The Bancroft and Born-Digital: Preserving the Archives of Tomorrow

Mary Elings and Kate Tasker
The Bancroft Library
12.14.2017
Research IT Reading Group
Introduction

The Bancroft Library’s Digital Collections Unit:

★ Large scale digitization initiatives
  ○ Calisphere
  ○ 1906 Earthquake and Fire
  ○ JACS (Japanese American Confinement Sites)

★ Born-digital program
  ○ Obsolete computer media
  ○ Hard drives
  ○ Websites, social media

http://news.lib.berkeley.edu/Topics/bancroft-library/digital-collections/
Born-Digital Collection Holdings

• 2,272 floppy disks (1,762 3.5” + 510 5.25”)
• 1,460 optical discs (CDs, CD-ROMs, DVDs)
• 190 Zip disks
• 48 hard drives / flash drives
• 11 websites / social media feeds
• 8 computers
• 3 email accounts
• ...and counting!
18.9 TB

Estimated total size of all Bancroft’s Born-Digital Collections

(= 18,900 gigabytes)

(= 18,900,000 megabytes)
Born-Digital Collections Include Materials From...

~ Bruce Conner Papers
~ Karen Lee Wald Research Collection on Cuba
~ Randall Milliken Papers
~ Columbia Foundation Records
~ UC Berkeley Department of Nuclear Engineering Records
~ Ira Nowinski California Native American Photograph Archive
~ Cynthia (Kiki) Wallis and René Yañez Real Mission School Archive
~ Congressman George Miller Papers
~ San Francisco Chronicle Two Cents Project Records
~ Rex Ray Archive
~ Digital Files from the Barbara Boxer Photograph Collection
Average Size: 140 GB

Estimated Processing Time

1 Gb = 5 hours (minimal processing)
1 Gb = 7 hours (moderate processing)
1 Gb = 10 hours (maximum processing)
Pathways to Preservation

Digital Forensics
Fixity Checks
Emulation
File Format Migration

Forensic Recovery of Evidence Device (FRED)
Building Bancroft’s Born-Digital Program

Collaborate with donors, curators, archivists, technologists

Manage acquisition, authentic data recovery, storage, and preservation

Perform data analysis and description

Screen for sensitive data

Research new formats and tools

Design methods for access
Workflow Overview

1. Evaluate digital media at the point of acquisition
2. Conduct archival appraisal
3. Inventory source media
4. Plan strategies for file capture and preservation
5. Perform archival processing
6. Screen for sensitive/personal content
7. Migrate files to preservation/access formats
8. Create catalog records and finding aids
9. Provide access...
Evaluate Digital Media at the Point of Acquisition

Digital Material Donor Survey
The Bancroft Library, UC Berkeley

Digital Material Creation and Function

1. What are the functions and activities which led to the creation of your digital materials?
2. What kinds of records are created? (Word documents, spreadsheets, databases, email, images, video, web, etc.)
3. Are personal files separated from work files?
4. Are you solely responsible for creating these digital materials? If not, who are the other creators, and how are their files separated? Are there multiple user accounts on your computer?
5. What is the date range of the digital files?
6. Are there any files you would like to transfer to us in the future? Please provide details.

Digital Material Storage and Organization

1. How are the digital materials organized?
Conduct Archival Appraisal
Inventory Source Media

BANC DIG 2013.008.002
CD
Memorex 2004, Bruce Conner Scrapbook #1 1963-1970
Label Cont. Firm + looking for mushrooms + cosmo ray + Marilyn Times 5. Images 80-100 (total unknown)

Imaging Notes
HFS+
File Size Total: 634.12 MB
File Location:
Inventory Filename:
File Types Present:

DISK IMAGE MEDIA TRACKING
Date Delivered: 11/7/2013
Date Returned: 12/3/2013
Del By: JG
Del To: LSO
PU By: JG
PU From: LSO
Plan Strategies for File Capture & Preservation
Perform Archival Processing

- Forensic Toolkit (FTK)
- DROID
- BitCurator
- ePADD
Screen for Sensitive/Protected Content
Migrate Files to Preservation/Access Formats
Create Catalog Records & Finding Aids

To improve the clarity and usefulness of finding aids and to promote consistency across campuses, a working group of digital archivists under the aegis of the UC Born-Digital Content Common Knowledge Group (CKG) have collaborated to develop a UC-wide descriptive standard for born-digital archival material.

---

**Bruce Conner papers**

**Basic Information**

- **Title**: Bruce Conner papers
- **Identifier**: BANC MSS 2000/36
- **Level of Description**: Collection
- **Language**: English
- **Purchased?**: True
- **Restrictions**: False

**Dates**

- Creation: 1940-2008
- Creation: Ink/1960s-2010

**Extents**

- 50 Linear Feet
- 20 Cartons, 1 Box, 2 Oversize boxes, and 7 Oversize Folders
- 1.4 Gigabytes (556 MBs)
Future Archives

Providing Access to Born-Digital Collections
Challenges & Opportunities

1. Scalability and automation
2. Digital forensics
3. Long-term secure preservation storage
4. Processing priorities
5. Access
Thanks!

Questions?

Mary Elings, Assistant Director & Head of Technical Services
melings@library.berkeley.edu

Kate Tasker, Digital Archivist
ktasker@library.berkeley.edu

[Bancroft Building, 721 Market Street, San Francisco], Framed items from the collections of The Bancroft Library, BANC PIC 1905.11569a--FR, The Bancroft Library, University of California, Berkeley.