UC Guidelines for
Born-Digital Archival Description

> **Note to readers:** *Version 1.0 of these guidelines are available on GitHub, where pull requests are welcome and encouraged.* **Please submit any written feedback to:** *Annalise Berdini, Charles Macquarie, Kate Tasker, & Shira Peltzman.*

Table of Contents:

Developed by

Annalise Berdini, UC San Diego
Charles Macquarie, UC San Francisco
Shira Peltzman, UC Los Angeles
Kate Tasker, UC Berkeley

With Contributions from

Controlled Vocabulary Lightning Team:

Courtney Dean, UC Los Angeles
Margaret Hughes, UC Los Angeles
Kelly Kress, UC Los Angeles

Document Revision History

A draft of the Guidelines was sent out to the UC Born-Digital Common Knowledge Group (CKG hereafter) for comment and feedback on May 15th, 2017. After a three week-long comment period, the document was revised to incorporate the CKG's feedback. On June 27th, 2017 the revised draft was sent out to everyone throughout the UC system involved in collection management via CKG representatives from each campus. This three week-long comment period yielded a series of suggested revisions that were incorporated into the document. As part of these revisions, a Lightning Team was asked to develop a controlled vocabulary for commonly used born-digital source media and related terms.

The UC Guidelines for Born-Digital Archival Description version 1.0 was passed along to the Heads of Special Collections Common Knowledge Group for final review and approval on September 11, 2017.

INTRODUCTION

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. The result of this work is a set of guidelines for creating and updating finding aids to include born-digital archival material.

Currently there is no widely accepted descriptive standard that adequately addresses born-digital archival material. Although *Describing Archives: A Content Standard* (DACS) is meant to apply to all material types, it is lacking in guidance addressing the various practices of digital processing and description as they have emerged at the participating UCs.[1] As a result, institutional practices for creating finding aids vary substantially throughout both the UC System and the field at large. The lack of clarity and agreement in this area has resulted in a proliferation of finding aids that fail to accurately express the quality, quantity, and usability of digital material. Not only does this diminish accessibility, it is also inefficient; each organization must 're-invent the wheel' when it comes to describing born-digital archival content.

Background and Methodology

As of January 2017, four Digital Archivist positions had been created within the UC system at UC Berkeley, UCLA, UC San Diego, and UC San Francisco. Recognizing common workflows, challenges, and questions, these archivists began scheduling monthly phone calls to exchange ideas and brainstorm solutions. The challenge of integrating description for born-digital archival materials immediately stood out as a particularly complex issue at all campuses and became the focus of discussion.

In early 2017 the group decided to tackle this challenge with the goal of producing a shared set of guidelines on integrated description of UC born-digital archival material. The group scheduled weekly web conferences to identify elements of born-digital description; map these elements to existing standards, including DACS, EAD, and MARC; review related work such as the UC Guidelines for Efficient Archival Processing and the UK Archives & Records Association Descriptive Standards Roundtable's "Best Guess Guidelines for Cataloguing Born Digital Material"; examine and evaluate a wide variety of online finding aids at other institutions; consult

---

[1] For example, DACS 4.3, "Technical Access", outlines metadata creation practices which assume that patrons would be required to access digital materials using the original hardware which created that material, but the field of digital archives has gone in an almost entirely different direction than this. Additionally, DACS 7.1.8, "Processing Information Note" (within "Notes" element), contains rules applying most closely to the EAD <processinfo> element. This is arguably the most important descriptive element for born-digital material, however DACS does not mention digital material at all in this section.

local processing manuals and workflows; and discuss current approaches to born-digital description.

Eight months of rigorous discussion and collaborative work produced these guidelines, which are based on existing descriptive standards, emerging best practices for born-digital materials, and archivists' practical considerations.

Goals

The UC Guidelines for Born-Digital Archival Description will help the UCs:
- Evolve current descriptive and processing practices
- Raise awareness and deepen understanding of the unique challenges and requirements that born-digital material presents to archivists
- Suggest a set of baseline descriptive requirements for born-digital material
- Align descriptive practices throughout the UC system
- Improve the overall quality and usability of finding aids
- Enhance intellectual control

How to Use These Guidelines

These guidelines were designed to be used as an extension of the Guidelines for Efficient Archival Processing in the University of California Libraries.[2] However, there is a key difference between these documents: while the latter provides recommendations to achieve a "golden minimum" by determining the appropriate processing level (from minimal to highly intensive), these guidelines do not attempt to define processing levels for born-digital materials. [3] Instead, recognizing that there is currently no widely accepted method for applying descriptive standards to born-digital archival materials, the authors present comprehensive details for each descriptive element, to help establish norms and provide a thorough overview of the process.

The highly-detailed guidelines are not intended to prompt or promote highly intensive processing for all collections. Each section establishes a required minimum baseline for archival description, as well as recommended and optional components of each descriptive element, with the expectation that users will tailor these guidelines to fit the needs of their given

---

[2] Bachli, K., Eason, J., Light, M., McAnnaney, K., Morrison, D., Seubert, D. "Guidelines for Efficient Archival Processing in the University of California Libraries," Version 3.2, University of California Libraries. (2012): 1-55.
http://libraries.universityofcalifornia.edu/groups/files/hosc/docs/_Efficient_Archival_Processing_Guidelines_v3-1.pdf

[3] There are no standard guidelines for determining appropriate processing levels and processing tasks for born-digital materials, although a Digital Processing Working Group has been established to produce a processing framework for this purpose.

collections, paring them back or scaling them up where necessary. Components listed as "required" are intended to serve as a UC-wide baseline. Some will be listed as "required" even if there is no change in their use between physical and born-digital materials (meaning that this document does not add any new information to the instructions for using the descriptive element). Examples of multi-level description were composed with the intention that these can be customized to best meet the needs of any institution or collection, regardless of the level to which a given collection is processed.

In the past, institutions both within and outside the UC system have taken an ad hoc approach to born-digital processing. While some of the included fields or instructions in these guidelines may differ from repositories' existing descriptive practices, this guide has been developed for purpose of unifying and aligning born-digital description standards across UC campuses. Along with the Guidelines for Efficient Archival Processing in the University of California Libraries, the UC Guidelines for Born-Digital Archival Description will establish the foundation of archival processing and descriptive practices throughout the UC system and beyond.

Systems of Record

Stewards of born-digital material rely on metadata. This information is captured at various points in the preservation lifecycle, according to different schemas, by different tools, and is used for distinct purposes. Preserving born-digital material in the long-term requires collection managers to coordinate the strategies and systems used to collect, store, and disseminate this information over time.

These guidelines do not attempt to outline all necessary preservation actions or metadata requirements for managing born-digital archival material. Rather, they focus exclusively on describing born-digital collections material in the context of a finding aid so that archival collections can be discovered and used.

Finding aids exist to help researchers determine whether information within a collection is relevant to their research and readily accessible. They are not and should not be used as the system of record for all the descriptive and technical metadata generated during processing. In many cases, metadata will be generated during processing that is ultimately important in the context of a collection management, but inessential and perhaps inappropriate within the context of the finding aid. An example of this would be a complete record of the preservation events that occur during processing. This information is clearly important, but including it in a finding aid in its entirety--particularly for collections with larger digital components--would quickly render the finding aid top-heavy and jargon-filled. However, some of this information can provide clarity and context to researchers as they first encounter born-digital material in collections. It is up to the archivist to make a determination about where to draw the line between providing an appropriate level of detail and too much information.

