of digital-rescue work at a typical library or archives: You don't need a FRED. You don't need a Kryoflux. There's cheaper equipment that works."[28] We look forward to seeing these voices more prominently represented in future iterations of this collaborative workstation documentation. The management of digital cultural heritage is not solely the charge of large academic institutions, and conversations about this work should not be limited to this privileged subset of the field.

By opening up the spreadsheet in an effort to solicit contributions from the broader archival and digital preservation community, we hope to promote a continued conversation about evolving tools and strategies employed by archivists in a variety of institutional settings. We respect and

---

[26] "Digital Processing Workstation Specs (Google sheet),"
https://drive.google.com/open?id=1jcsXOymePNRVXhkgffQta4jRfDZPzDxZwxouqeL4M2k.
[27] "POWRR" (project website)," Preserving Digital Objects With Restricted Resources (Digital POWRR), http://digitalpowrr.niu.edu/.
[28] Dorothea Salo (@LibSkrat), "For the giant majority of digital-rescue work at a typical library or archives: You don't need a FRED . . . ," Twitter, December 1, 2017, 10:52 a.m., https://twitter.com/libskrat/status/936669258100215808.

are grateful to archivists who are engaged in developing new workflows, digital or otherwise, and who dedicate time to assist other archivists by sharing their strategies. We hope that centralizing some of this core information will make it easier to share up-to-date workstation specifications. While a shared Google Sheet is a crude but functional place to start, future work could link institutional specifications to collaborative software tool registries such as COPTR or incorporate workstation research into other projects currently underway to map digital processing workflows, such as OSSArcFlow, which is exploring the integration of open-source software tools.[29] Open documentation of workstation components could also support tool-sharing collaborations among regional consortiums. Distributing the burden of purchasing expensive or niche workstation components across a collective of institutions could be another way for archivists to support one another across institutional boundaries. We look forward to seeing how this spreadsheet develops and to continuing to work with our colleagues across the field to iteratively develop and document practical workstations for digital processing.

---

[29] "OSSArcFlow" (project website), Educopia Institute, https://educopia.org/research/ossarcflow .

Appendix A. Dell and FRED Comparison

	Dell Precision Tower 7910	FRED DX Single RAID
Processor	Dual Intel® Xeon® Processor E5-2640 v4 (10C, 2.4GHz, 3.4GHz Turbo, 2133MHz, 25MB, 90W)	Dual(2) Intel Xeon E5-2620 v4 CPU, (8 Core) 2.1 GHz, 20MB Cache, 8.0 GT/s Intel QP
Memory	64GB (16x4GB) 2400MHz DDR4 RDIMM ECC	32 GB (2x16GB) PC4-17000 DDR4 2133 MHz ECC Memory
Drives	2.5" 256GB SATA Class 20 Solid State Drive	1 x 256 GB Solid State SATA III Drive - OS Drive
	2TB 3.5" Serial-ATA (7,200 RPM) Hard Drive	1 x 256 GB Solid State SATA III Hard Drive - Temp/Cache/DB Drive
	2TB 3.5" Serial-ATA (7,200 RPM) Hard Drive	1 x 2.0 TB 7200 RPM SATA III Hard Drive - Data Drive installed in HotSwap Bay 1
RAID	RAID 1	RAID Set #1 2 TB 7200 RPM SATA III Hard Drive Set (Set of 5 Drives)
Accessories	IDE/SATA/Firewire write blockers, media card reader, external cd/dvd drive	None needed
Price	$6,640.45 (computer) + accessories (~$1800) = $8440.45	$10,549.00

Bibliography

Arroyo-Ramirez, Elvia, Kelly Bolding, and Faith Charlton. "Moving Beyond the Lone Digital Archivist Model through Collaboration and Living Documentation." Google Slides presentation at Code4Lib Conference, UCLA, Los Angeles, CA, March 9, 2017. https://blogs.princeton.edu/techsvs/2017/03/28/moving-beyond-the-lone-digital-archivist-model-through-collaboration-and-living-documentation/.

