Digital Accessioning: Developing Basic Processes

Society of California Archivists

Annual General Meeting, April 2012

Nancy Enneking

Institutional Records and Archives

Getty Research Institute

Current Environment

- Trust and its programs: Conservation Institute; Foundation; Museum; Research Institute (GRI)

- Departments with born-digital material:
  – Special Collections
  – Institutional Archives

- Technology assistance
  – Information Technology Services (Trust)
  – Library Information Systems (GRI)

Recent Digital Materials Initiative

To define, recommend, and, where possible, implement the policies, procedures, and technological structures/systems required to govern the process of managing, from acquisition to permanent preservation, the born digital materials of the GRI, particularly Special Collections and Institutional Archives.

Primary focus on born-digital rather than digitized records

Physical Media

Institutional Records and Archives is increasingly receiving materials in digital format on a variety physical storage devices in diverse file formats, many of which are dependent on obsolete technologies.
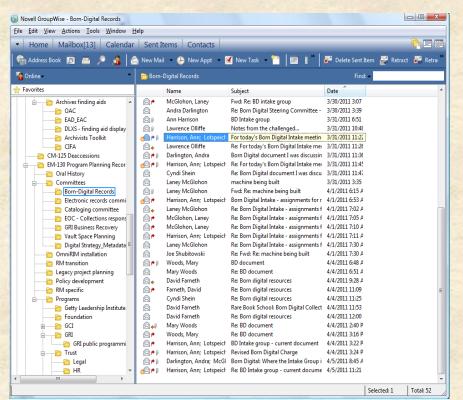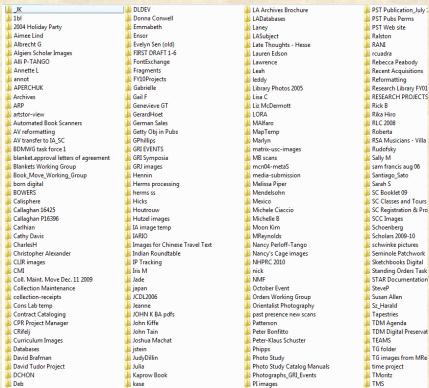
Physical Media

Images and oral history interviews with artists stored on hard drives

Networked Information

Ingest of Digital Material

Best Practices:
(http://tinyurl.com/DigitalArchivesBestPractices1)

- Move Data Off At-Risk Media
- Capture Bit-for-Bit Image of the Data
- Quarantine and Inventory Incoming Material
- Protect Authenticity and Integrity of Data

- That's great!! ………. How?

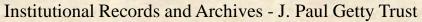# Initial Acquisition: Media

For material on media:

Building forensic workstations: Fluffy and MacFluffy

- Platforms for data acquisition
- Intent is to move toward accessioning the content, not the carrier
- Utilizing any expertise/skills of existing staff
- Reusing /finding inexpensive equipment to allow access to a variety of media
- Scrounging the web for relevant open-source software
- Purchasing write blockers; using existing virus scanners
- Not networked, purchased transfer media

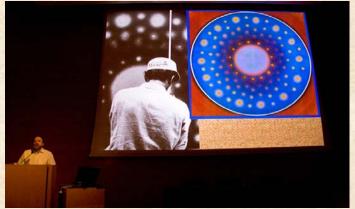# Initial Acquisition: Fluffy

Initial Acquisition: Media

Basic process:

- Use forensic software to view content and determine how to proceed
- Create a disk image and set the physical media aside
- Depending on the materials in question:
  - Save the disk image as such until we can manage the content
  - Use the forensic software to export the desired files or to create a smaller custom image
- Run the content through the Duke DataAccessioner
- Save the content & directory structure produced by the forensic software to transfer media and take to a networked machine for storage
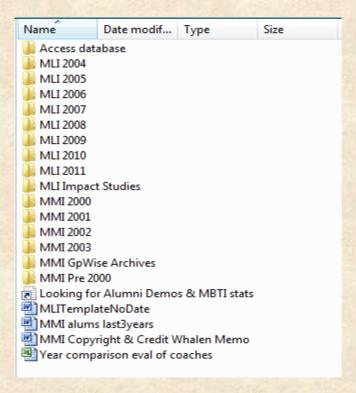
Initial Acquisition: Networked

Most content is currently:
• Event-related (audio, video, and/or image with hard copy ephemera and contracts)
• Oral Histories

Initial Acquisition: Networked

Beginning to take internal documents and portions of network drives

Initial Acquisition: Networked

Basic process:

- Established a workbench and backed-up storage location (6 terabytes total) for digital content

- We do not create disk images, do not run separate virus scans

- Beginning to use Duke DataAccessioner (but not for AV)

- Creating preservation copies for AV and some doc files as we go (if we can) using a combination of SoundForge, Streamclip, Mp3tag

- Checking rights and contracts (as we would for anything else!)

