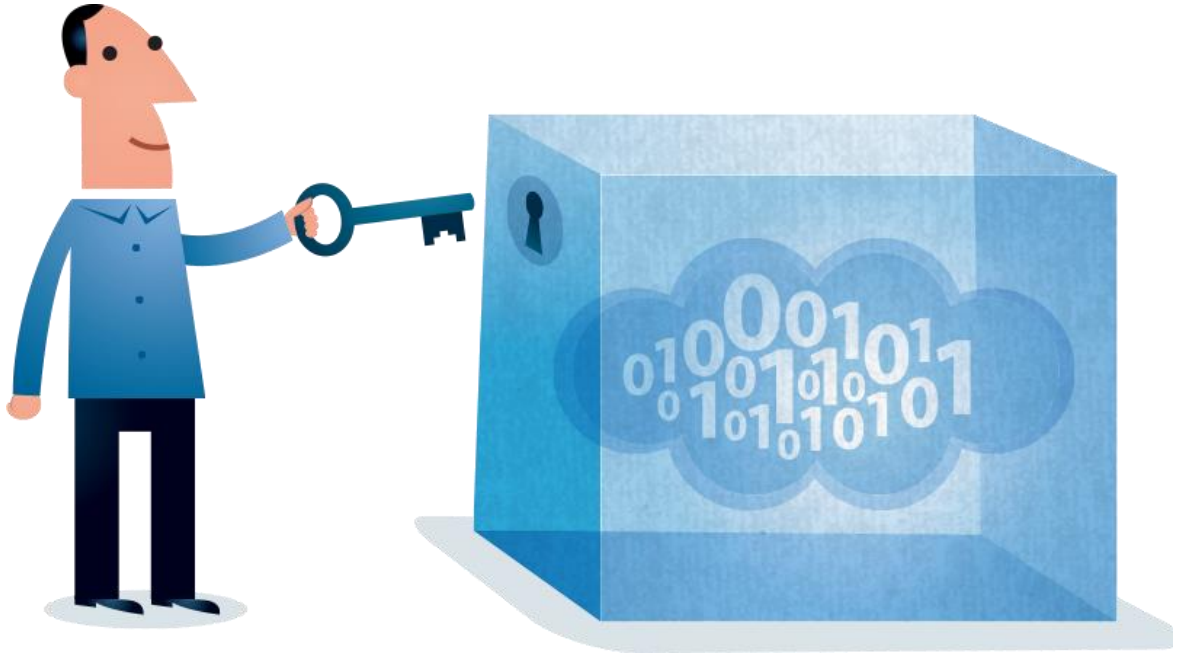
Personal Digital Archiving

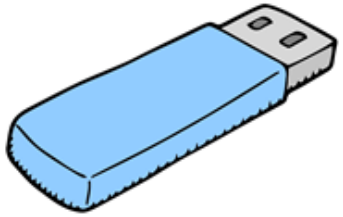
Train-the-Trainer Workshop

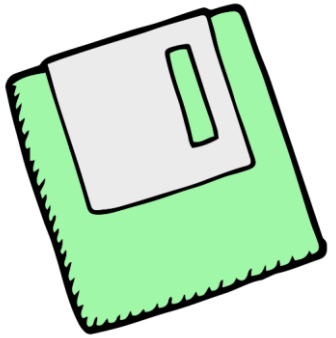
July 31, 2014



Brought to you by:

The Society of Georgia Archivists
The Georgia Library Association
The Atlanta Chapter of ARMA





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.

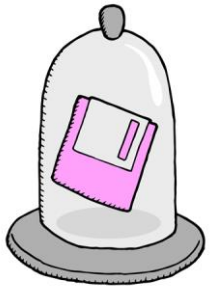
Part I:

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?

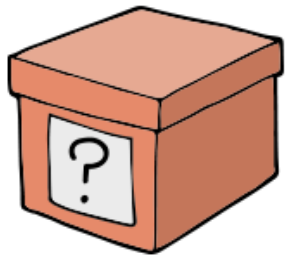
What Qualifies as a Personal Record?



