

SOLOMON R. GUGGENHEIM FOUNDATION

Electronic Records Management Start-Up Project Generously funded by the NHPRC

Electronic Records Processing and Obsolete Media Ingestion Workstation

August 25, 2014 – Prepared by Anthony Cocciolo, Electronic Records Consultant

Objective: To plan for and install an electronic records processing and obsolete media ingestion workstation that can double as an archive viewing workstation.

Problem Statement:

The research and inventory project conducted from September through December of 2013 revealed removable media, such as CDs, DVDs, and floppy disks, are held both in the archives as well as by select staff in their work areas. Content with archival value contained on removable media will need ingested into the electronic records repository by first imaging the media per best practices in born-digital archiving.¹

Analysis of Environment for Solution:

The most common solution to imaging removable media is to create a workstation dedicated to this task and purchasing the requisite drives for reading the media.² Efforts are underway to create shared workstations that can be used by multiple institutions such as through Metropolitan New York Library Council. However, since these services are not yet available, and the cost for establishing such a workstation is not prohibitive, SRGF will establish its own workstation for this purpose.

Solution:

Hardware

The consultant and Director of Library and Archives decided that the electronic records processing workstation should be a laptop that could be moved between various SRGF locations. The machine purchased is described below:

HP EliteBook 840 G1 - 14" - Core i5 4200U - Windows 7 Pro 64-bit / 8 Pro do - 4 GB RAM
Cost: \$936.03

Working with Alejandra Montoya-Boyer (Systems Analyst from IT Department) and Paul McAlpine (Support Analyst from the IT Department), the following hardware was also purchased:

- IBM External 3.5 floppy drive with USB connector, approx. cost \$12.00
- 5.25 Floppy drive controller with connection to USB port from Device Side Data³, cost \$56.00
- 5.25 Floppy drive, purchased used on EBay, approx. cost \$50 (TEAC FD55GFR)
- 5.25 Floppy drive power supply,⁴ cost \$25.00
- Samsung External Blu-Ray player with backwards support for CD and DVD connectable to laptop, cost: \$97.00 (Samsung SE-506BB - BDXL drive - Hi-Speed USB)
- Iomega 2GB Jaz Drive with Oakwood SCSI to USB Adapter, cost \$66.00 for drive and \$179.00 for adapter

An external 750 MB Iomega Zip drive was found in Guggenheim storage and was also connected to the laptop. The workstation with select devices setup is shown in Figure 1.

¹ <http://oclc.org/content/dam/research/publications/library/2013/2013-02.pdf>

² <http://www.bitcurator.net/2013/08/02/building-a-digital-curation-workstation-with-bitcurator-update/>

³ <http://shop.deviceside.com/prod/FC5025>

⁴ <http://shop.deviceside.com/prod/PSU5>



Figure 1. Obsolete media workstation with select drives connected.

Best-practice in imaging media is using a write-blocker so that the new information is not inadvertently written to the removable media. The use of a hardware write-blocker was explored, but ultimately decided against. Specifically, the use of the WiebeTech USB Writeblocker was tested, with the hope being that all the USB devices could be connected to it (Iomega Zip drive, Iomega Jaz Drive, 3.5 floppy) and it would provide write protection (cost: \$195, see Figure 8 in the appendix). Unfortunately, only devices that register as “USB Mass Storage Device,” such as USB thumb drives, would be read by Windows when connected to the write blocker, and other devices such as the Zip, Jaz or 3.5 floppy drives did not (Figure 2). This product would be useful for USB thumb drives, however, during the inventory stage of this project there was no evidence that USB thumb drives were being used for storing content with permanent value.



Figure 2. Jaz media found in the archives

For 3.5 floppy drives and 5.25 floppy drives, the recommendation being made is that the write protection be enabled on the disk (opening the write-protection tab). Also, the controller for the 5.25 drive only provides for read ability; thus, the chances of a disk getting written to are very low. The other drives do not have write protection tabs that can be enabled. Testing of Zip drives indicates that in Windows 7 they are by default already write protected because special software is needed to make them writeable. Testing of the Iomega Jaz drives indicates that they are writeable, thus, caution will need to be exerted with this media. However, because of the

complex setup of this media (use of SCSI to USB converter), it is unlikely that a hardware write blocker could be found that would work with this medium. BluRays, CDs and DVDs are by default read-only, so no write blocker is needed with them.

