

“OUT OF MANY, ONE”

Creating a UC-Wide Descriptive Standard for Born-Digital Archival Material

Annalise Berdini (Princeton, formerly UCSD)

Charlie Macquarie (UCSF)

Shira Peltzman (UCLA)

Kate Tasker (UCSF, formerly UCB)

🐦 @annarchives

🐦 @advnturelibrary

🐦 @shirapeltzman

🐦 @kate_tasker

Hi everyone!

- Background & origin
- Methodology
- Working process
- UC Guidelines → overview, highlights, & future plans
- Q&A (?)

Origins

- Born out of UC Born Digital Content [Common Knowledge Group \(CKG\)](#)
 - Guideline developers realized shared need for standardization in this area
 - Recognized opportunity for select group of experts to work independently while being supported by larger CKG structure
 - Chance to get in on the ground floor to help define practice going forward

Limits of Existing Guidelines

- Describing Archives, A Content Standard (DACS) and Encoded Archival Description (EAD) are not always clear, do not always agree
 - DACS 7.1.8, Processing Note: barely mentions digital material
 - DACS 4.3, Technical Access: assumes different processing practices than have emerged in the field
 - DACS 7.1 (Notes elements) do not always clearly map to EAD, and some mappings (<control> vs <eadheader> for example) have changed from EAD 2002 to EAD3

Varying Practices for Description at UCs

- Many programs not yet started or formalized
- Digital had not been written into many people's processing manuals
- Entirely different practices from institution to institution
 - Extent → Storage? Number of files? Number of media?
 - Abstract → Should digital nature of material be mentioned at all?

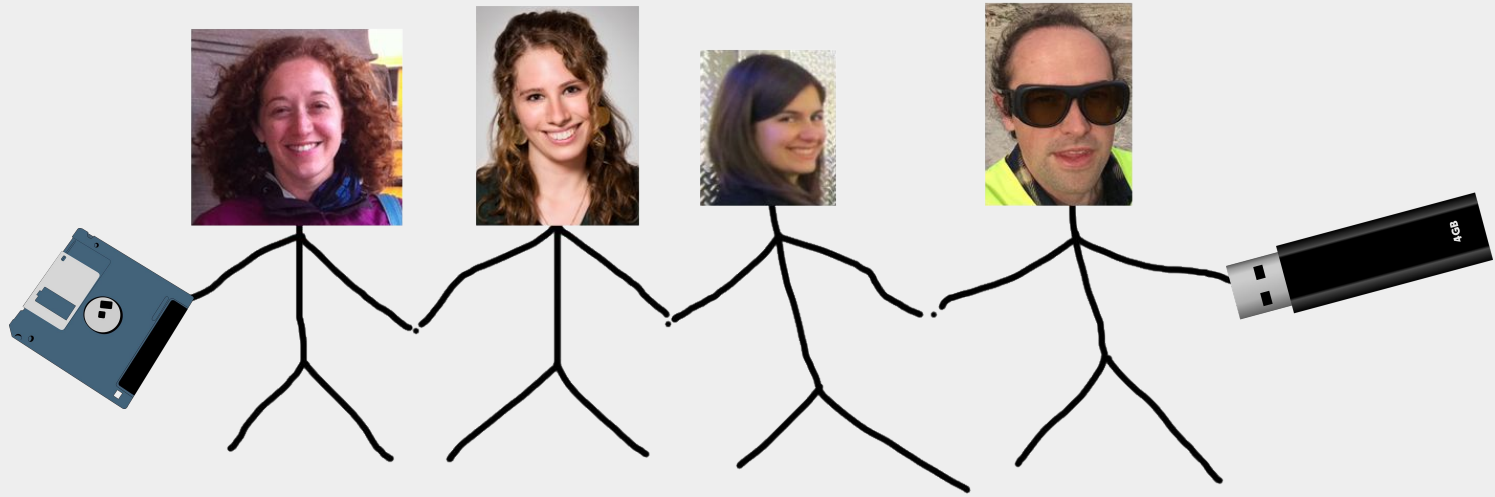
New territory...

- This is a unique context
 - Many existing rules, standards, guidelines for item-level description of digital objects
 - Very few guidelines for describing these materials in archival contexts
 - Collection-level
 - Aggregate
 - Still accessed in reading room

Familiar struggles...