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

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

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)

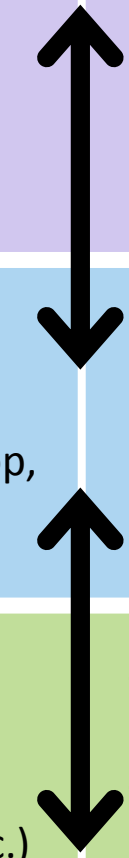
- Digital records are rendered, represented, experienced

- 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

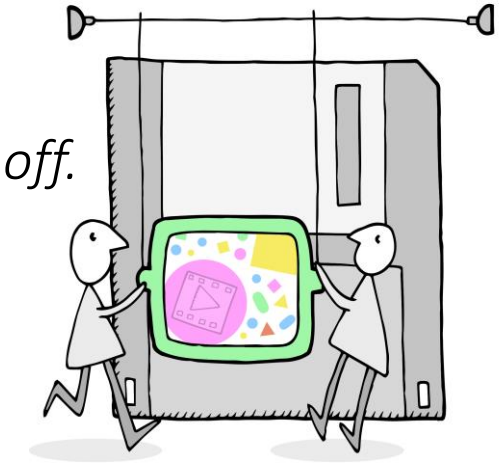
	LOCAL	OFFSITE
HARDWARE	• Personal computer -Internal: CPU, RAM, motherboard, storage drive (magnetic, optical, or solid state), etc. -Peripheral: USB devices, keyboard, monitor, etc. • Other devices: phone, camera, external drives, wearable devices, etc.	• Remote servers & storage (including cloud)
SOFTWARE	• Filesystem • Operating System (Windows, OS X, Linux) • Applications (MS Word, iMovie, Adobe Photoshop, AutoCAD, Firefox, etc.)	• Web Applications (Gmail and Google Drive, DropBox, iCloud, Facebook, Twitter, Flickr, etc.)
DATA	• Data & metadata intentionally created by user (image, video, document, code, etc.) • Data & metadata created by system (log files, 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…

Part III:

Best Practices for Creating Personal Digital Records

Think Ahead

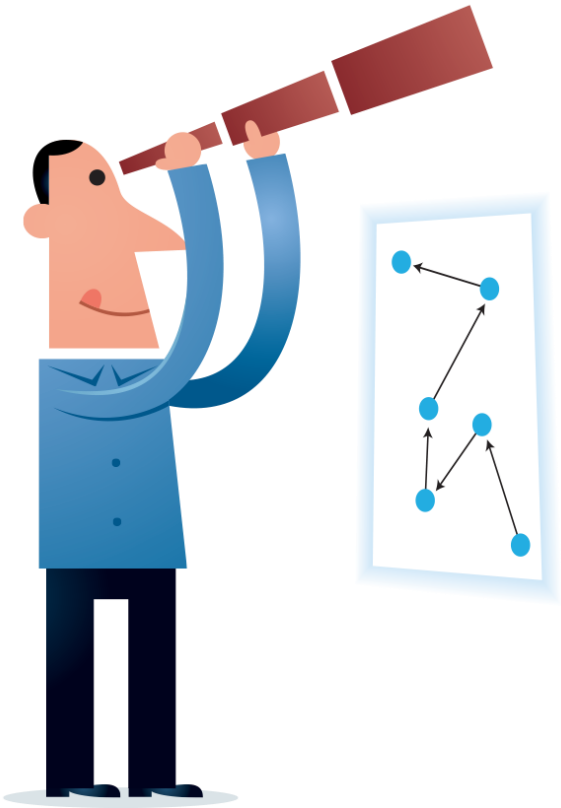
- How will you be able to most easily find your records?
- How will you be able to use your records at a later date?