With that in mind, one solution to this dilemma is to document this information elsewhere and provide readers with a link that would allow them to access this information if they wish. For instance, institutions should consider using standardized language in the Processing Information section that includes a link to their institutional standards and policies for processing born-digital collection materials rather than trying to outline them in the finding aid. (See, for example, the Bentley Historical Library's Digital Processing Note[4]). This increases transparency, makes the process of creating finding aids more efficient, and helps researchers better understand how a collection was processed.

Other instances in which archivists should consider linking out to institutional policy or practice include, but are not limited to:
- File formats accessible onsite
- Processing practices or standards
- Digital archives or digital preservation policies
- Policies pertaining to personally identifiable information (PII), protected health information (PHI)

Level of Description

For wholly unprocessed collections (whether hybrid or purely born-digital) the authors recommend following the UC Guidelines for Efficient Archival Processing (especially section 4.G) to determine an appropriate level of description.[5]

However, many special collections departments continue to grapple with a large backlog of unprocessed born-digital materials which have been separated from processed paper-based materials of the same provenance. Working through this backlog to reunite born-digital content with parent collections will be an ongoing process over a number of years, so specific guidelines addressing these types of materials are also needed.

For unprocessed born-digital materials which are part of a partially-processed hybrid collection, it is helpful to review the existing level(s) of description for paper-based material. Depending on the collection's complexity and size, can born-digital content be easily integrated into this finding aid structure? If paper-based material is fairly homogenous and is sufficiently described at the series level, consider if the same level of description can be helpfully applied to born-digital

---

[4] Bentley Historical Library Digital Processing Note (2013). https://deepblue.lib.umich.edu/bitstream/handle/2027.42/93344/ProcessingNote_20131312.pdf?sequence=4 (Version 2.0).
[5] Guidelines for Efficient Archival Processing in the University of California Libraries (2012). http://libraries.universityofcalifornia.edu/groups/files/hosc/docs/_Efficient_Archival_Processing_Guidelines_v3-1.pdf

material. Refer to standards for multilevel description found in ISAD(G)[6], DACS, and the Guidelines for Efficient Archival Processing in the University of California Libraries, section 4.G.

For example, if a collection includes a series of regular annual reports in both paper and electronic form, it may be more efficient to describe this content together at the series level rather than listing each report separately. If file lists or directory lists are useful and can be easily generated and attached to the finding aid, then they should be included as attachments. For the purpose of these descriptive guidelines however, the authors differentiate between a file list and archival description. A key function of archival description is intellectual control. This should be carried out at whatever level is warranted by a given collection. The level of description should not necessarily be tied to the question of whether or not a file list can be generated with ease. See the Container List & Inventory section of this document for more information.

DESCRIPTIVE ELEMENTS

Processor

DACS: Processing Information in Notes Element, 7.1.8
EAD3: Processing Information <processinfo> ; Control <control> (replaces <eadheader> from EAD 2002)
ArchivesSpace: Processing Information Note, and/or Collection Management -> Processors (internal field; not exported in EAD)
MARC: N/A
ISAD(G): 3.7.1
RDA: N/A

REQUIRED
If the digital materials were processed either at a later date and/or by a different person than the rest of the collection, specify when and by whom they were processed.
- Ex. **"Digital materials processed by [processor] in [year]"**. If processed by a student, consider including, "**under the supervision of [supervisor]**".

Some organizations may have a local practice of recording the processor's name and the period of time over which that processing took place in the front matter of the finding aid. In that case, use the EAD header <author> element. The header element <author> can also be used to note the actual author of the finding aid if this differs from the processor, and can be helpful to note

---

[6] International Council on Archives. (2000). *ISAD(G): General International Standard Archival Description* (2nd ed.).

additions where new description has been added to an existing finding aid. [7] Note that this element contains information on the processor only, and that more specific information on the Processing Information should be contained within the "Processing Information" element, which is outlined in the "Processing Information" section of this document.

Examples:

Susan Sontag papers, circa 1933-2004

2002 installment processed by Lorain Wang and Catherine Lee, August 2002; 2005 installment processed and integrated with 2002 installment by Lauren McDaniel, February 2007-April 2008; 2012 installment processed by Mitchell Erzinger, October 2013-March 2014; **digital materials processed by Lori Dedeyan in 2014.**

Barbara Jones papers, 1942-2007 (bulk 1945-1986)

Processed by Timothy Smith in 2011. **Digital materials processed in 2016 by Victoria Maches and Scott Reed in the Center for Primary Research and Training (CFPRT), under the supervision of Shira Peltzman***.*

Physical Description and Extent

DACS: Extent Element 2.5
EAD3: Physical Description <physdesc>
ArchivesSpace: Physical Description Note
MARC: 300
ISAD(G): 3.1.5
RDA: Extent 3.4

REQUIRED

Physical Description[8] must include units of measure for born-digital material. This should *always* include both the size of the digital material in GB as well as the total number of files that have been preserved. Use 'GB' ' instead of "Gigabytes", "gigabyte", "Gb", "GBs", or any other

---

[7] These guidelines distinguish "Processor" as a separate descriptive element from "Processing Information". Note that there is no corresponding "Processor" element in DACS or EAD. Information about the processor may be publicly recorded in the <author> element of the EAD header, or privately noted in internal collection files or collection management systems. ArchivesSpace (version 2.0) provides a staff-only field labeled "Processors" in the Collection Management sub-record.

[8] "Physical Description" is used here to mean a brief narrative summary noting the type and number of containers present, their physical attributes and/or dimensions, or other information related to their size, shape, or appearance. Please note that "Physical Description" and "Extent" actually map to two *separate* EAD3 elements: <physdesc> and <physdescstructured>. These two elements are independent of each other, and are differentiated in order to record different kinds of information about physical description. See https://www.loc.gov/ead/EAD3taglib/index.html#elem-physdesc and https://www.loc.gov/ead/EAD3taglib/index.html#elem-physdescstructured

variations thereof. Consult a digital data storage conversion tool if necessary, such as: http://www.thecalculatorsite.com/conversions/datastorage.php

When calculating size, round to three decimal points only when content is less than 1 GB. If content is smaller than 1 MB, default to ".001 GB". Otherwise, round to two decimal points (ex. 9.25 GB).

For unprocessed material where capacity is unknown or difficult to estimate, include a count of the unprocessed media formats.

> Example:
> **"Physical Description: 3 unprocessed hard drives (100 GB, 3000 GB, and 1000 GB) and 14 unprocessed compact discs."**

In certain cases, processors can also include other units of measure that may help a researcher better gauge or contextualize the amount of digital material present in the collection. This could include, for instance, total runtime or duration (for audiovisual files), total number of disk images, or total number of emails.