- Existing UC-wide document outlining “efficient processing” practices
 - Drafted, reviewed, and accepted by another CKG
 - Few protocols for maintaining, updating, and promoting this valuable information however.
- Opportunity to learn from this history
 - Similar UC-wide goals
 - Establish shared investment in implementation
 - Clear procedures and responsibilities for making changes, and maintaining “final” version.

All these things brought us together



UNIVERSITY
OF
CALIFORNIA

Libraries

Methodology

- Compare finding aids from organizations around the world
- Goals:
 - Locate areas of relative consensus or agreement
 - Better understand how each descriptive element is being used
 - Identify approaches to description that we liked

Methodology

- Compare finding aids from organizations around the world
- Goals:
 - Locate areas of relative consensus or agreement
 - Better understand how each descriptive element is being used
 - Identify approaches to description that we liked
- Huge shoutout & thank you to CFPRT grad students [Tori Maches](#) [Scott Reed](#) & [Patricia Ciccone](#) for their work on this project!

Methodology

FINDINGS:

- SO. MUCH. VARIATION.
- Describing born-digital was essentially a boutique procedure
- This was affecting the accessibility and usability of the material being described.

Other resources consulted

- [Revised DACS Statement of Principles \(Proposed\)](#)
- [SAA-ACRL/RBMS Joint Task Force on the Development of Standardized Holdings Counts and Measures for Archival Repositories and Special Collections Libraries](#)

Other resources consulted

- [Revised DACS Statement of Principles \(Proposed\)](#)
- [SAA-ACRL/RBMS Joint Task Force on the Development of Standardized Holdings Counts and Measures for Archival Repositories and Special Collections Libraries](#)
 - *aka SAA-ACRL/RBMS JTFDSHCMARSCL*

Other resources consulted

- [Revised DACS Statement of Principles \(Proposed\)](#)
- [SAA-ACRL/RBMS Joint Task Force on the Development of Standardized Holdings Counts and Measures for Archival Repositories and Special Collections Libraries](#)
 - *aka SAA-ACRL/RBMS JTFDSHCMARSCL*
- [ISAD\(G\)](#)
- [Best Guess Guidelines for Cataloging Born-Digital Material](#)

Best Guess Guidelines

- Didn't address everything to the level of detail required
- Guidelines frequently expressed ambivalence:
 - *“For born digital archives, we are a bit torn.”*
 - *“On the one hand [...]. On the flip side, (and perhaps less problematically)...”*
 - *“Theoretically...”*

We Know What We Want...Now What?

- Prioritization
 - Had to devote time and focus to meeting, writing, research despite other priorities
 - Co-writing made this more of a possibility
- Weekly to bi-weekly meetings by phone, starting February 2017
- Identified desired sections and developed main structure of document first, then split writing up section by section

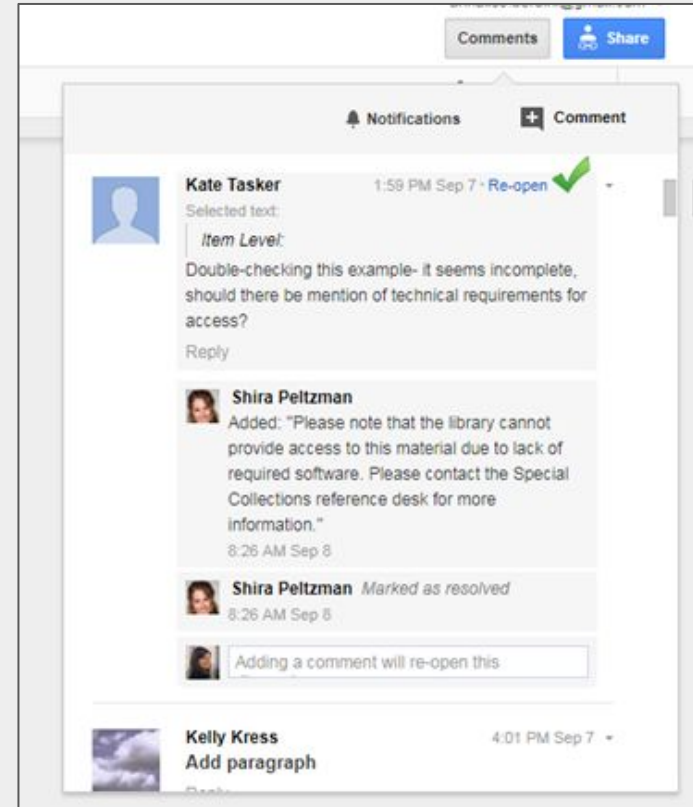
UNIVERSITY
OF
CALIFORNIA

Libraries



Google Docs!