Barrera-Gomez, Julianna, and Ricky Erway. *Walk This Way: Detailed Steps for Transferring Born-Digital Content from Media You Can Read In-house*. Dublin, OH: OCLC Research, 2013. http://www.oclc.org/content/dam/research/publications/library/2013/2013-02.pdf.

BitCurator Users Google Group. https://groups.google.com/forum/#!forum/bitcurator-users.

Bolding, Kelly. "Digital Archives Workstation Update: KryoFlux, FRED, and BitCurator Walk into a Bar . . ." *This Side of Metadata* (blog). Princeton University Library, October 3, 2017. https://blogs.princeton.edu/techsvs/2017/10/03/digital-archives-workstation-update-kryoflux-fred-and-bitcurator-walk-into-a-bar/.

Bolding, Kelly. "Tooling Up: Building a Digital Processing Workstation." *This Side of Metadata* (blog), Princeton University Library, July 19, 2016. https://blogs.princeton.edu/techsvs/2016/07/19/tooling-up-building-a-digital-processing-workstation/.

"BDAX2016." Google Drive Folder containing documents from the Born Digital Archiving & eXchange, held at Stanford University, Stanford, CA, July 18–20, 2016. https://drive.google.com/open?id=0By4GpQlPYswAbDJHLUsxcXZPOUk.

Charlton, Faith, Kelly Bolding, and Chloe Pfendler. "Jump in 3 Report." Society of American Archivists, 2015. https://www2.archivists.org/groups/manuscript-repositories-section/jump-in-3-2015-results.

"Community Owned digital Preservation Tool Registry (COPTR)." Last modified November 1, 2018. http://coptr.digipres.org/.

Digital POWRR. "About POWRR." Last modified August 22, 2017. http://digitalpowrr.niu.edu/.

Educopia Institute. "OSSArcFlow." Last modified 2018. https://educopia.org/research/ossarcflow.

Erway, Ricky. *You've Got to Walk Before You Can Run: First Steps for Managing Born-Digital Content Received on Physical Media*. Dublin, OH: OCLC Research, 2012. http://www.oclc.org/research/publications/library/2012/2012-06.pdf.

Exline, Eleta. "Gender Composition and Salary Gaps in Association of Research Libraries (ARL) Institutions." Masters' thesis, University of New Hampshire, 2014. https://scholars.unh.edu/thesis/3/.

Gelman, Juan. Papers. Princeton University Library. https://findingaids.princeton.edu/collections/C1511.

Goldman, Ben. "Outfitting a Born-Digital Archives Program." *Practical Technology for Archives*, no. 2 (2014). https://practicaltechnologyforarchives.org/issue2_goldman/.

Harvard Library Digital Forensics Working Group. "At Home with Digital Forensics: A Conversation about Harvard Workflows." Google Slides, 2016. https://docs.google.com/presentation/d/1sttFttVdSCjZK05YMQWm9FsxlQLmIF4CN5CKf2ma0UY.

Morrison, Toni. Papers. Princeton University Library. https://findingaids.princeton.edu/collections/C1491.

Olsen, Porter. "Building a Digital Curation Workstation with BitCurator (update)." *BitCurator* (blog). August 2, 2013. https://www.bitcurator.net/building-a-digital-curation-workstation-with-bitcurator-update/.

Olsen, Porter. "Digital Curation Workstation." *Maryland Institute for Technology in the Humanities* (blog). University of Maryland, November 26, 2012. http://mith.umd.edu/digital-curation-workstation/.

Prael, Alice, and Amy Wickner, "Getting to Know FRED: Introducing Workflows for Born-Digital Content." *Practical Technology for Archives*, no. 4 (2015). https://practicaltechnologyforarchives.org/issue4_prael_wickner/.

Walsh, Tim. "Data-Driven Reporting and Processing of Digital Archives with Brunnhilde." *Practical Technology for Archives*, no. 8 (July 2017). https://practicaltechnologyforarchives.org/issue8_walsh/.

Weintraub, Jennifer. "Bibliography on Digital Forensics (from the Harvard Library Digital Forensics Working Group)." Last modified January 12, 2017. https://wiki.harvard.edu/confluence/pages/viewpage.action?pageId=213390414.