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

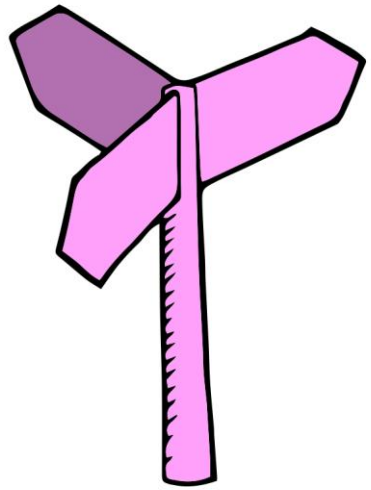
Findability

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

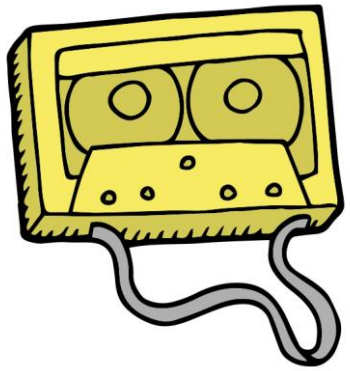


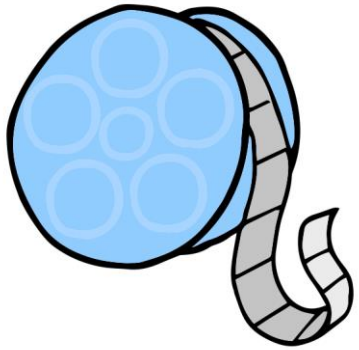
Usability

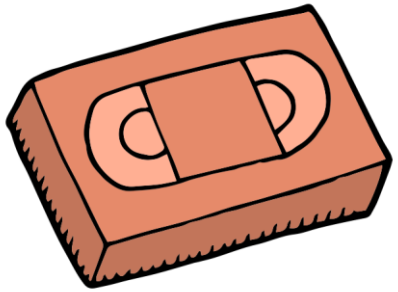


- Will you be able to open and use your records 5 years from now? 10 years? 20 years?

- Be mindful of the following:
  - File Formats
  - Storage Media
  - Storage Location







Part IV:

Ownership and Copyright of Personal Digital Records

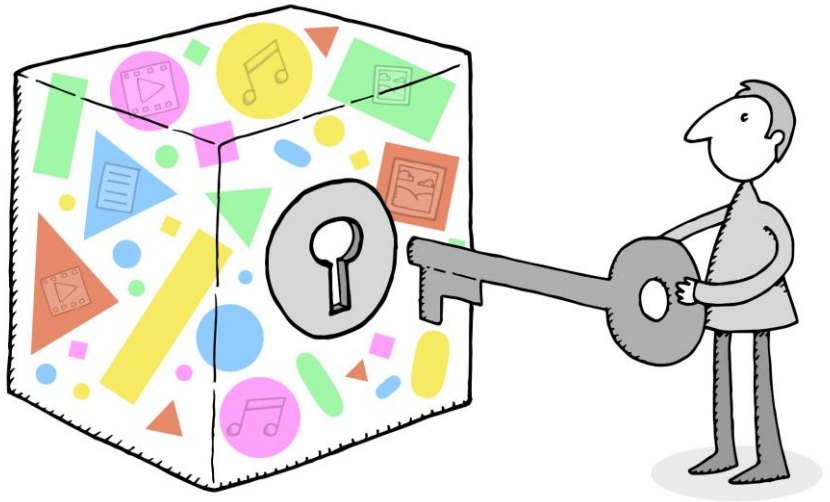
Ownership and Copyright

Ownership	Copyright
Owning the bits you make	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

Ownership and copyright don’t always go hand-in-hand

Copyright and Ownership Questions to Ask Yourself about the Digital Records You Create

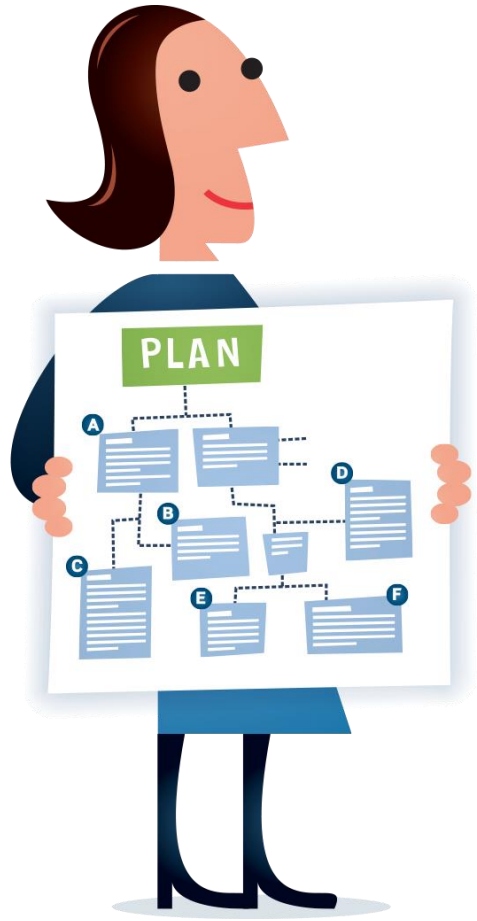
Ownership	Copyright
Do you own the digital record?	Do you own copyright to the digital record?



