Personal Digital Archiving
Train-the-Trainer Workshop
July 31, 2014
Instructors:

Oscar Gittemeier, Youth Services Librarian, Atlanta-Fulton Public Library, East Atlanta Branch
Wendy Hagenmaier, Digital Collections Archivist, Georgia Tech Archives
Michelle Kirk, Records Manager, VP Corporate Records and Information Management, SunTrust Banks, Inc.
Part I:
The What and the Why of Personal Digital Archiving
Personal Digital Archiving

• What is retention?

• What is archiving?

• Why should we care about archiving our personal records?
What Qualifies as a Personal Record?

Records are things constituting pieces of evidence about the past, especially an account of an act or occurrence kept in writing or some other permanent form (*Google Dictionary*)

Personal vs. Business Records

- An organization owns all its records that are created as evidence of its business transactions
- Likewise, individuals own records of their personal business transacted, and any other records created for historical purposes
Record Identification and Inventory

• Questions to ask yourself when determining if something is a record:
  • Might I need this to substantiate a claim?
  • Is there a legal or financial/tax reason why I should preserve this?
  • Does this have intrinsic or historical value which makes me want to keep it indefinitely?

• Make a list of the records
  • Group into records and non-records
  • Determine how long to keep the records
## Types and Retention of Personal Records

<table>
<thead>
<tr>
<th>Record Type</th>
<th>Examples</th>
<th>Retention Period (how long to preserve)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial</td>
<td>Loan Payoffs, Tax Returns, Cancelled Checks, Bank Statements, Paycheck Stubs, Investment Statements, Medical Bills</td>
<td>Varies</td>
</tr>
<tr>
<td>Legal</td>
<td>Wills, Trust Documents, Marriage Licenses, Adoption Papers, Death Certificates, Deeds</td>
<td>Varies</td>
</tr>
<tr>
<td>Medical</td>
<td>Test Results</td>
<td>Indefinite</td>
</tr>
<tr>
<td>Historical</td>
<td>Photos, Videos, Audio, Scanned or “born digital” Documents</td>
<td>Indefinite</td>
</tr>
</tbody>
</table>

*See the handouts for great resources on how long different types of personal records should be retained.*
Formats of Records

- Records are increasingly being “born digital”
- Most records can be preserved digitally instead of in paper format
- It’s important to understand that unlike paper records, long term digital records need special treatment
Part II:
The Landscape of Personal Digital Records
“Instructions for Future Interaction”

“Digital objects are sets of instructions for future interaction.”
(Cal Lee, “Digital Forensics Meets the Archivist (And They Seem to Like Each Other),” Society of Georgia Archivists Annual Meeting, 2012)

• Digital records are rendered, represented, experienced

• Think of digital records as interactions at various technical and social levels:
  • interactions between hardware and software
  • interactions between software and files
  • interactions among record creator, record steward, and record user

• So, it’s important to understand the ecosystem of personal digital records
# Personal Digital Records Ecosystem

<table>
<thead>
<tr>
<th>LOCAL</th>
<th>OFFSITE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HARDWARE</strong></td>
<td><strong>SOFTWARE</strong></td>
</tr>
</tbody>
</table>
| • Personal computer  
  - Internal: CPU, RAM, motherboard, storage drive (magnetic, optical, or solid state), etc.  
  - Peripheral: USB devices, keyboard, monitor, etc.  
  • Other devices: phone, camera, external drives, wearable devices, etc. | • Filesystem  
• Operating System (Windows, OS X, Linux)  
• Applications (MS Word, iMovie, Adobe Photoshop, AutoCAD, Firefox, etc.) | • Data & metadata intentionally created by user (image, video, document, code, etc.)  
• Data & metadata created by system (log files, etc.) |
| | | • Remote servers & storage (including cloud)  
• Web Applications (Gmail and Google Drive, DropBox, iCloud, Facebook, Twitter, Flickr, etc.)  
• Data & metadata intentionally created by user (tweet, image, video, document, code, etc.)  
• Data & metadata created by system (log files, transactional data that assists with management of distributed systems, etc.) |
Multiple Ways to Represent a Digital Record

“Though computer systems maintain ‘an illusion of immateriality’...it is essential to recognize that digital objects are created and perpetuated through physical things (e.g. charged magnetic particles, pulses of light). Digital materials can be considered and encountered at multiple levels of representation, ranging from aggregations of records down to bits as physically inscribed on a storage medium...”