- Working with Google Docs
 - Flexibility
 - Comment feature
 - Track edits
 - Assign tasks/email notifications
- Review and revision meetings



Review Process

- Completed first draft of document in May 2017
- First round of review:
 - Introduced document to Born-Digital CKG mid-May and gave the group one month to submit feedback and suggested edits (June 2017)
- Second round of review:
 - Born-Digital CKG members presented document to collection management colleagues and other stakeholders at all UCs
 - One month to submit feedback and suggested edits (July 2017)

Review Process

- Results of two rounds of review:
 - Additional appendices
 - Revisiting assumptions
 - UCLA Controlled Vocabulary Lightning Team
- Final step: Review and approval by UC Heads of Special Collections
 - Introduced document September 2017
 - Met on October 8 to discuss -- received full support and approval to share and implement document

The UC Guidelines for Born-Digital Archival Description

- **Detailed guidelines** for describing born-digital material

The UC Guidelines for Born-Digital Archival Description

- **Detailed guidelines** for describing born-digital material
- **Minimum descriptive requirements** for finding aids

The UC Guidelines for Born-Digital Archival Description

- **Detailed guidelines** for describing born-digital material
- **Minimum descriptive requirements** for finding aids
- **Crosswalks** to other standards

The UC Guidelines for Born-Digital Archival Description

- **Detailed guidelines** for describing born-digital material
- **Minimum descriptive requirements** for finding aids
- **Crosswalks** to other standards
- **Controlled vocab** for born-digital terms

The UC Guidelines for Born-Digital Archival Description

- **Detailed guidelines** for describing born-digital material
- **Minimum descriptive requirements** for finding aids
- **Crosswalks** to other standards
- **Controlled vocab** for born-digital terms
- **Sample finding aid**

The UC Guidelines for Born-Digital Archival Description

- **Detailed guidelines** for describing born-digital material
- **Minimum descriptive requirements** for finding aids
- **Crosswalks** to other standards
- **Controlled vocab** for born-digital terms
- **Sample finding aid**
- **Links to resources**

Extent and Physical Description?

- **37,364 digital files** (486 GB) | ([Hoover](#))
- **61 digital audio files** | ([Harvard](#))
- 0.2 linear feet and **1027 digital files** | ([UCI](#))
- **35.0 computer file(s)** 777 KB | ([Stanford](#))
- 15 Linear Feet, 5550 Physical items; **180.3 gigabytes** | ([Duke](#))
- [digital extent is not listed] | (many)

(More) Extent and Physical Description

- 344.0 boxes (318.75 linear feet and **2.8 gigabytes**) | ([University of Georgia](#))
- 66 Cubic Feet The Jan Karon papers consist of 120 legal Hollinger boxes, 3 oversize boxes, 17 oversize folders, 17 audio cassettes, 26 videocassettes, **46 disks, and 2 hard drives.** | ([University of Virginia](#))
- 6.0 linear feet and **10.84 GB** | ([University of Maryland](#))

(Even More) Extent and Physical Description

- 2 series consisting of **15 and 6 files respectively**, containing **188 WAV files in total**, approximately **33 hours in duration** | ([British Library](#))
- 136.25 linear ft. (252 boxes), 143 oversized papers (OP), 10 bound volumes (BV), 5 oversized bound volumes (OBV), 2 framed items (FR), AV Masters: 5.5 linear ft., and **5.8 MB of born digital materials** (484 files) | ([Emory](#))

Extent: Research & Recommendations

- DACS, EAD, MARC
 - [SAA/ACRL-RBMS Joint Task Force on the Development of Standardized Holdings Counts and Measures](#)
- ➔ **Born-digital extent should be described in gigabytes (GB), or fractions thereof**

Example

Hybrid Collection:

Extent: 109 linear feet

Extent: 985 GB

Physical Description: 109 linear feet (204 boxes) and 985 GB (11,905 digital files, including 17 disk images and 209 digital video files that total approximately 19 hours in duration)

Physical Characteristics & Technical Requirements

- *Where the heck is <phystech>?!!*
- Rarely used element (...but one that digital materials have been looking for all their lives...)
- Describes “the physical condition of the materials **and/or technical requirements that affect their use**” ([EAD3](#))
- Use to inform researchers about aspects of born-digital files or formats which will affect their ability to access them

Example

Physical Characteristics and Technical Requirements:

“Born-digital content is found in Series 3) Writings and Series 5) Filmmaking. These records must be accessed in the Special Collections and Archives Reading Room. Some of the files in Series 5) Filmmaking were created for Smith’s Later Days project using Avid Matador. The library is unable to provide access to Avid Matador due to the lack of hardware and software required. Please contact the reference desk for more information about viewing this material.”

