Personal Digital Archiving: Managing Born-digital audiovisual collections across the lifecycle

Howard Besser
Moving Image Archiving & Preservation
New York University
http://besser.tsoa.nyu.edu/howard/Talks/
http://www.nyu.edu/tisch/preservation/

General problems of born-digital personal content

In the analog world

- Traditionally, we have come to understand the work of writers and scientists by scholars studying their papers in Special Collections and Archives
- Their correspondence and progressively different drafts of papers reveal their changing thoughts and craft
- But how do we gather these in the Digital Age?

Alasdair Gray's Lanark (Glasgow U Library)

Correspondence
Correspondence

Where can we find these today?

- Do people write letters on paper? Can we see the iterations of changes on manuscripts?
- Where can we find today’s equivalent of these?

- This will require
  - new interventions (like changing authors’ workflow, or intervening in email handling software)
  - New tools (like for analyzing email)
  - new approaches like digital archeology, forensics

Stages of the problem

- Stage #1: People write on computers instead of paper
- Stage #2: People no longer store their digital works in places over which they have absolute control
  - Email services (gmail, yahoo)
  - Cloud storage for documents (google docs)
  - Social network services (FLICKR, YouTube)

Our Changing Environment

If we're lucky:
Our Changing Environment

• Rise of Online Services and Social Media is changing where this content resides (and is imposing restrictions that go beyond the rightsholder)

Core Multi-location Problems

• It’s difficult enough when someone’s photos are spread throughout their hard disk. But today some photos there, but others on their phone(s), Instagram, Flickr, Facebook, in Tweets, etc.
• Similar problems plague email
• Most Social Network TOS policies prohibit the owner from giving their password to anyone else (even Library)

And how do we handle donations after an important person dies?

And these issues are also true for Community Grps & Assns

• w/Social Media, group activity is more important than ever
• But each person in the group is an individual collector. And frequently a set of individual collections forms the group collection.

Famous visitor to UCLA

Documenting Protests

• photo from Activists Guide to Archiving Video
Every Image Collector has a Different Approach
• Different file-naming conventions
• Different file formats
• Different compression schemes
• Different metadata
• Stored in different arrangements/hierarchies
• Stored in different places (cellphone, personal hard disk, Instagram/Facebook, …)

THE PDA CONFERENCES

PDA Covers born-digital
• Correspondence/email
• Personal photos and group photo albums
• Manuscript drafts
• Personal documents
• Home movies
And has been extended to encompass:
• Family history
• Community/Ethnic history & Movements
• Genealogy
• Digital humanities.

PDA: Who Attends & Presents
• Citizen Archivists
  — People who want to step in and rescue content in peril
  — People who like to create software/Apps/Guidelines to help others facing similar problems
• Community or Ethnic groups and activists wanting to save portions of their heritage
• Professional librarians & archivists (and their programming support staff)
• Regular software developers
• Researchers (both academic and computer industry)

PDA Goals—Sharing knowledge
• What worked and what didn’t; what parts turned out to be more difficult than anticipated
• New and different types of content to collect
• Guidelines, procedures, workflows, methodologies
• Software

PDA History
Initially started by Internet Archive with co-sponsorship from Netherlands Sound & Vision, LC/NDIIPP and CNI
• 2010 Internet Archive
• 2011 Internet Archive
• 2012 Internet Archive
• 2013 Univ of Maryland
• 2014 Indiana State Library
• 2015 New York University
PDA 2015


PDA 2015 Keynote—Digital Diaspora

http://artsfwd.org/digital-diaspora-family-reunion/

PDA Keynote II

PDA 2015-

- Preserving Digital Photos
- Digital Preservation of Art
- Creative use of Open Source tools
- College Library sponsored Personal Digital Archiving Days
- Community Engagement
- Video Preservation
- Digital Humanities & Social Science
- Workshops

Preserving Digital Photos

- Todd Wimmen, "Methods and Projects for Recording Audio Related to Family Photographs" (PDF)
- Meredith Powers, "Instagram as Quarion, Participation, and Documentation" (PDF)
- Melody Condon, "Digitalization & Curatorial of Personal Photo Collections: Recommendations, Methods & Tools from a Recent Project"
- Sarah Stevenson, "How I Used Flickr, WordPress and Embedded Metadata to Build My Family Archive" (PDF)
- Rita Nix, "Methods for Organizing, Cataloging, and Developing Metadata for Family History Photos" (PDF)
- Joel Neville Anderson, "The Circulation of Domestic Portraiture Following Japan’s Triple Disaster of 2011" (PDF)

Digital Preservation of Art

- Jo Ana Morfin, "Encounters between Conservators and Digital Artists in Latin America" (PDF)
- Lauren O’Connor and Kelly Hayton, "Models for Collaboration in the Digital Preservation of Audiovisual Materials" (PDF)
- Jason Keener, "Preserving the History and Theory of Digital Art in Video and Online" (PDF)
- Sam Mieras, "Embracing Collaboration to Address the Personal Digital Archive Needs of a Contemporary Artist" (PDF)
- Pamela Vizner, "Stretched Resources, Flexible Solutions: XFR Collective In Action"
- Dana Haugh, "Navigating the Challenges of Widescale Dissemination in the Performing Arts"
- Shannon Luddy, "AtCHiving: A Not-for-Profit Art Community’s Approach to Digital Archiving"
- Blaikley McDowell, "Dance Heritage Coalition"