> Examples:
> *"**Physical Description:** 4.5 linear feet (6 oversize boxes, 1 manuscript box), **and 3400 GB (37,364 digital files)**"*
>
> *"**Physical Description:** 13.4 linear feet (26 document boxes, 10 half document boxes, and 1 oversize flat box), **and 385 GB (12,938 digital files)**"*
>
> *"**Physical Description:** 19.5 linear feet (29 boxes) **and 3750 GB (58,439 digital files, including 20,879 WAV files that total approximately 75 hours in duration)**"*
>
> *"**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 Description:** .5 linear feet (1 box) and **113 GB (1,097 WARC files representing periodic crawls of approximately 193 websites)**"*
>
> *"**Physical Description:** 17 linear feet (25 boxes) **and .011 GB (31 digital files)**"*

Extent

DACS: Extent Element 2.5

EAD3[9]: Structured Physical Description <physdescstructured>
ArchivesSpace: Extents
MARC: 300
ISAD(G): 3.1.5
RDA: Extent, 3.4; File size, 3.19.4

REQUIRED IF APPLICABLE
Additionally, record a quantitative measurement of processed digital content in an Extent statement. This measurement should be recorded separately from any Extent statement(s) for physical material. Use parallel Extent statements, if necessary. At minimum, record the extent of digital content in gigabytes (GB). Optionally, add an Extent statement to record the number of digital files.

> Examples:
> **Hybrid Collection:**
> *Extent: 109 linear feet*
> *Extent: 985 GB*
>
> ─────────────────────
>
> **Multiple Extent Statements for a Completely-Digital Collection:**
> *Extent: 3750 GB*
> *Extent: 58,439 digital files*

Abstract

DACS: Scope and Content Element, 3.1
*EAD3: Abstract *
ArchivesSpace: Abstract Note
MARC: 520
ISAD(G): N/A
RDA: Summarization of Content, 7.10

RECOMMENDED
If there is significant born-digital material present in a collection, the Abstract should reflect this. Try to differentiate the descriptive language in the Abstract from the Scope and Content note to avoid redundancy if you will be using both elements separately. For instance, if you have listed

---

[9] Though the Extent element (<extent>) has been completely deprecated in EAD3, for the purpose of this document the authors will continue to use the language "Extent" and give that name to the descriptive element outlined here. As descriptive standards change the authors may also change the terminology of this document, but for now the authors will continue to use the terminology that is most widely-shared among the community, which is "Extent."

the formats present in a given collection in the Scope and Content note, there is no need to repeat that information in the Abstract.

> Examples:
>
> *"[...]The collection includes the following: personal and professional correspondence **including email**; journals; documents relating to her political activism**; and digital audiovisual recordings related to her work at the Humane Society."**
>
> *"Jane Smith is an engineer, scientist, and professor based in Los Angeles. The collection consists of sketches, correspondence, published and unpublished teaching materials, **email correspondence, and born-digital datasets related to her research on radio frequencies**."*

Conditions Governing Access and Restrictions

DACS: Conditions Governing Access Element, 4.1
EAD3: Conditions Governing Access <accessrestrict>
ArchivesSpace: Conditions Governing Access Note
MARC: 506
ISAD(G): 3.4.1
RDA: Restriction on Access, 4.4.

REQUIRED[10]
Exactly as described in DACS, this element provides information about access restrictions on files due to the nature of the information in the materials being described. For the purpose of digital collections, use of this field does not change. If there are any restrictions or conditions governing access, this should be noted at every level to ensure that researchers are aware of them.

> Examples:
> **Example #1**
> *"Open for research, with the following exceptions: Boxes 136 and 137 include hard drives that contain files of Smith's digital journals, which are restricted until 25 years after Smith's death (September 15th, 2037). Box 352 of medical files are restricted until 50 years after Smith's death (June 11, 2061). Certain digital files are restricted until June 2056."*
>
> **Example #2**

---

[10] Please note that this field is always required for institutions submitting finding aids to the OAC: http://www.cdlib.org/services/access_publishing/dsc/contribute/docs/oacbpgead_v2-0.pdf. If there are no access restrictions on digital content, add a note such as "Collection is open for research."

"The UCSF Archives and Special Collections policy places access restrictions on material with privacy issues for a specific time period from the date of creation. Access to records that contain personal and confidential information about an individual or individuals is restricted for 75 years from date of creation or until the death of the individual mentioned in the records, whichever is longer. Medical records are restricted for 50 years after an individual's date of death, if known. If the date of death is unknown, access is restricted for 100 years from the individual's date of birth or 100 years from the date of record creation, whichever occurs first.

Some audio recordings have been restricted to protect personal identifiable information or Protected Health Information (PHI) in accordance with federal regulations and UCSF Archives and Special Collections privacy policies. Restrictions are noted at the item level. Please contact the Head of Archives and Special Collections for more information."

Example #3
"By donor request, permission to access materials on disk mss0270d01 -- Dell OptiPlex 755, 2007-2015 will only be granted after review by The McGuinness Foundation, effective until 2050. Please contact the Special Collections and Archives Reference Librarian for access inquiries."

Conditions Governing Reproduction and Use

DACS: Conditions Governing Reproduction and Use Element, 4.4
EAD3: Conditions Governing Use <userestrict>
ArchivesSpace: Conditions Governing Use Note
MARC: 540
ISAD(G): 3.4.2
RDA: Restriction on Use, 4.5
RDA: Restriction on Use, 4.5

RECOMMENDED IF APPLICABLE
For the purpose of describing born-digital materials, this element should be used in the same way as to describe physical materials. Follow existing directions laid out by DACS and/or EAD, and refer to your organization's own policies and procedures for reproducing or using digital or physical material.

Please note that this field is <u>always</u> required for institutions submitting finding aids to the Online Archive of California.[11]

---

[11] http://www.cdlib.org/services/access_publishing/dsc/contribute/docs/oacbpgead_v2-0.pdf

Physical Characteristics and Technical Requirements

DACS: Physical Access Element, 4.2 ; Technical Access Element, 4.3
EAD3: Physical Characteristics and Technical Requirements <phystech>
ArchivesSpace: Physical Characteristics and Technical Requirements Note
MARC[12]: 336-338 ; 347 ; 538
ISAD(G): 3.4.4
RDA: Media Type, 3.2 ; Carrier Type, 3.3 ; Digital File Characteristic, 3.19 ; Equipment or System Requirement, 3.20.

REQUIRED IF APPLICABLE (see below)
The Physical Characteristics and Technical Requirements is an essential field for digital collections that combines information from the EAD tag <phystech> and the Physical Access (4.2) and Technical Access (4.3) DACS fields, which are sub-fields of Conditions Governing Access. Use this field to inform researchers about the physical or technical characteristics of digital materials that will affect their ability to access them. This is especially important for open-access collections in which users may want to interact with digital materials on their own home computers, or in cases where the storage media or format cannot be made accessible on library computers.

Physical Characteristics and Technical Requirements is also the field in which to note that users may need to contact reference staff to make use requests in advance of their visits. This is important for large, complex collections or collections with many varied software and hardware dependencies. For more complex collections where physical and/or technical requirements impact use, it is required that this field be added to the finding aid for more complete description.

If any portion of a collection contains digital material that cannot be readily accessed by researchers, then a PhysTech note must be placed at every level of the collection (ie series, box, folder, etc.) to notify researchers that this is the case. This ensures that researchers will see the requirements necessary to access material. Notation at each level could also be helpful to reference staff, as it can act as a reminder of the special requirements for serving up born-digital records. See below for note examples at each level.

