ANYTHING IS BETTER THAN NOTHING
MINIMUM VIABLE ACTIONS FOR BORN DIGITAL ACCESSIONING

1. PROCESSING WORKSTATION
Emily Vigor

2. APPRAISAL
Kate Dundon

3. CAPTURE & PRESERVE
Steve Duckworth

4. STABILIZE
Steve Kutay

5. ACTIVITY

6. Q&A
PROCESSING WORKSTATION

What do I need?
MINIMUM REQUIREMENTS

Computer

Storage

Open source software
Grant funded
Supplies under $2000
IT Support :(
SUPPLIES & MATERIALS

COMPUTER [$1000]
Requires an internal optical media drive to read CDs and DVDs, as well as several USB ports to connect disc drives to

EXTERNAL HARD DRIVE(S) [$200]
1TB (minimum) hard drives to be formatted for Mac and PC

OPEN SOURCE SOFTWARE [$0]
Make sure you have the IT support or training to help guide installation

FLOPPY DISK CONTROLLER CARD & DISK DRIVE [$200]
Helps newer machines read media connected via obsolete disk drives

WRITE-BLOCKER [$300]
Protects legacy media from being overwritten once connected to a computer for imaging

MEMORY CARD READER [$25]
Supports reading flash drive formats such as SD, CompactFlash and MicroSD

Ben Goldman post:
https://practicaltechnologyforarchives.org/issue2_goldman/
OTHER THINGS TO CONSIDER

Access to IT

Storage

Priorities
NO MONEY? NO PROBLEM!

Open Source Tools
- ePADD
- Archivelt
- California Revealed

Research and Training
- Literature review
- Webinars
- Build internal documents

Raise $$$
- Grants
- Crowdfunding campaigns

Collaborate
- Put a call out on Listservs
- Check your own institution
God grades
pass/fail.

APPRAISAL

Kate Dundon
UC Santa Cruz Special Collections & Archives
PRE-ACQUISITION APPRAISAL

✓ Collect born digital intentionally
✓ Consider your technical capacity
✓ Exclude out of scope content
PRE-CAPTURE APPRAISAL: ITERATIVE WEEDING

Duplicate or out of scope software
blank disks
backups
INTRA-CAPTURE APPRAISAL

Use Data Accessioner to exclude unwanted content from your transfer
POST-TRANSFER APPRAISAL

Manually identify and delete PII and clutches of non-collection content

Use DROID csv export and Excel to identify duplicate files
APPRAISAL: TAKEAWAYS

APPRAISAL IS ITERATIVE
Take many passes

BUILD A KNOWLEDGE BASE
Maintain continuity by processing both paper and digital

DOCUMENTATION IS ESSENTIAL
Record pre and post-processing extents

RESPECT YOUR RESOURCES
It is resource-intensive to steward digital content
CAPTURE & PRESERVE

Accessible control and preservation

Steve Duckworth, @archivesteve
Oregon Health & Science University
ACCESSIONING: real talk

Moving/copying digital records can be scary

What if I mess up the metadata and ruin everything forever?

Leaving everything on deteriorating media is scarier

Move it or lose it.
INGESTING RECORDS with DATA ACCESSIONER

◇ migrate data
◇ capture metadata
◇ create preservation data (checksums)

◇ assists with description
INGESTING RECORDS with DATA ACCESSIONER

❖ fill in your metadata
❖ “Migrate” does all the work for you
BONUS TALK:
DATA ACCESSIONER METADATA TRANSFORMER

◮ use the XML file from the migration to create more useful documents
**BONUS TALK:**