(Cal Lee, Kam Woods, Matthew Kirschenbaum, and Alexandra Chassanoff. "From Bitstreams to Heritage: Putting Digital Forensics into Practice in Collecting Institutions.")
For Example, a Few Different Ways to Represent a Cat Video...

01100011
01100001
01110100
01110110
01101001
01100100
01100101
01101111
meowpurr
Part III:
Best Practices for Creating Personal Digital Records
Think Ahead

- How will you be able to most easily find your records?
- How will you be able to use your records at a later date?

If you can’t find them or you can’t open them, they are of no use to you!
Findability

Things that can increase findability and factors that assist with this:

• Ability to be able to find using a search tool
  • Metadata
  • Full Text Indexing
  • Intelligent and Standard file naming

• Ability to find manually
  • Intelligent and Standard Organization
  • Intelligent and Standard file naming
Usability

• Will you be able to open and use your records 5 years from now? 10 years? 20 years?

• Be mindful of the following:
  • File Formats
  • Storage Media
  • Storage Location
Part IV:
Ownership and Copyright of Personal Digital Records
Ownership and Copyright

<table>
<thead>
<tr>
<th>Ownership</th>
<th>Copyright</th>
</tr>
</thead>
<tbody>
<tr>
<td>Owning the bits you make</td>
<td>Having the exclusive right to reproduce, distribute, perform, display, or create derivatives of a digital record, and the exclusive right to authorize others to do so</td>
</tr>
</tbody>
</table>

Ownership and copyright don’t always go hand-in-hand
Copyright and Ownership Questions to Ask Yourself about the Digital Records You Create

<table>
<thead>
<tr>
<th>Ownership</th>
<th>Copyright</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do you own the digital record?</td>
<td>Do you own copyright to the digital record?</td>
</tr>
</tbody>
</table>
Be Proactive about Your Rights

How can you make proactive choices about the digital records you own and to which you own copyright?

• Use a Creative Commons license standard: https://creativecommons.org/choose/

• Check out the DPLA’s Getting it Right on Rights project: http://dp.la/info/about/projects/getting-it-right-on-rights/
Be Proactive about Your Rights

Do you know what digital records you have licensed to another entity?

Pay close attention to terms of service and user agreements

Casey Fiesler and Amy Bruckman, GVU Center at Georgia Tech
For Example…

<table>
<thead>
<tr>
<th>Ownership</th>
<th>Copyright</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do you own the digital record?</td>
<td>Do you own copyright to the digital record?</td>
</tr>
</tbody>
</table>

- A record you create with a local copy of purchased, copyrighted software (Photoshop, AutoCAD, etc.)
- A record you create, share, and store using offsite/remote/web-based copyrighted software as a service (social media records, records in Google Drive, etc.)

Also ask yourself:
- Have you licensed the record to another entity?
- Will you own or have access to the software required to edit or view the record indefinitely?
- Will you be able to preserve the record indefinitely?
Part V:
Privacy and Security of Personal Digital Records
The digital landscape introduces new privacy and security challenges and intensifies existing issues

- Personal records have always contained private and personally identifying information, but technological factors and the sheer quantity of digital records can make it harder to delete or even know about the private information that exists in a digital record collection.