Software

Software installed on the workstation includes:

Default Software installed by IT:

- Windows 7
- Adobe CreativeSuite (InDesign, Illustrator, Photoshop, Acrobat)
- Microsoft Office 2010 (used by scripts to convert office documents to preservation and access formats)
- Kapersky Endpoint Security v10 (institution-wide virus scanning software)
- Firefox web browser

Software for Accessing files particular to SRGF:

- DWG TrueView (free tool from AutoDesk for viewing AutoCAD drawings)⁵
- MBox Viewer (for reading MBOX files, which is preservation format for Email boxes)⁶
- 7-Zip (free, open-source software for reading 7Zip files, which is used to compress Archivematica AIPs)⁷

Digital Forensics Tools:

- VirtualBox⁸ (for running BitCurator⁹, with BitCurator VBOX image also installed)
- FTK Imager (free software for creating disk images)¹⁰
- OSFMount (free software for reading disk images)¹¹
- TRID and TRIDNet (useful for identifying file extensions, free software)¹²
- "Disk Image and Browse," software for reading 5.25 disks (provided by vendor of FC5025 controller)
- Guggenheim Verlag Fonts and fonts formerly in use (Guggenheim HTF Book and Guggenheim HTF Bold)
- Quick View Plus Standard Edition (allows expedient viewing/appraisal of computer files, costs \$49.99)¹³
- Bagger Software (GUI for BagIt software provided by Library of Congress)¹⁴

Web Archiving Tools

- WAIL (Web Archiving Integration Layer)¹⁵ – For running Heritrix and WayBackMachine locally
- GrabThemAll (Firefox extension for creating screenshots)¹⁶
- createURLs.bat (Created by Consultant for generating a list of URLs to be captured. Uses the open-source program WGet and Grep).

General System Utilities:

- WinSCP (free software for copying files to Archivematica server)¹⁷
- VLC Player¹⁸ (video player) and Handbrake¹⁹ (video encoder) (for handling submissions including Video)
- MacDrive (for reading Mac formatted disks, cost: \$49.99)²⁰

⁵ <http://usa.autodesk.com/adsk/servlet/pc/index?id=6703438&siteID=123112>

⁶ <http://sourceforge.net/projects/mbox-viewer/>

⁷ <http://www.7-zip.org/>

⁸ <https://www.virtualbox.org/>

⁹ <http://www.bitcurator.net/>

¹⁰ <http://www.accessdata.com/support/product-downloads>

¹¹ <http://www.osforensics.com/tools/mount-disk-images.html>

¹² <http://mark0.net/soft-trid-e.html>

¹³ <http://www.avantstar.com/metro/home/Products/QuickViewPlusStandardEdition/Info/HowToBuy>

¹⁴ <http://sourceforge.net/projects/loc-xferutils/files/loc-bagger/>

¹⁵ <http://matkelly.com/wail/>

¹⁶ <https://addons.mozilla.org/en-US/firefox/addon/grab-them-all/>

¹⁷ <http://winscp.net/eng/index.php>

¹⁸ <http://www.videolan.org/vlc/index.html>

¹⁹ <http://handbrake.fr/>

²⁰ <http://www.mediafour.com/products/macdrive>

Not all software programs used in SRGF's files were installed to the laptop. Because of the cost of VectorWorks (approx. \$1,800 USD), it was cost-prohibitive to acquire a license for the workstation. The consultant recommends its purchase when additional funding is possible. VectorWorks would be useful for converting VectorWorks files into AutoCAD files for long-term preservation; however, a Guggenheim computer with this software already installed (e.g., in Exhibition Management) could potentially be used in a limited-funding scenario.

Rhino is required for viewing 3D models, and 2-dimensional representations (e.g., PDF) are diminished representations. Thus, the consultant recommends that this software be purchased when additional funds are available, thus allowing 3D models to be both preserved and accessible to researchers/staff. A Rhino software license costs approximately 1,000 USD.

Additional fonts may need to be installed on the workstation to render some documents and convert them to PDF. These font collections are held on various network shares, and will be installed on an as-needed basis.

Conclusion

The total approximate cost for the hardware and software to establish the obsolete media laptop is \$1,600.00.

Appendix: Images of drives connected to Obsolete Media laptop



Figure 3. 3.5 TEAC floppy with USB controller from Device Side Data



Figure 4. Iomega 750 MB Zip disk



Figure 5. Iomega 2GB Jaz Drive with SCSI to USB connector



Figure 6. Samsung USB BluRay player with backwards support for DVD and CD

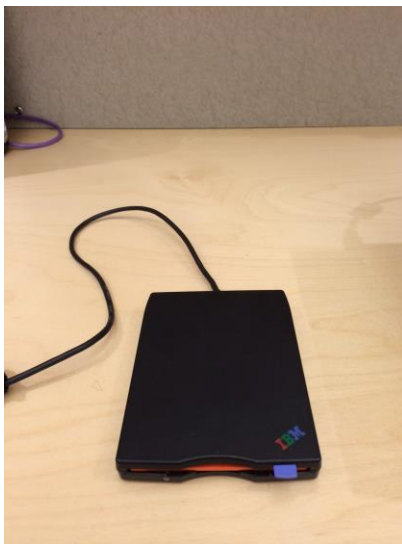


Figure 7. IBM 3.5 Floppy drive



Figure 8. WiebeTech USB WriteBlocker with SanDisk Thumbdrive connected. This device was tested, but not included in the obsolete media setup.