REQUIRED IF APPLICABLE

---

[12] Multiple MARC fields may be appropriate and useful for recording aspects of physical characteristics and technical requirements. The authors suggest 336 (Content Type), 337 (Media Type), 338 (Carrier Type), and 538 (System Details Note). MARC field 347 (Digital File Characteristics) may be useful for single-item digital collections or relatively homogeneous collections where it is possible to manually describe all file types and encoding formats. Note that it can become infeasible to manually record all file types and formats in repeating 347 fields for large collections with dozens or hundreds of file types. See *MARC 21 Format for Bibliographic Data*, Library of Congress: https://www.loc.gov/marc/bibliographic.

- **Unprocessed materials**
  Collections may contain some digital media that cannot be processed. This could be due to issues with the media itself, such as damage or degradation that prevents it from being imaged. It could also be due to technological constraints, such as the lack of appropriate technology to read, transfer, or preserve certain formats in-house (for instance, SyQuest disks, data punched cards, or 8 inch floppy disks).
  - If portions of the collection include media that cannot be read or transferred in-house, note that the library cannot provide access to this due to lack of required hardware. You can include the series/box/folder or use the phrase "content includes", etc. to clarify the scope and contents of the accessible media, if known.
    - Ex. *"The library cannot provide access to all media formats in Series 3: Creative Writing due to lack of required hardware. For more information, please contact the reference desk."*
  - If a collection contains unprocessed born-digital material that will be restricted, note that the material is awaiting processing in the finding aid. Also include, "If interested in viewing this material, please contact the reference desk."
    - Ex. *"The library cannot provide access to all media formats in Series 1: Filmmaking due to lack of required hardware. Unprocessed content includes edit decision lists (EDLs) and source media files from a variety of projects spanning 2004-2008. For more information, please contact the reference desk."*
  - If materials are unprocessed due to other non-technical constraints, note this in Conditions Governing Access rather than in this field.

- **Processed material that cannot be readily accessed**
  All files have external dependencies including the format into which a file has been encoded, the software required to render it, the operating system on which that software is able to run, or the hardware on which a particular operating system runs. Some collections may include files that have been preserved (perhaps at a bit-level) but that cannot be readily rendered or accessed by researchers because of one or more of the above dependencies. If this is the case, be sure to include a note in the finding aid explaining that the library is unable to provide access to all file formats due to lack of required hardware/software/system specifications.
  - Ex. *"Please note that the library is not able to provide access to the donor's scheduling software, Movie Magic, due to lack of required software. If interested in viewing this material, please contact the reference desk for more information."*

- **Special instructions for viewing digital materials**
  If the procedure for requesting access to digital material differs in any way from the procedure for requesting paper-based collections, it is important to include this information in the finding aid.

○ Ex. ***"Access to digital materials is available on-site in the Special Collections reading room. Requests to access born-digital materials must be made at least 2 days in advance. Once request has been made, please confirm viewing arrangements with Special Collections reference desk staff."***

RECOMMENDED IF APPLICABLE
- **Hardware and system requirements** (Operating systems, file systems, RAM requirements, media readers, peripherals, etc.)
  A list of required hardware to access born-digital files in a collection may be useful to researchers, and even to reading room staff. If a particular collection requires certain tools to access content, however rarely, it should be noted.
  - Ex. ***"Macintosh Computing Platform with OS System 7.0 or Higher."***

- **Software or other rendering requirements**
  A list of the required software to access the born-digital files in a collection will allow the user to determine whether or not they can access materials at home (for open collections) or will need to access them in a reading room. Note whether or not the library has the necessary software to access all files. (This may overlap or coincide with special instructions for viewing digital materials, as detailed below.) Include any major software requirements, including version and creator.
  - Ex. ***"Microsoft Word 2013, HandBrake Version 1.0.1 - 64 bit OS."***

- **Storage location** (Extra retrieval time required)
  This condition could apply to off-site storage OR to collections in which files must be transferred from preservation media to access copies. (E.g. Born-digital video stored on optical media that must be extracted for access.) See also "Special instructions for viewing digital materials", below.

  Examples:
  Collection-level:
  "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.

  Requests to access any born-digital files must be made at least 2 days in advance. Once a request has been made, please confirm viewing arrangements with Special Collections reference desk staff."

  Series-level:

"Series 10) Digital Media is currently awaiting processing. This series includes 3.5 inch Mac-formatted floppy disks, 5.25 inch Mac-formatted disks, and one Mac hard drive, mostly containing Word and Notepad files. If interested in viewing this material, please contact the reference desk. Please note that the library cannot provide access to all file formats due to lack of required software and/or system specifications."

Series-level:

"Series 3) correspondence contains email that can be viewed using ePADD software. This materials must be accessed in the Archives & Special collections reading room. Requests for viewing this materials must be made 2 business days in advance, please contact Archives staff to schedule an appointment or for more details."

Box-level:

"Box 113 includes a hard drive that contained Jorgenson's digital video project files for her 2011 documentary 'Swindled', which was edited using Final Cut Pro version 7.0. Although these files have been preserved by LSC, the library cannot provide access to all file formats due to lack of required software and/or system specifications. If interested in viewing this material, please contact the reference desk for more information."

Box-level:

"Box 12 contains Jacobson's personal digital assistant, the contents of which have not been processed due to current technical limitations. Unprocessed digital materials are not available for researcher use, and so this item is currently restricted. Please contact the reference desk for more information."

Item-level:

"Floppy disks ucla_ua_545_49_001 through ucla_ua_545_49_027 contain coded data utilizing DbaseIII software, which was initially developed and used to digitally sort, classify, code, manage, and facilitate retrieval of extensive government and research documents. Case 1 (ten 3.5 inch floppy disks; ucla_ua_545_49_001 through ucla_ua_545_49_012) appears to contain "Input Data" and two (2) program disks. Case 2 (15 3.5 inch floppy disks; ucla_ua_545_49_013 through ucla_ua_545_49_027) appears to include backups of index, inputs, and data files; also contains yellow sheet with notations regarding display structure. Please note that the library cannot provide access to this material due to lack of required software. Please contact the Special Collections reference desk for more information."

Immediate Source of Acquisition

DACS: Immediate Source of Acquisition Element, 5.2
EAD3: Acquisition Information <acqinfo>

ArchivesSpace: Immediate Source of Acquisition Note
MARC: 541
ISAD(G): 3.2.4
RDA: Immediate Source of Acquisition of Item, 2.19

OPTIONAL
Record general information about the acquisition of born-digital material, such as the source, date, and type of acquisition, in an Immediate Source of Acquisition Note. However, since acquisition of born-digital content often involves technical processing and data capture after media has been physically transferred, the specific methods and processes of born-digital acquisition and data capture should be detailed in a Processing Information Note. Refer to the Processing Information section in this document. Do not enter specific technical details in Immediate Source of Acquisition.