Processing Information

- Goal: make digital processing actions transparent to users
- Demystify born-digital processing for others
- Link out to standard workflows or policies

```
<archdesc level="collection">
  <did>
    <unittitle>
      Upworthy records,
    </unittitle>
    <unitdate>
      2012-2014.
    </unitdate>
  </did>
  <processinfo>
    <p>
      Acidic Boxes, Piles Of Papers,  
And A Whole Mess Of Flash Drives.  
What Our Archivist Did Next  
Will Amaze You.
    </p>
  </processinfo>
</archdesc>
```

Example: Processing Information (Normalization)

“... .JPEG, .PICT, and .PSD files were normalized to the .TIFF format for preservation and to the .JPEG format for access using Archivematica v 1.5.1, according to the Library’s File Format Policy. For more information, see [*link to policy*].”

Example: Processing Information (File Renaming)

“...Many of the filenames in this collection were extremely similar [e.g. “Screenplay_Final.docx, Screenplay_Final(1).docx]. To avoid confusion, all original filenames in this collection have been appended with ‘_[Last Date Modified]’ to help researchers distinguish between drafts.”

But Wait, There's More!

Other Descriptive Elements

- Processor
- Abstract
- Conditions Governing Access
- Conditions Governing Use
- Immediate Source of Acquisition
- Appraisal Information
- Scope & Content
- Organization & Arrangement
- Container List & Inventory

Supplemental Material

- Metadata Crosswalk
- Controlled vocabulary for born-digital terms
 - Courtney Dean
 - Maggie Hughes
 - Kelly Kress
 - Shira Peltzman
- Sample finding aid



DACS
EAD3
ArchivesSpace
MARC
RDA
ISAD(G)

Using the Guidelines: UC System

- Implement guidelines at each of the 10 University of California campuses
- Provide expert, practical guidance for processing staff
- Provide a common standard throughout the UC system
- Work with Born-Digital Common Knowledge Group to share and support the guidelines

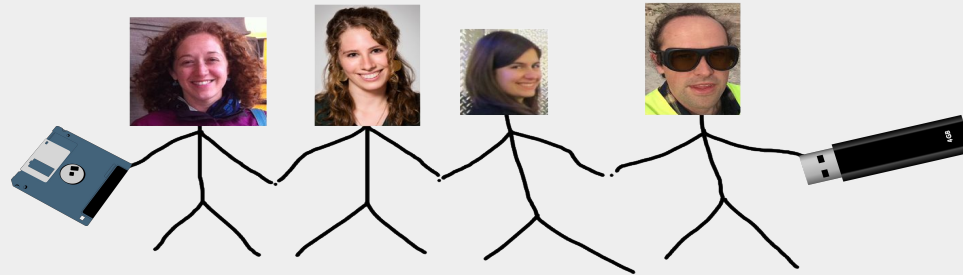
Using the Guidelines: UC System - and Beyond!

- Share the Guidelines as a companion to the widely-used Guidelines for Efficient Archival Processing in the UC Libraries
- Illustrate UC-wide practices
- Make Guidelines available for adoption and adaptation by other institutions
- Seek feedback from our professional community

Now it's your turn!

- Take a look at the full guidelines on GitHub:
<https://github.com/uc-borndigital-ckg/uc-guidelines>
- Feel free to reach out to learn more about the document, ask questions about implementation, and anything else -- you can contact us through this Google form: <http://tinyurl.com/yb5cyay7>

THANKS!



Charlie Macquarie (UCSF)

Shira Peltzman (UCLA)

Kate Tasker (UCSF, formerly UCB)

Annalise Berdini (Princeton, formerly UCSD)

🐦 @advnturelibrary

🐦 @shirapeltzman

🐦 @kate_tasker

🐦 @annarchives

UNIVERSITY
OF
CALIFORNIA

Libraries

Society of California Archivists
Annual General Meeting
Yosemite, CA. April 13, 2018
#SCAYosemite