**DATA ACCESSIONER METADATA TRANSFORMER**

◇ CSV file can help with quick descriptions

<table>
<thead>
<tr>
<th>directory path</th>
<th>file name</th>
<th>file extension</th>
<th>last modified</th>
<th>size [bytes]</th>
<th>md5</th>
<th>file format</th>
<th>format version</th>
<th>format registry key/PUID</th>
<th>Identified By</th>
<th>MIME Type</th>
<th>ID Warning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pictures/cats/</td>
<td>cool-cat-with-sunglasses-768x410.jpg</td>
<td>jpg</td>
<td>2019-02-20T</td>
<td>37118</td>
<td>b71ce62288b4ba00b99:fb08619e831c</td>
<td>JPEG File Interchange Format</td>
<td>1.01</td>
<td>fmt/43</td>
<td>Droid v6.1.5</td>
<td>image/jpeg</td>
<td></td>
</tr>
<tr>
<td>Pictures/cats/</td>
<td>moustache-cat(1).jpg</td>
<td>jpg</td>
<td>2019-02-20T</td>
<td>67519</td>
<td>8837e8508cc0a3a352042a171ec17c7f</td>
<td>JPEG EXIF</td>
<td>2.2.1</td>
<td>fmt/645</td>
<td>Droid v6.1.5</td>
<td>image/jpeg</td>
<td></td>
</tr>
<tr>
<td>Pictures/cats/</td>
<td>ifS002es28b4576f13409f16e75445a44---shark-hat-the-shark.jpg</td>
<td>jpg</td>
<td>2019-02-20T</td>
<td>8285</td>
<td>ae9d904085d1380a0736029e8446511a</td>
<td>JPEG File Interchange Format</td>
<td>1.01</td>
<td>fmt/43</td>
<td>Droid v6.1.5</td>
<td>image/jpeg</td>
<td></td>
</tr>
<tr>
<td>Pictures/cats/</td>
<td>5bad360c26000036008038a2.jpg</td>
<td>jpg</td>
<td>2019-02-20T</td>
<td>56176</td>
<td>b808cb7451d86bdf5:7f0f159c92</td>
<td>JPEG File Interchange Format</td>
<td>1.01</td>
<td>fmt/43</td>
<td>Droid v6.1.5</td>
<td>image/jpeg</td>
<td></td>
</tr>
<tr>
<td>Pictures/idris_elba/</td>
<td>idris_elba-4580_(cropped).jpg</td>
<td>jpg</td>
<td>2019-02-20T</td>
<td>522043</td>
<td>6f369b6b04a3523a2b06e5c73e2c4d591e</td>
<td>JPEG File Interchange Format</td>
<td>1.01</td>
<td>fmt/43</td>
<td>Droid v6.1.5</td>
<td>image/jpeg</td>
<td></td>
</tr>
<tr>
<td>Pictures/idris_elba/</td>
<td>Unknown.jpg</td>
<td>jpg</td>
<td>2019-02-20T</td>
<td>4601</td>
<td>e4b93c66623c54033cd6ca8935549b66</td>
<td>JPEG File Interchange Format</td>
<td>1.01</td>
<td>fmt/43</td>
<td>Droid v6.1.5</td>
<td>image/jpeg</td>
<td></td>
</tr>
<tr>
<td>Pictures/idris_elba/</td>
<td>Unknown-1.jpg</td>
<td>jpg</td>
<td>2019-02-20T</td>
<td>6239</td>
<td>ea79b8ad7d379d7a03ba73bfa5953a0b</td>
<td>JPEG File Interchange Format</td>
<td>1.01</td>
<td>fmt/43</td>
<td>Droid v6.1.5</td>
<td>image/jpeg</td>
<td></td>
</tr>
<tr>
<td>Pictures/</td>
<td>.DS_Store</td>
<td>DS_Store</td>
<td>2019-02-20T</td>
<td>6148</td>
<td>c336b888c5835e19e671bf7f7f40988b</td>
<td>DS_store file (MAC)</td>
<td>fmt/394</td>
<td>Droid v6.1.5</td>
<td>application/octet-stream</td>
<td>DROID Identification Warning</td>
<td></td>
</tr>
<tr>
<td>Pictures/hats/</td>
<td>fox.png</td>
<td>jpg</td>
<td>2019-02-20T</td>
<td>129022</td>
<td>3ff51c054376fc93257df43eb6e3720</td>
<td>JPEG File Interchange Format</td>
<td>1.01</td>
<td>Exiftool v0.2</td>
<td>image/jpeg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pictures/hats/</td>
<td>article-2336777-1A2D3088000005DC-195_b34d08B.jpg</td>
<td>jpg</td>
<td>2019-02-20T</td>
<td>96785</td>
<td>a944fd9db63a3fab1d9e0d8288b23e8</td>
<td>JPEG File Interchange Format</td>
<td>1.01</td>
<td>Exiftool v0.2</td>
<td>image/jpeg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pictures/hats/</td>
<td>brains.png</td>
<td>jpg</td>
<td>2019-02-20T</td>
<td>137047</td>
<td>74a2596e52f2ab03b51753180228337</td>
<td>JPEG File Interchange Format</td>
<td>1.01</td>
<td>Exiftool v0.2</td>
<td>image/jpeg</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
PRESERVATION with BAGGER

◇ group together your data (records) and metadata (description)
◇ like a digital archival box
◇ not compressed
PRESERVATION with BAGGER

◇ more metadata
PRESERVATION with BAGGER

◇ save that bag!
structure of the process
ACCESSIONING: KEY TAKEAWAYS

DATA ACCESSIONER
transfer records without messing them up
capture metadata

DATA ACCESSIONER METADATA TRANSFORMER
reuse already captured metadata
less work for you

BAGGER
group files and metadata together for preservation and portability
Migrating from storage media

Steve Kutay, California State University, Northridge
BARRIERS TO ACCESS
A REVIEW

RELIABLE

LESS RISK

RELIEABLE

OBsolete & Hardware Dependent

MORE RISK

FAILURE PRONE

ACTIVE & Hardware Independent

27
MIGRATION WORKFLOW
Creating a preservation package from source to vault

Prepare media
Describe and analyze
Export
Preserve
PART 1: PREPARE MEDIA

1. Assess drive and jumper block (if necessary)
2. Take photos
3. Document
4. Connect
5. Create destination directory
PART 2: DESCRIBE and ANALYZE