> **Example:**
> *The digital files from the Reginald E. Zelnik papers were donated to the Library by Elaine Zelnik on January 21, 2011, on six 5.25 inch floppy disks. The floppy disks were disk imaged on March 16, 2015 to create master preservation copies and to extract the content files.*

Appraisal Information

DACS: Appraisal, Destruction, and Scheduling Information Element, 5.3
EAD3: Appraisal Information <appraisal>
ArchivesSpace: Appraisal Note
MARC: 583
ISAD(G): 3.3.2
RDA: N/A

OPTIONAL
Note any general information about actions relating to appraisal, deduplication, or weeding of digital files. Refer to or link to library policy if applicable. Do not include specific technical details about the process of de-duplication or weeding in Appraisal Information. Use the Processing Information Note to provide additional information if necessary.
- Ex. ***"Temporary and deleted files were removed from this collection, according to the Library's digital preservation and privacy policies."***

Scope and Content

DACS: Scope and Content Element, 3.1

EAD3: Scope and Content <scopecontent>
ArchivesSpace: Scope and Content Note
MARC: 520
ISAD(G): 3.3.1
RDA: Summarization of Content, 7.10

RECOMMENDED
The Scope and Content note should address the 'who, what, where, when, why, and how' of digital materials. The level and kinds of detail included in the Scope and Content note will depend largely on the collection and the nature and extent of the materials being described. Therefore, it's possible that certain information in the Scope and Content note may be repeated elsewhere in the finding aid such as Processing Information, Extent, etc. Some repetition is fine, but the focus of this section is more about the creation of the records than technical access requirements. The level or levels at which scope and content notes are applied should be chosen according to existing guidelines laid out in DACS, which should be the guide for any other scope and content questions not covered in this document as well.

Oftentimes, a processor will be updating or editing an existing Scope and Content note written by the original processing archivist. In cases like this, try to balance the tone and description of the born-digital content with the existing description provided for the analog material, to maintain a co-authorship with the original processor.

According to DACS 3.1, the Scope and Content note briefly details how the materials were created. For digital materials, this might include the type of computer, software, other significant media, operating systems, version numbers, etc., and how many years of the donor or content creator's digital life are included amongst the data. In this vein, an important consideration is the method used to determine when digital material was created. Please see "Date range of the materials" bullet point below for specific guidance on this subject.

For larger, more complex collections, breaking down numbers of files by format might be helpful.
- Ex: "**3,498 documents (.pdf), 4,312 images (.jpg), and 29 spreadsheets (.xlsx)**."

Consider briefly describing any highly proprietary software or hardware or unusual creation techniques. (E.g. art made on fax machines; digital video art that includes errors and glitches inherent to the work).

In general, when noting specific hardware or software, be as clear as possible about what the software/hardware is -- it might not exist in 5, 10, or 20 years, and even if it does, users might not have any point of reference for it (but use your judgement--what matters in the context of the

collection and the creator?). Consider including its primary function, software provider/vendor, whether proprietary or open-source, software version, date if known.[13]

RECOMMENDED

- **Description of functions, processes, or activities that led to creation of materials themselves**
  Keep in mind that you should focus on the materials themselves, not the context in which they were created. That information will be recorded in the Administrative/Biographical note. However, in the case of born-digital records, this could be a helpful section in which to describe the operating systems, hardware, software, and significant media used to create the materials being described. This differs from the Physical Characteristics and Technical Requirements note in that it pertains to the creation of the records, not to the access requirements and limitations pertaining to the collection as it exists in the archives.

- **The documentary form(s) or types of records being described**
  This may include correspondence, minutes, reports, data sets, web pages, videos, photographs, etc. Be sure to consider adding file format types in a large, complex collection (.jpg, .pdf, .xlsx) and the number of files.

- **Date range of the materials**
  Although DACS advises processors to use the date(s) of creation, this is not always straightforward for digital material.[14] Exercise caution when assigning dates to digital material, and include in the Scope and Content note how dates were determined and/or assigned: Date Created? Date Last Opened/Modified? This can be confusing when a donor has migrated materials from one computer or hard drive to another, or in cases where an error has altered the Date Created to erroneously read "December 31, 1969".[15]

---

[13] At the time of writing, the IMLS-funded Software Preservation Network (http://www.softwarepreservationnetwork.org), the 2008 JISC-funded study 'The Significant Properties of Software' (https://www.webarchive.org.uk/wayback/archive/20100624233431/http://www.jisc.ac.uk/media/documents/programmes/preservation/spsoftware_report_redacted.pdf) , and the Yale University Library CLIR/IMLS-funded project to explore Wikidata as a digital preservation knowledge base (http://openpreservation.org/blog/2016/09/30/wikidata-as-a-digital-preservation-knowledgebase/) were helpful resources.

[14] DACS chapter 2.4 instructs processors to describe a date or date range using the date(s) of creation, but does not yet address dates of born-digital material. The "creation" date recorded by file systems can be easily and inadvertently changed simply by copying a file from one location to another. The "last modified" date persists when files are copied, and arguably could be considered the more precise date, as it indicates when the last changes were made.

[15] This date commonly appears because it refers to the Unix operating system's epoch date. See http://mentalfloss.com/article/26316/why-does-my-gadget-say-its-december-31-1969.

Prefer the "last modified" date of the files if possible, as this is generally more reliable than the "created" date. Use the "created" date if "last modified" is not available.

- **Authenticity and related content**
  Although the concept of authenticity has always played an important role in archives, given digital media's inherent vulnerability and the ease with which files can change, be duplicated, or become corrupt, it is an especially crucial aspect of born-digital processing and preservation. Broadly defined, authenticity refers to digital material being, "what it purports to be."[16] Establishing authenticity for digital material can be a complex and multifaceted process. Although local policies and requirements will vary throughout the UC system, it is important to include information about the authenticity of digital material if this information has been gathered and recorded. This might include retaining technical details associated with the file's creation, arrangement, or acquisition, and/or recording details associated with the file's accompanying metadata such as the assignation of Universally Unique Identifiers (UUIDs), hash values or checksums.
  - Ex. ***"Smith's digital material relates to various writing projects she undertook between 1989-2006. These project files include research and notes, drafts, edits, and supplementary material. Smith routinely backed up her old computer files onto newer machines, which resulted in some file duplication across systems. Duplicates have been retained to maintain original order, but to avoid confusion, all files were assigned Universally Unique Identifiers (UUIDs) and checksums during processing, and their 'Last Modified' dates have been preserved and retained."***

DO NOT INCLUDE:
- Information about the context in which the material being described was created, used, "and so on" (put this in Administrative/Biographical History).
- Information about gaps in the material resulting from archival appraisal decisions (i.e. weeding; put this in Appraisal note).