- Staying aware of the digital records you have and proactively managing them will empower you to deal with privacy and security challenges that might arise.
Encryption

One tactic often used to protect privacy and security is encryption (using algorithms to transform digital records into formats that are intentionally harder to read)

- Some types of encryption: application, operating system, storage system

<table>
<thead>
<tr>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Can enhance privacy and confidentiality</td>
<td>Can hinder your ability to maintain control over your records</td>
</tr>
<tr>
<td>Can protect the integrity and authenticity of a record</td>
<td>If you or someone in the future loses the encryption key</td>
</tr>
<tr>
<td>Can help you maintain control over your records</td>
<td>If software required to decrypt or render encrypted records becomes obsolete or unavailable</td>
</tr>
</tbody>
</table>
Part VI:
Best Practices for Storing Personal Digital Records
Storage Options

- Cloud service
- External storage
- Personal server
Cloud Storage

Your photos, documents, music, email, etc. are stored and managed on servers that belong to someone else.
Cloud Storage: Examples

- Google Drive - 15 GB Free Storage
- Apple iCloud - 5 GB Free Storage
- Dropbox - 2 GB Free Storage
- Box - 5 GB Free Storage
- Microsoft One Drive - 7 GB Free Storage
- Amazon Cloud Drive - 5 GB Free Storage

Comparison of Cloud Services:
http://en.wikipedia.org/wiki/Comparison_of_online_backup_services
# Cloud Storage: Pros and Cons

<table>
<thead>
<tr>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sync data from multiple devices and access data anywhere from any device</td>
<td>Security and privacy concerns</td>
</tr>
<tr>
<td>Provides backup and recovery of data</td>
<td>File type restrictions</td>
</tr>
<tr>
<td>Inexpensive storage option for small amounts of data</td>
<td>Companies and services are not permanent and can change</td>
</tr>
</tbody>
</table>
External Storage

Stores data outside of your computer, laptop, camera, phone, or other device
External Storage: Examples

• External hard drive
• USB flash drive, DVD, Blu-ray, CD (these are sometimes considered less desirable for preservation)

Options:
http://www.pcworld.com/article/248921/need_more_storage_expand_with_external_drives.html
## External Storage: Pros and Cons

<table>
<thead>
<tr>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>You can store an unlimited amount of data outside of your device</td>
<td>Drives can be damaged or lost</td>
</tr>
<tr>
<td>Provides a backup for your data if your device is lost or damaged</td>
<td>Data can rot and decay over time</td>
</tr>
<tr>
<td>Cheaper than cloud storage</td>
<td>More expensive than an internal drive of equal capacity</td>
</tr>
</tbody>
</table>
Personal Server

Hardware and software that provide network service and centralized access to data
Personal Server: Examples

• Dedicated Server
• Virtual Private Server (VPS - Hosts data remotely)
• Embedded Miniserver (ex. Raspberry Pi - Low cost, but slow)
# Personal Server: Pros and Cons

<table>
<thead>
<tr>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Privacy and security</td>
<td>Upfront cost</td>
</tr>
<tr>
<td>Centralized storage for all devices</td>
<td>Must manage software and security</td>
</tr>
<tr>
<td>Access files remotely</td>
<td>Maintenance: power outages and damage</td>
</tr>
</tbody>
</table>

Personal Server Pros and Cons: [https://citizenweb.is/guide/srv/1-why](https://citizenweb.is/guide/srv/1-why)

Part VII:
Best Practices for Access and Ongoing Management of Personal Digital Records
Lifecycle of Digital Stewardship

Create, Organize, Describe

Access, Re-Use, Transform

Appraise and Select

Actively Preserve

Store to Preserve
Store to Preserve

Archivematica Format Policies:
https://www.archivematica.org/wiki/Format_policies