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…

Ownership	Copyright
Do you own the digital record?	Do you own copyright to the digital record?

- 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:

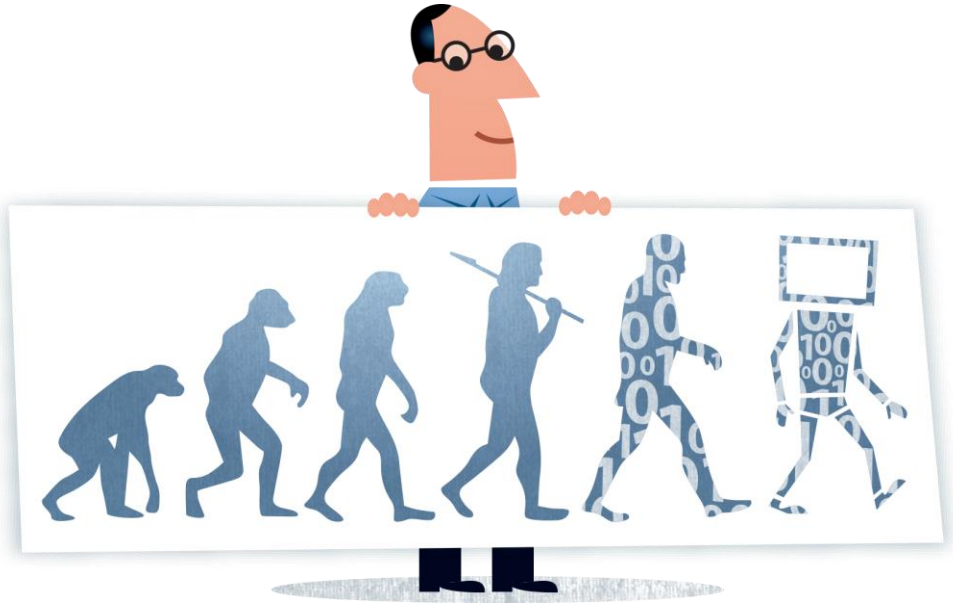
- Have you licensed the record to another entity?
- Will you own or have access to the software required to edit or view the record indefinitely?
- Will you be able to preserve the record indefinitely?

Part V:

Privacy and Security of Personal Digital Records

Digital = New Privacy and Security Challenges

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

Pros	Cons
Can enhance privacy and confidentiality	Can hinder your ability to maintain control over your records
Can protect the integrity and authenticity of a record	If you or someone in the future loses the encryption key
Can help you maintain control over your records	If software required to decrypt or render encrypted records becomes obsolete or unavailable

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

Pros	Cons
Sync data from multiple devices and access data anywhere from any device	Security and privacy concerns
Provides backup and recovery of data	File type restrictions
Inexpensive storage option for small amounts of data	Companies and services are not permanent and can change

External Storage

Stores data outside of your computer, laptop, camera, phone, or other device

External Storage: Examples

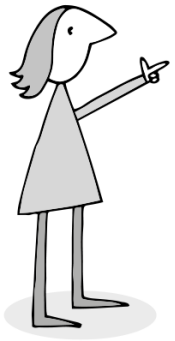
- External hard drive
- USB flash drive, DVD, Blu-ray, CD (these are sometimes considered less desirable for preservation)

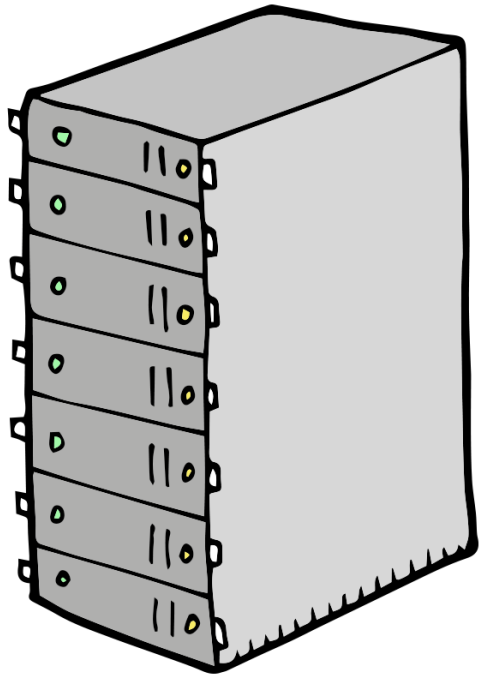
Options:
http://www.pcworld.com/article/248921/need_more_storage_expand_with_external_drives.html