Examples:
Example #1 Collection-level (hybrid collection):
Sample text from Duke University: *"Collection Overview: The Stephanie Strickland Papers include printed journals and anthologies featuring Strickland's poetry, programs and posters from Strickland's publications and performances; school materials from high school to graduate school; articles and anthologies; electronic media (removed); videotapes; and other miscellaneous materials. There are source materials from works such as Zone : Zero, True North, Sand Soot, and Vniverse; conference programs; journals; and gallery catalogs and other sources used in her work, as well as files, proofs, and drafts from Strickland's V : Vniverse, True North, Zone : Zero, Red Virgin,*

---

[16] http://www.dpconline.org/handbook/glossary#A

Give the Body Back, V : WaveSon.nets / Losing L'una, Dragon Logic and other Strickland poems, essays, and compilations. There is a small amount of print correspondence and event material. There is a larger amount of electronic correspondence. Electronic media has been separated from the accession and housed on Duke's electronic records server but has been described in this collection guide. There are several VHS and cassette recordings of interviews with Stephanie Strickland and related to her work." (Duke)

Example #2 Series-level (hybrid collection):
[Sample text from UC Irvine](#): *"This series comprises Richard Rorty's electronic word-processing files. Included are letters, many drafts of writings, lecture notes, syllabi, and exams. Also included are bibliographies of his work, a few administrative files, and a couple of documents relating to his children. This series also contains drafts of writings done by his peers and colleagues. Researchers will notice some data loss and corruption has occurred in some of the earliest files."*

Example #3 Series-level (hybrid collection):
[Sample text from UC San Diego](#): *"Series 21) Born-digital material consists of computer drafts of Mac Low's major works, including Forties and Steins, email files, poetry sourced from other authors, including Jim Cayley and Charles Doria, and custom computer programs written to generate poetry, mostly for MS-DOS. Files have been renamed to correctly reflect the format in which they were created. Formats mostly include Word documents (.doc), Outlook email files (.pst), and text documents (.txt). These were all created on Mac Low's Power Macintosh 6100 and originally stored on 3.5 inch and 5.25 inch floppy disks."*

Example #4 Sub-series level (hybrid collection):
"The material contained in this sub-series consists of documents, graphics, spreadsheets, databases, software programs, video, audio, and other material created and/or stored on computers, storage media, and hard drives managed by Berenstain and their assistants. The digital files are forensically captured images of those stored on hard drives, 3.5 inch floppy disks, compact disks, zip disks, and other media in the Berenstain papers."

Example #5 Sub-Series Level:
"The material contained in this subseries consists of documents, graphics, spreadsheets, databases, software programs, video, audio, and other material created and/or stored on computers, storage media, and hard drives managed by [creator]."

Organization and Arrangement

DACS: System of Arrangement Element, 3.2

EAD3: Arrangement <arrangement>
ArchivesSpace: Arrangement Note
MARC: 351
ISAD(G): 3.3.4
RDA: System of Organization, 7.8

RECOMMENDED
Include a sentence or two about how the digital materials were organized and arranged.

Note whether or not the original order of the files has been maintained.

Note whether the digital materials have been segregated into their own series or whether the digital material has been subdivided into multiple series, and if so on the basis of what (ie, content, format, etc.)? This is especially important to note in the context of hybrid collections, as well as in cases where there may be duplicative or overlapping material, which is often due to a donor's migration and/or backup routines.

If the files have been arranged by the processor into a folder structure, be sure to include a description of any major necessary rearrangement--especially filename or structure changes, the creation of new "folders", etc. As a point of reference, use the following scenarios, as outlined by the Descriptive Standards Roundtable, of different degrees of intervention and arrangement[17]:

1. *No intervention by the archivist: The original order of the material has been retained without any intervention by the archives staff. This does mean that the collection may include filenames containing spelling mistakes and duplicate files that were placed in the wrong folder(s) by the creator.*

2. *Minimal intervention by the archivist: The original order of the material has been retained without any intervention by the archives staff. With the material being received over time from the depositors some automatic processing has been undertaken to identify and remove files that are exactly the same as files already contained within the archive.*

3. *Intervention by the archivist: The original order of the material as received by the archive service has not been retained as the arrangement of the material reflected the creator's current working practice. The material has been re-organised following consultation with the depositor into an order comprising a number of sections and sub-sections to allow*

---

[17] The majority of this section was taken verbatim from 3.3.4 ("System of Arrangement") in the "Best Guess Guidelines for Cataloguing Born Digital Material." Descriptive Standards Roundtable, 2016. http://www.archives.org.uk/images/Data_Standards/Best_Guess_Guidelines_v1.0_160325.pdf

subsequent digital material to be easily integrated into a single system of arrangement.

4. *Post-ingest intervention: The material was ingested into our digital repository in its original order. As part of the processing work conducted by the archives staff some intellectual reorganisation of the material was undertaken to allow subsequent material to be easily integrated into a single system of arrangement.*

    The arrangement presented here is a result of this intervention by archives staff, but the original file paths are recorded in the descriptions at lower levels. [The file path would then be added to the system of arrangement field at those lower levels].

Please note that there is potential for overlap between this field and the "Processing Information" field. Duplication may not be avoidable, but whereas "Processing Information" provides as complete a record as possible of the actions undertaken on the material, both before and after its arrival at the archive, "Arrangement" relates to recording decisions on arrangement only and to explaining what the arrangement (or perhaps more accurately the level of processing) is so that its impact on the material can be assessed.

Examples:
Example #1
"Born-digital materials are integrated into their corresponding series based on content. The original order of the files has been retained."

Example #2
"Collection is arranged into series: Correspondence, which includes Strickland's emails; Writings, including Strickland's projects, collaborations, essays, publications, and writings by others; Education and Academia; Miscellany; and Audiovisual."

Example #3
"Ávila's digital material has been arranged into two series: Photography, which includes various versions of image files she created for her 1994 exhibition, "Present Past", and Writings, which includes drafts of artists statements, art criticism, and essays. In each series the files have been arranged into a single directory and are organized alphabetically by filename."

Example #4
"The original structure, naming, and contents of the material have been maintained. Sontag migrated files from one computer to the next; accordingly, duplication and variation in files are common. Files and folders with the same or similar titles may not have exactly the same contents; it is also not uncommon for segments of text to be repeated across files of different names, as many of the file instances are drafts and often build upon one another."

Example #5

"Digital materials have been placed into the 'writings' series based on their content. No other changes have been made to the original arrangement, structure, or naming of the files."

Processing Information

DACS: Processing Information Note in Notes Element, 7.1.8
EAD3: Processing Information <processinfo>
ArchivesSpace: Processing Information Note
MARC: 583
ISAD(G): 3.7.1
RDA: N/A

The Processing Information section is one of the most important aspects of any finding aid that describes born-digital materials. Decisions made during processing can greatly affect who, what, where, when, why and how researchers access and understand the digital material within a given collection. In particular, processing legacy born-digital material can often involve changing the nature of the data to make it preservable and accessible: this may include migrating to different file formats, redacting or removing Personally Identifiable Information (PII), extracting files, or normalizing filenames. It is essential that this information be recorded to ensure that future archivists and users understand where the materials came from, how they were created, and the process by which they are able to access the materials.[18]

For the cases in which libraries are not able to make born-digital materials accessible, even after preserving content, this section will also allow users to understand access limitations of old file formats or hardware. It is essential that organizations include information in their finding aids about some of the key preservation actions and techniques employed during processing.

As is clear from the lengthy list of bullet points that follow, this can sometimes mean that the Processing Information section can become top-heavy, and often with information that is not collection-specific. For example, acquisition procedures, PII scans, virus scans, and file normalization methods may be standard processing steps for any collection with born-digital materials. Therefore, the authors recommend using the section to link to a Processing Information document that lives **outside of the finding aid**, potentially with other standard processing procedures. Please see the Systems of Record section of this guide for more information about why this option can be useful.