<table>
<thead>
<tr>
<th>Media type</th>
<th>File formats</th>
<th>Preservation format(s)</th>
<th>Access format(s)</th>
<th>Normalization tool</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audio</td>
<td>AC3, AIFF, MP3, WAV, WMA</td>
<td>WAVE (LPCM)</td>
<td>MP3</td>
<td>FFmpeg</td>
</tr>
<tr>
<td>Email</td>
<td>PST</td>
<td>MBOX</td>
<td>MBOX</td>
<td>readpst</td>
</tr>
<tr>
<td>Email</td>
<td>Maildir**</td>
<td>Original format</td>
<td>MBOX</td>
<td>md2mb.py</td>
</tr>
<tr>
<td>Office Open XML</td>
<td>DOCX, PPTX, XLSX</td>
<td>Original format</td>
<td>PDF for PPTX</td>
<td>Tool search in progress</td>
</tr>
<tr>
<td>Plain text</td>
<td>TXT</td>
<td>Original format</td>
<td>Original format</td>
<td>None</td>
</tr>
<tr>
<td>Portable Document Format</td>
<td>PDF</td>
<td>PDF/A</td>
<td>Original format</td>
<td>Ghostscript</td>
</tr>
</tbody>
</table>

Library of Congress Recommended Format Specifications:
http://www.loc.gov/preservation/resources/rfs/
Lifecycle of Digital Stewardship

Create, Organize, Describe

Access, Re-Use, Transform

Appraise and Select

Actively Preserve

Store to Preserve
Part VIII:
Best Practices for the Digital Afterlife
What to Consider When Giving Personal Digital Records to Family/Heirs

• Create a summary description of the files
• Create intelligent file names that include date, location, and context
• Use open formats (PDF, TIFF, JPEG)
• Provide 2 copies in 2 different formats that can be maintained in 2 separate locations
• Keep in stable and moderate temperatures
• Create new media copies every 5 years to prevent data loss
• Pass along digital passwords

Library of Congress on Personal Digital Archiving:
http://digitalpreservation.gov/personalarchiving/
What to Consider When Giving Personal Digital Records to an Institution (Archives, Library, etc.)

• Create a summary description of the files
• Create intelligent file names that include date, location, & context
• Remove inappropriate material
• Use open formats (PDF, TIFF, JPEG)

Society of American Archivists guide, “Donating Your Personal or Family Records to a Repository”:
http://www2.archivists.org/publications/brochures/donating-familyrecs

http://www.clir.org/pubs/reports/pub159/pub159.pdf
References

Wonderful illustrations courtesy of:

• Tom Woolley, created for the "Digital Preservation Business Case Toolkit http://wiki.dpconline.org/ (Creative Commons Attribution-NonCommercial 3.0 Unported License)

• Jørgen Stamp, created for www.digitalbevaring.dk (Creative Commons Attribution 2.5 Denmark license)
References


• Lee, C. (2014) "Digital Forensics Meets the Archivist (And They Seem to Like Each Other)," Provenance, Journal of the Society of Georgia Archivists: Vol. 30: Iss. 1, Article 2. http://digitalcommons.kennesaw.edu/provenance/vol30/iss1/2


• Creative Commons. (2014). “Creative Commons: Choose a License.” http://creativecommons.org/choose/

References


References

• Archivematica. (2014, April 1). "Format policies."  
  https://www.archivematica.org/wiki/Format_policies

  http://www.loc.gov/preservation/resources/rfs/

  http://digitalpreservation.gov/personalarchiving/

• Society of American Archivists, Manuscript Repositories Section. (2013). “Donating Your Personal or Family Records to a Repository.”  
  http://www2.archivists.org/publications/brochures/donating-familyrecs

  http://www.clir.org/pubs/reports/pub159/pub159.pdf
Questions?
Activity

Intro
Activity:
Find the Person in the Personal Digital Archive: Murder Mystery Edition! (25 minutes)
Activity:
Small Group Reflection and Discussion about Personal Digital Archiving (10 minutes)
Activity:
Large Group Sharing and Discussion (10 minutes)
Wrap Up:
Invitation to Host Workshop and Complete Survey
Thank you!
And best of luck with your personal digital archiving...