External Storage: Pros and Cons

Pros	Cons
You can store an unlimited amount of data outside of your device	Drives can be damaged or lost
Provides a backup for your data if your device is lost or damaged	Data can rot and decay over time
Cheaper than cloud storage	More expensive than an internal drive of equal capacity

Personal Server

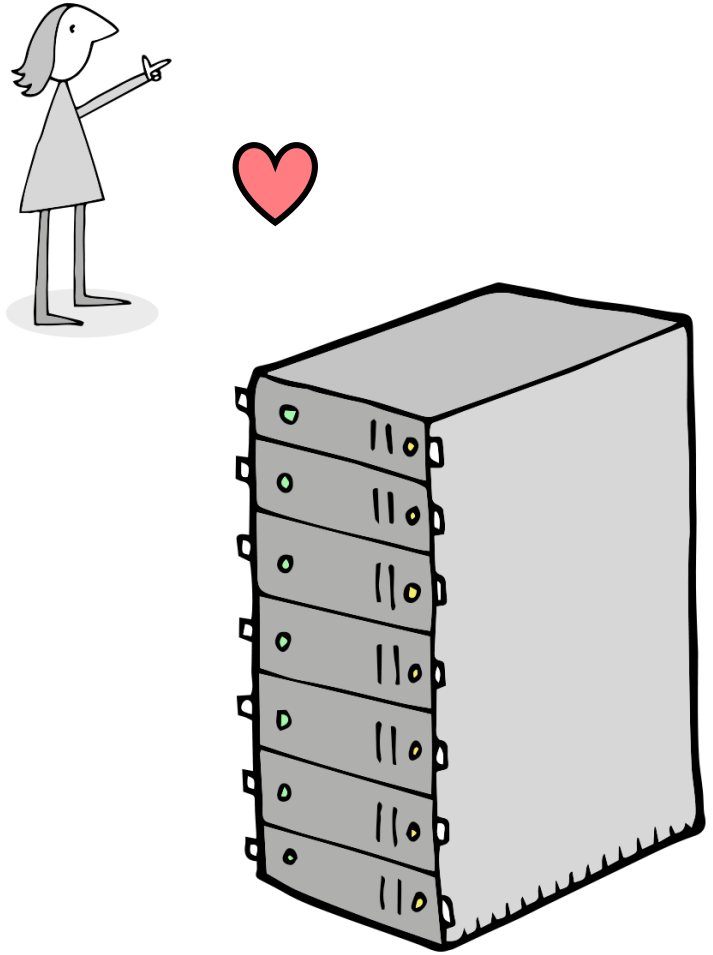




Hardware and software that provide network service and centralized access to data