REQUIRED:

---

[18] Note that information about the processor of the collection should be recorded as described in the "Processor" section of these guidelines, and not in the Processing Information element.

- **Removal or redactions**
Briefly note if any files were redacted or removed during processing as a result of PII screening or because of donor requests. Try to keep this to one sentence or part of one sentence, and avoid going into specifics. (Detailed information about restrictions should be recorded in the Conditions Governing Access and Restrictions section).
    - If any material is removed or redacted either as a result of the PII screening--such as contact information, medical/student records, or bank account numbers--or in accordance with departmental or institutional policies, consider indicating the reasoning or policies in place that guided this decision.
        - Ex. *"Some files have been redacted to protect personal identifiable information or Protected Health Information (PHI) in accordance with federal regulations and UCSF Archives and Special Collections privacy policies."*
    - Briefly explaining what happens to the restricted or redacted content.
        - Ex. "*Content was removed from access copy and quarantined with master copy for preservation.*"
    - Use the Restriction note to document these actions in greater detail.

- **Unprocessed material**
Some collections may contain some digital material that has not been or cannot be processed. In cases like this, it is important to provide researchers with a clear understanding of what item(s) remain unprocessed.
    - Note the number and format of any unprocessed items.
    - If processing was attempted but ultimately not possible, note the reason(s) why this material could not be processed. (E.g. institutional policy, lack of identifying labels or description, an item's UC Value Score, a lack of appropriate technology, a technological error such as an unrecognized file system, physical degradation that prevents an item from being read, or a combination thereof.)
    - In cases where a repository has the ability to process material but has not yet processed it, indicate that the materials are not available in their original format and must be reformatted to a digital access copy.

- **Reformatting or normalization**
Note whether files have been reformatted and/or normalized for preservation or access. In addition, note tool, process or application used to transcode files, as well as the types changed.
    - Ex. "*.TIFF, .JPEG, .PICT, and .PSD files were normalized to the .TIFF format for preservation and the .JPEG format for access using Archivematica v 1.5.1*."
If your organization has a policy governing reformatting or normalization, consider mentioning or linking to it.

- ○ Ex. "***Files were normalized to preservation-standard formats according to the Library's File Format Policy***. ***For more information, see [permalink].***"

- ● **Any other choices made during processing that may affect a researcher's understanding of the digital material in question**
  No two collections are the same. There are a number of factors that may influence processing decisions including, but not limited to, institutional policy, available resources, format of the source media, available contextual information, technological constraints, etc. In certain instances, processors may make choices in order to facilitate access or preservation that, without context or explanation, could impact a researcher's understanding of the files in a given collection.

  Examples of this could include:
  - ○ Modifying or cleaning the filenames in any way.
    - ■ Ex. ***"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.***"
  - ○ Altering (or choosing not to alter) the original file structure.
    - ■ Ex. ***"Files have been stored alphabetically by filename in a single directory."***
    - ■ Ex. ***"The materials have not been arranged beyond file type."***
  - ○ Deleting (or choosing not to delete) empty directories, folders, and/or duplicates.
    - ■ Ex. "***Sontag migrated files from one computer to the next; accordingly, duplication and variation in files are common. Files and folders with the same or similar titles may not have exactly the same contents; it is also not uncommon for segments of text to be repeated across files of different names, as many of the file instances are drafts and often build upon one another.***"
  - ○ Reformatting or normalizing file formats for preservation and/or access.
  - ○ Changing (or choosing not to change) permissions on files or folders.
    - ■ Ex. ***"Content on the "LCROSS" DVD was locked, so a disk image could not be created."***
  - ○ Any instances where processors have chosen to pursue unique processing methods or procedures that diverge from typical processing policies or protocols.
    - ■ Ex. the decision to retain both an original set of proprietary file formats as well as normalized versions of those files in anticipation of future emulation-based access strategies.

RECOMMENDED

- **Mode of acquisition and/or file capture**[19]
  Understanding the process(es), hardware and software used to acquire media and/or capture files is important. Not only does this information help contextualize an acquisition, it can also help demonstrate chain of custody, authenticity, and provenance. It is especially important to document this information for files that were originally received on storage media (e.g. optical disks, floppy disks, flash drives, etc.), since the characteristics of these storage media can impact readability and integrity of the files in distinct ways. Include any information that may help a researcher better understand how the files arrived at an archive and what they were like prior to processing. This could include:
  - Whether the files were received as a single group or in multiple installments.
  - The number and format(s) of storage media on which the files were received.
  - The technique, hardware and software used to transfer files from the storage media on which the files were received.
    - Indicate whether disk images were created or whether files were transferred using another tool or strategy such as applications that employ FTP, an email client, or via a third party cloud storage service like Google Drive or Box.
    - Consider noting to where the files were transferred and/or where they're being stored (e.g. a digital repository).

- **Virus scan**
  Indicate if a virus scan was performed. Performing a virus scan during processing is a standard practice, so if *no* virus scan was performed be sure to note this. You may also include any of the following:
  - Whether any viruses or malware were found.
  - Whether any suspicious files were deleted or removed.

- **Personally Identifiable Information (PII)**
  Note if the files have been screened for PII and indicate the software used to do this.

- **File weeding and/or deletion**
  It is important to note whether or not any files were weeded or deleted during processing. This information is especially useful in cases where the size of processed files differs from the file size of the original acquisition.

---

[19] General information about the acquisition of born-digital material, such as the source, date, and type of acquisition, may be appropriately recorded in an Immediate Source of Acquisitions Note (DACS: Immediate Source of Acquisition Element, 5.2; EAD 2002: Acquisition Information Element <acqinfo>; ArchivesSpace: Immediate Source of Acquisition Note; MARC: 541). However, since acquisition of born-digital content often involves technical processing and data capture after media has been physically transferred, the specific methods and processes of born-digital acquisition and data capture should be detailed in Processing Information. Do not enter specific technical details in Immediate Source of Acquisition.

- ○ Include the reason(s) for weeding or deleting files: Were they duplicates? Did they lack research value? Was the library unable to provide access due to lack of required hardware, software, or system specifications? Were they inessential, system files, or proprietary program files? Were they deleted due to copyright concerns?
  - ○ If duplicate files exist that have NOT been weeded or deleted, please note this information as well.
    - ■ Ex. "***The original file structure has been maintained; duplicate files have not been deleted.***"

- **Checksums[20]**
  Consider noting the files and/or disk images for which checksums were created. Generating checksums during processing is a standard practice, so if *no* checksums were created be sure to note this.
  - ○ If they were created, include what checksum format was used, (e.g. MD5, SHA-256, etc.) and note what tool was used to create the checksums, the level at which checksums were created, and when the checksumming took place.
    - ■ Ex. *"MD5 checksums were assigned to the disk images at the point of creation using FTK Imager, after which files were extracted from the disk images and assigned SHA-256 checksums using Archivematica version 1.5.1.*

- **Filename cleaning and/or normalization**
  Often during processing filenames with special characters or undesirable formatting must be normalized to prevent technical glitches from occurring. If any of the original filenames were cleaned, changed, or amended during processing, it is important to provide a record of this. This will often occur when processing files from older media because modern file systems don't allow certain characters (/,&,etc.) in filenames.
  - ○ If filenames were changed at all during processing, mention why the names were changed. (E.g. institutional practice, to facilitate access and discovery, forced or automated change by file system, etc.)
  - ○ Assuming a record of the original filenames has been maintained, be sure to note that researchers may see the original filenames upon request.