Initial Acquisition: Networked

Accession Documentation
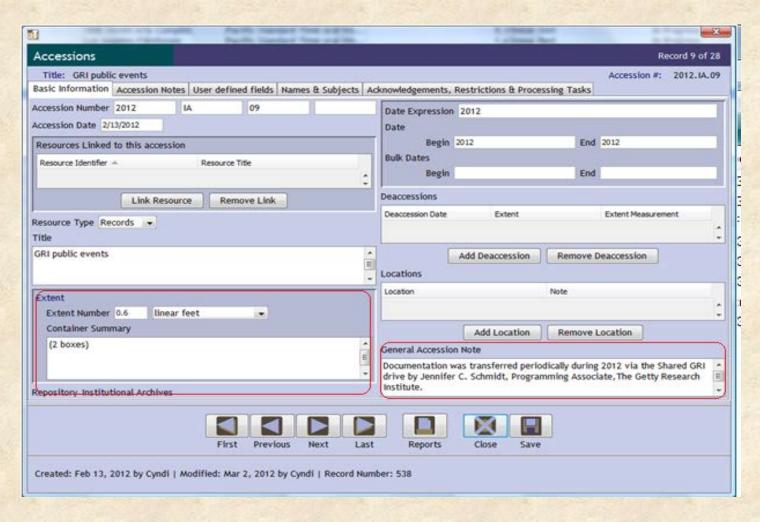
Two primary administrative records for each accession:

- Hard copy accession file
- Archivists' Toolkit record

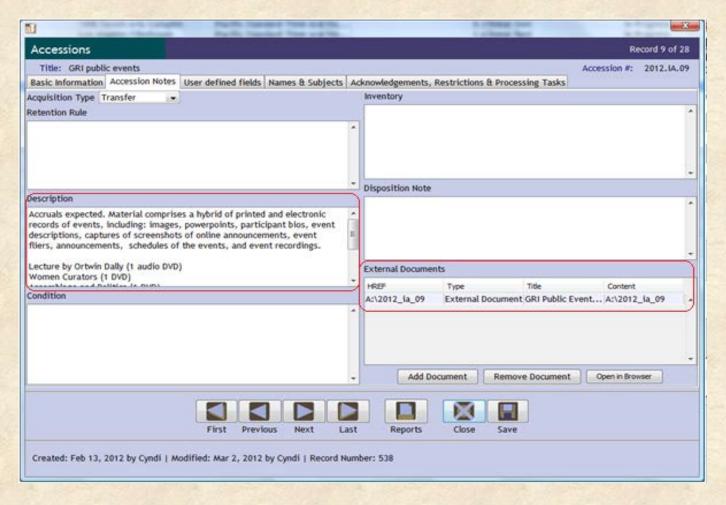Additionally we create an electronic accession file

- ReadMe files that we create
- Spreadsheets, directory lists, etc. created by the software we are using and, in some cases, Karen's Directory Printer

Accession Documentation: Toolkit

Accession Documentation: Toolkit

Accession Documentation: Toolkit

Institutional Records and Archives - J. Paul Getty Trust

Accession Documentation: Toolkit
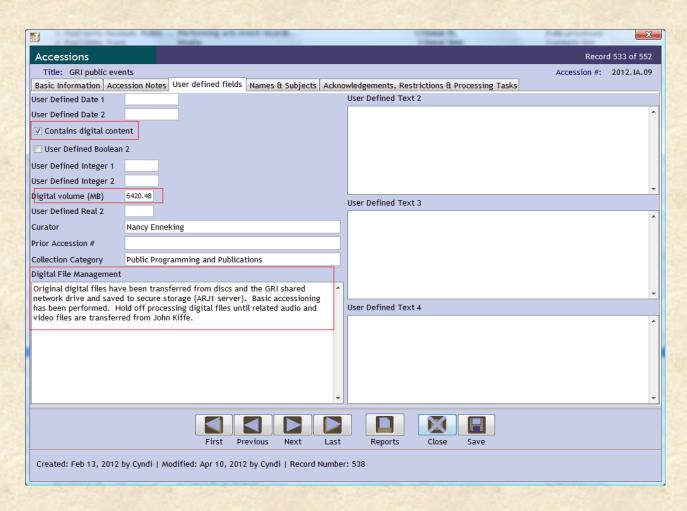
Work in Progress: Policy and Process

Legal and Ethical Issues

- Rights and take-down policy
- Agreements with donors/sellers - agreements made by people who are familiar with paper, not technology
- Do we recover deleted files?

Appraisal and Retention

Best Practices

- Appraisal should be performed
- Retain original media/carriers

Our Approach: It Depends

Work in Progress: Policy Development

Metadata harvesting/enhancement

- Identified the standards we want to use
- Harvest what we can on intake, but how do we enhance????
- Exploring KEA for Keyword metadata extraction

Description and Access

- Convert directory listings in Excel to EAD components and load into Toolkit
- Use Toolkit to intellectually combine paper and digital content using DAO's
- Initial example: http://hdl.handle.net/10020/cifaia40011

Work in Progress: Policy Development

Permanent Preservation

Best Practices

- Establish secure storage spaces; back up the data; emulation v. migration?

Our Approach

- We have elements but not a full plan
- Determining systems
- Migrate most content, emulate only as required
- Planning for a digital policy audit retreat using DRAMBORA (Digital Repository Audit Method Based on Risk Assessment)

Work in Progress: General

- Investigating Archive-IT and the CDL web archiving services

- Currently writing policy and procedure manuals

Major Lessons for Ourselves

Progress was impossible until we started getting over the fear of imperfection.

As with everything else, we'll do the best we can with the resources we have.

Questions, Comments?

Nancy Enneking

Head of Institutional Records and Archives

Getty Research Institute

J. Paul Getty Trust

(310) 440-7066

nenneking@getty.edu