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

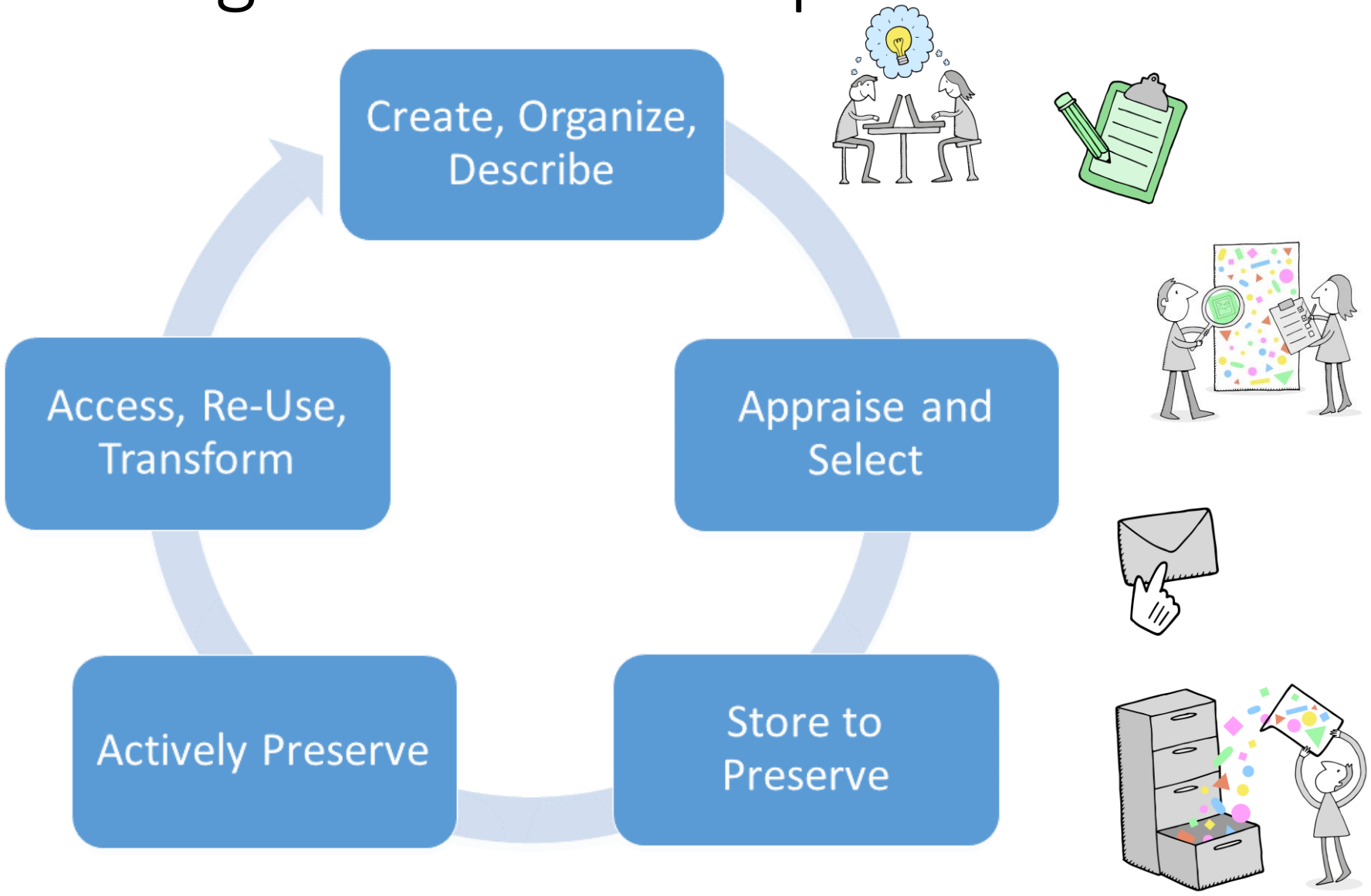
Pros	Cons
Privacy and security	Upfront cost
Centralized storage for all devices	Must manage software and security
Access files remotely	Maintenance: power outages and damage

Personal Server Pros and Cons: https://citizenweb.is/guide/srv/1-why
PersonalServer.com: https://www.personalserver.com/web/en/home

Part VII:

Best Practices for Access and Ongoing Management of Personal Digital Records

Lifecycle of Digital Stewardship



Store to Preserve

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

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

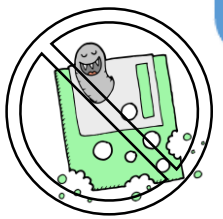
Library of Congress Recommended Format Specifications:
http://www.loc.gov/preservation/resources/rfs/

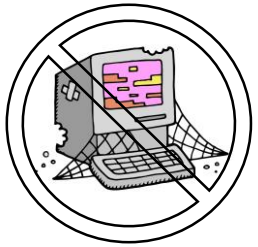
Lifecycle of Digital Stewardship

Create, Organize, Describe

Appraise and Select

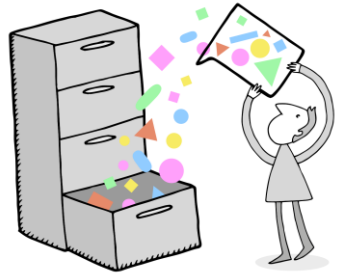
Access, Re-Use, Transform





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

Council on Library and Information Resources report, “Born Digital: Guidance for Donors, Dealers, and Archival Repositories”: 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

- Google Dictionary. (2014). “Record.” http://www.google.com/search?q=define+record
- 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
- Lee, C., Woods, K., Kirschenbaum, M., & Chassanoff, A. (2014). “From bitstreams to heritage: Putting digital forensics into practice in collecting institutions.” http://www.bitcurator.net/docs/bitstreams-to-heritage.pdf
- Creative Commons. (2014). “Creative Commons: Choose a License.” http://creativecommons.org/choose/
- Digital Public Library of America (DPLA). (2014). “Projects: Getting it Right on Rights.” http://dp.la/info/about/projects/getting-it-right-on-rights/