USING SOFTWARE:
1. Add contextual and administrative metadata regarding files
2. Extract technical metadata
3. Generate fixity and other analytic reports
4. OPTION: Consider a forensic disk image for further analysis and storage

FORENSIC ANALYSIS:
- Virus scans
- Checksums
- Format validations
- Hashes/de-duplication
- File inventories
- Directories/structure
- Private information
- Hidden/deleted files
- Web history
- And more
**THE FORENSIC DISK IMAGE**

*(optional copy method)*

**WHAT IS IT?**
A bit-for-bit copy (bitstream) of all files (visible and hidden), folders, and free space as a single file. A stable and mountable preservation format.

**WHEN DO I USE IT?**
Acquire images when access to hidden information in the storage medium is desired. Most useful for media containing system files and logs. *Consider donor privacy.*

**HOW DO I CREATE IT?**
Open source and proprietary software applications. Some are format-specific.
PART 3: EXPORT

EXPORT ALL AS A PACKAGE
Data packaging software (e.g. BagIt, Bagger) bundles disk image or source files with metadata, reports, and other sidecar files for transmission as part of an archival information package (AIP) of the collection.

CONFIRM/MAINTAIN FILES
- Check destination folders for accuracy
- Duplicate preservation file package for use as access files, but remove redactions*
- Write-protect your access drive, or purge and replace access package after each use*

*access policies will vary by institution
PART 4: PRESERVE

SECONDARY MIGRATION
Migrate AIP (disk image, files, metadata, reports) to long-term storage destinations.

REDUNDANCY
Protect against human/mechanical errors and catastrophic damage by storing 2-3 copies in 2 or more locations.

PHYSICAL MEDIA PLAN
Create policy a policy regarding the retention or disposition of the physical storage media once redundant preservation is achieved.
STILL WANT MORE?

Tools to consider to develop your capacity...

**BitCurator**

Grant-funded, open source virtual machine bundles apps for disk imaging, integrity checks, analysis, redaction, directory reports, deduplication, and more.

**FORENSIC WORKSTATIONS**

Assembled for forensic computing. Contains numerous write-protected ports and storage bays for acquiring, analyzing, and storing disk images. $5000 - $8000
ACTIVITY

Small group discussion
THANKS

Emily evigor@berkeley.edu
Kate dundon@ucsc.edu
Steve D duckwors@ohsu.edu
Steve K stephen.kutay@csun.edu
Anything is better than nothing: Minimum viable actions for accessioning born-digital

Resource list

Getting started

- Digital POWRR Project (Preserving digital Objects With Restricted Resources). [https://digitalpowrr.niu.edu](https://digitalpowrr.niu.edu)
- Outfitting a Born Digital Archives Program. [https://practicaltechnologyforarchives.org/issue2_goldman/](https://practicaltechnologyforarchives.org/issue2_goldman/)
- Museum of Obsolescent Media. [https://obsoletemedia.org/](https://obsoletemedia.org/)

Processing and description

- Digital Processing Framework. [https://ecommons.cornell.edu/handle/1813/57659](https://ecommons.cornell.edu/handle/1813/57659)

Workflow planning

- AIMS White Paper. [https://dcs.library.virginia.edu/aims/white-paper/](https://dcs.library.virginia.edu/aims/white-paper/)
- OSSArcFlow “as-is” workflows. [https://educopia.org/ossarcflow/](https://educopia.org/ossarcflow/)

Education

- DigiPET: A Community Built Guide for Digital Preservation Education + Training. [https://docs.google.com/document/d/16KGPT-isQ39PSFNIp3eujIWHQccAydI86uyWJmWyA7E/edit](https://docs.google.com/document/d/16KGPT-isQ39PSFNIp3eujIWHQccAydI86uyWJmWyA7E/edit)
- DLF Born Born Digital Access Bootcamp. [https://www.diglib.org/groups/born-digital-access-group/](https://www.diglib.org/groups/born-digital-access-group/)

Tools referenced in presentation

- Bagger. [https://github.com/LibraryOfCongress/bagger](https://github.com/LibraryOfCongress/bagger)
- Bitcurator. [https://bitcurator.net/](https://bitcurator.net/)
- Duke Data Accessioner. [http://dataaccessioner.org](http://dataaccessioner.org)
- Data Accessioner Metadata Transformer. [http://dataaccessioner.org/da-mt.htm](http://dataaccessioner.org/da-mt.htm)
Activity questions

1. Describe your current capacity for accessioning born digital content, focusing on the processes of appraising, capturing, and ingesting files into your repository.

2. Describe one thing you need/want to be able to do in accessioning born digital. Pick something specific, big or small.

3. What do you need to accomplish this at your institution?

4. Describe one thing that you learned today that you can realistically apply to help you achieve this goal.

Discuss your answers to these questions with your group. Consider: what is standing in your way of achieving your goal in question 2?