Creative use of Open Source tools

- Jason Scott, "When the Emulators Broke Free. "Scott's basic message was that the Internet Archive has made significant advances with emulation of old games online.
- Justin McIvor, Mark Simon Hauck and Ashley Brawner, "BitTorrent, and P2P as a tested, effective tool to transfer large files over the internet online."
- Cal Lee and forkwood, "For The Layperson" (PDF)
- Jenifer Lauer, "Building Digital Collections in Oregon For The Layperson"
- Ashley Brawner, "No, you don’t know about you, but I’m feeling the links right now.
- Peter Chan, "A Guide to P2P" (PDF)

College Library sponsored Personal Digital Archiving Days

- Rachel Appel, Bryn Mawr College
- Amy Bocko, Wheaton College
- Joanna DiPasquale, Vassar College
- Sarah Walden, Amherst College
- Kevin Powell, Brown University

Community Engagement

- Andrea Copeland and Lydia Spotts, "Community Engagement Methods for Systems Design and Archive Building" (PDF)
- Draegaan Espenschied, "Community-Based Web Archiving"
- Lauren Algee, "Community Archiving – The DC Punk Archive at DC Public Library" (PDF)
- Jennifer Douglas, "They Were Still Here: Archives and Online Grief Communities" (PDF) - parents grieving their stillborn
Video Preservation

- Paul Dougherty, "40 Years of Video Husbandry" (PDF)
- Zack Lischer-Katz, "Reflections on Digitizing a Personal Audiovisual Collection" (PDA)
- Rebecca Fraimow, "Digital Video Remix: A Risk Assessment." (PDF)

Digital Humanities & Social Science

- Bonnie Gordon, "Sustainable Digital Preservation at Interference Archive" (PDF)
- Eira Tansey, "Large-Scale Archiving and the Right to be Forgotten" (PDF)
- Susan Asman, "The Cultural Dynamics of Personal Archiving"
- Lori Kendall, "Memory is a Process, not an Artifact"
- Paula Murphy, "Personal Digital Archival Collecting Old and New"
- Julie C. Swierny, "On Preserving Technological and Social Forms Used in Social Media" (PDF)
- Kyung Eun Oh & Vanessa Reyes, "A Study of How Academics Manage their Personal Digital Information and Personal Paper-Based Information in their Work Spaces" [They wrote a blog post for The Signal about their research.]
- Ellysa Stern Cahoy/Smiljana Antonijevic, "The Zotero/ScholarSphere Project"
- Jessica Busher,"How Online Sociality and Free Terabytes are Shaping Personal Digital Archiving"
- Travis L. Wagner, "Remembering South Carolina Women’s History."

Workshops

- Workshop on Do-It-Yourself Personal Digital Archiving
- Workshop on Do-It-Yourself Personal Digital Archiving
- Workshop on Do-It-Yourself Personal Digital Archiving
Workshop on *Do-It-Yourself Personal Digital Archiving*

*Workshop on ArchiveMatica*

*Workshop*

*Workshop on Email Archiving*

**Make it easy for future contributors to create “archival-friendly” works**-

- Low-hanging fruit
- Easy instructional material that appeals to what they think is important

**SOME KEY IDEAS FROM ACTIVIST ARCHIVISTS**

[HTTP://ACTIVIST-ARCHIVISTS.ORG/](http://activist-archives.org/)
Low-Hanging fruit

- Turn GPS on
- Develop strategies for automating a profile and uploads (our ideal App)

7 Tips to Ensure Your Video Is Usable in the Long Term

- Collect details while filming
- Keep your original raw footage, unaltered
- Make your video discoverable
- Contextualize it
- Make it verifiable
- Allow others to collect and archive
- Or archive it yourself

WITNESS: Activists’ Guide to Archiving Video, Yvonne Ng

http://archivewguide.witness.org

Find smart ways to deal with Scale-

Tamiment YouTube collecting

- plug-in for FireFox (downloadhelper.net)
- As of April, ca. 250 items, policy: large events
- Fair Use: 2012 ARL Code of Best Practices
  - “transformative” collecting with context
- Tamiment has been selectively browsing through YouTube Occupy videos, trying to choose which ones to keep, then cataloging them with
  - Title, Creator, Creation Date, Upload Date, Description, URL, Youtube Username, License, Format, Codec, Source Media, On Internet Archive, CC License type

Tamiment YouTube Cataloging
But this won’t scale!

March 24, 2012 YouTube stats
(just 6 months after start of movement)

- “#Occupy” 169,000
- “Occupy Wall Street” 98,400
- “Occupy Protest” 70,500
- “Occupy Movement” 54,800
- “#OWS” 50,300
- “Occupy Oakland” 13,400
- “Zucotti Park” 6,690

Alternative approach to YouTube Selection process
- Develop categories of important YouTube videos
  - Celebrity visits, Internal workings (library, kitchen, media), Confrontations with police, Labor, Housing, etc.
- Have Occupiers fill in an online form listing the 5 most important videos in each category

Advantages of YouTube Collaborative Filtering Selection Process
- Scalable and manageable
- Consistent with Occupy ideas of inclusiveness and of managing own story
- Tamiment can still choose to be selective in collecting only a portion of what is voted in, but the total set for review is a manageable scale

Tips on Outreach to Communities
- Build trust
- Speak in their language (not MARC-speak)
- Identify ways you can meet needs they already perceive
- Approach projects as collaboration whenever possible
- Don’t only focus on content and metadata, but also rights that can be an impediment to preservation

Personal Digital Archiving 2015
http://personaldigitalarchiving.com/
Personal Digital Archiving: Managing Born-digital audiovisual collections across the lifecycle

- Special Thanks to rest of the PDA 2015 Organizing Committee (especially Mike Ashenfelder for coherent summarization)
- http://personaldigitalarchiving.com/
- http://besser.tsoa.nyu.edu/howard/Talks/