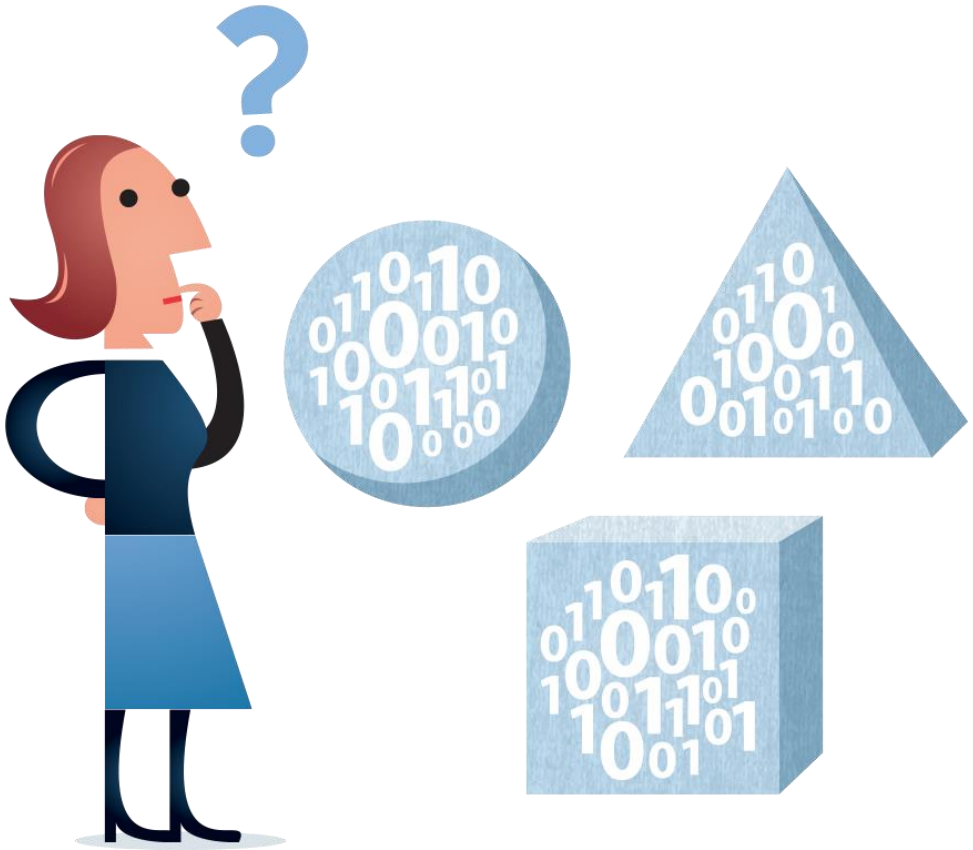
References

- Bruckman, A., Fiesler, C., & Georgia Institute of Technology Institute Communications. (2014, April 18). “Do You Read Terms of Service? Maybe You Should.” http://www.news.gatech.edu/2014/04/18/do-you-read-terms-service-maybe-you-should
- Wikimedia Foundation. (2014, July 23). “Comparison of online backup services.” *Wikipedia*. http://en.wikipedia.org/wiki/Comparison_of_online_backup_servicesBradley, T. (2012, January 29).
- “Need More Storage? Expand with External Drives.” *PCWorld*. http://www.pcworld.com/article/248921/need_more_storage_expand_with_external_drives.html
- Citizen Web. (2014). "3.1 – Why a Personal Server?." *The CitizenWeb Project*. https://citizenweb.is/guide/srv/1-why
- Blackbrick. (2014). "Personalserver.com: The Future of Personal Computing." https://www.personalserver.com/web/en/home

References

- Archivematica. (2014, April 1). "Format policies." https://www.archivematica.org/wiki/Format_policies
- Library of Congress. (2014). "Recommended Format Specifications." http://www.loc.gov/preservation/resources/rfs/
- Library of Congress. (2014). “Personal Digital Archiving.” 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
- Redwine, G., Barnard, M., Donovan, K., Farr, E., Forstrom, M., Hansen, W., John, J. L., Kuhl, N., Shaw, S., & Thomas, S. “Born Digital: Guidance for Donors, Dealers, and Archival Repositories.” Council on Library and Information Resources. 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…