- **Hardware and software used**
  Providing information about the hardware and software used to perform various aspects of processing digital media may be helpful to researchers. This may be particularly useful in the context of larger, more complex donations, as these can include a wide array of digital media that required several different tools/workflows to process.

---

[20] A checksum is a unique alphanumeric number that is used to help detect errors and ensure that a digital file remains unchanged over time. For more information, please see: "Fixity and checksums." Digital Preservation Handbook, 2nd Edition. Digital Preservation Coalition. (2015). http://www.dpconline.org/handbook/technical-solutions-and-tools/fixity-and-checksums

- ○ Note tool name and version, as well as its purpose. (E.g. disk image creation, metadata reports, import/export, filename changes, etc.)
- ○ Consider including the reason(s) a particular tool was chosen over another, if relevant. This could exist as boilerplate text that explains each of the common tools used institutionally for processing, with more specific information added in the core processing note.
    - ■ Ex. "***FTK Imager was used to create exact, bit-level disk images of born digital media, and to generate checksums both before and after image capture to ensure fixity***."

- ● **Institutional processing policies or guidelines**
  Include a link to your institution's processing policies and guidelines, provided that they exist in a stable location and are kept up-to-date. This could exist as boilerplate text that is appended to the end of every Processing Information section.
    - ○ Ex. ***"This collection was processed in accordance with UC XX's processing guidelines/policies. For more information, see [permalink].***"

Examples:
Example #1
"Fisher's born digital files were received on 25 3.5 inch floppy disks, 19 5.15 inch floppy disks, a Windows-formatted 500 GB hard drive, and one .pst email file. The .pst file was transferred to UCSD using the secure RDL Share tool. Disk images from the floppy disks and hard drive were created using a Windows instance of FTK Imager, and were scanned for PII using Identity Finder. Files with personal information (social security numbers, bank account numbers) were removed from the access copies of the disk images and quarantined on a secure non-networked server for storage. No other content was restricted or redacted from the born digital files. A virus scan using Clam AV was performed on both the master disk images and the content processed for access. 4 3.5 inch floppy disks were removed from the original donation due to containing proprietary software not created by Fisher. Mp3 files found on the hard drive were reformatted to .WAV for preservation, using Archivematica's normalization tool. The .pst file is currently awaiting processing for the completion of ePADD testing at UCSD."

Example #2
"The AIDS Treatment News records were received from John Jacobs on two external hard drives. The files were extensively screened by the donor prior to transfer to the archives, mainly for mailing list and other contact information. Archives staff screened the material for viruses using Malware Bites, and for personally identifiable information (PII), protected health information (PHI), and other protected information using Forensics Toolkit (FTK) version 5.5. Several files containing pieces of the publication mailing list have been removed. Email was ingested and processed using ePADD version 3.0, and named entities were noted and added as subject headings.

Some additional files were ingested on 20 compact disks, but were unable to be processed along with the rest of the collection due to an unidentifiable file system. The discs have been retained for further investigation."

Example #3
"Digital files from the Bruce Conner papers were acquired on seven compact disks, which were disk imaged to the .ISO format on 2013-11-21. The compact disks were scanned for viruses before imaging, with no results. Files were originally created on an Apple (Mac) computer using the HFS+ file system. The original file structure has been maintained. All files were screened for personally identifiable information (PII), and no files were restricted.

Some filenames were modified during processing to enable files to be exported from the disk images. A list of modified filenames and their original filenames is available. Files were normalized to preservation-standard formats according to the Library's File Format Policy, using Adobe Photoshop, Adobe Bridge, and Audacity. .TIFF, .JPEG, .PICT, and .PSD files were normalized to the .TIFF format for preservation and the .JPEG format for access. Word documents were saved in the .PDF format for preservation and access, and .MP3 recordings were saved to .WAV format for preservation and a second .MP3 copy created for access."

Example #4
"UCLA received Sarah Smith's files in two simultaneous installments. The first installment was a set of files that included drafts of Smith's most recent writing projects, totalling 235 GB, which were received directly from Smith via FTP. After being scanned for viruses and malware in Forensic Toolkit (FTK) version 5.5, the cleaned files were arranged into two series: Screenplays (series #1), and Short Stories (series #2). Files within each of these series have been organized according to project title, but have not been further arranged. Duplicate files detected using hash values were deleted. Special characters were removed from file names to facilitate researcher access (for instance, files associated with Smith's 2012 Western, "Never Apologize!", have been changed to "Never Apologize"). The majority of the files in these series consisted of drafts written in Microsoft Word and the screenplay software Final Draft. These files were converted to PDFs using Archivematica version 1.5.1.

The second installment consisted of 3 external hard drives (each 500 GB), 27 DVDs (each 4.7 GB), and 11 Zip disks (each 250 GB) that contained files related to Smith's charitable work including the non-profit she co-founded, Kids Write Now. Disk images were created of the DVDs using Guymager version 0.8.4; of these disk images, 2 disk images are unable to mount, which is likely due to file corruption inherent to the original media. In addition, 3 of the DVD-ROMs included files that were formatted as "read only", which prevented us from creating disk images. Disk images of the external hard drives

were created using Forensic Toolkit (FTK) version 5.5. Files were extracted from the disk images and were scanned for viruses using either Clam AntiVirus (ClamAV) or Forensic Toolkit (FTK) version 5.5. The files received in this installment constitute their own series, Kids Write Now ("series #3). The original order of these files has been maintained, and the files have not been further arranged.

In all cases, files were ultimately processed and packaged for long-term preservation using Archivematica version 1.5.1. As per UCLA LSC's processing guidelines, all files were scanned for Personally Identifiable Information (PII) and some information or files have been redacted where appropriate. Files were normalized for preservation and access according to departmental policy. Please note: At the time of processing, UCLA LSC does not have a Zip drive and was therefore unable to image Zip disks. If interested in viewing this material, please contact the reference desk for more information."

Container List & Inventory

Including a container list or inventory in a finding aid can provide an additional layer of insight into the digital materials within a given collection. A file list that includes filepaths may be especially useful in this regard, as allowing researchers to view filenames and directory structures remotely may help facilitate identification and location of the material in which they are interested. This has the potential to save both time and resources.

This information can be generated in a number of ways depending on the format, age, condition, and variety of the digital material.

RECOMMENDED:

MANUALLY
Even a basic list or spreadsheet of the media formats, dates (if known), and contents of the media in a collection can provide a useful point of reference. If the donor provided extensive description of or information about the digital portion of a collection, consider making it available as a container list.

AUTOMATICALLY
File lists or inventories can be easily generated automatically during processing and/or appraisal using tools such as Directory List and Print Pro.[21] Creating a file list does not take very much work to do, and can yield a surprising amount of information. However, these inventories can often be lengthy, so it is better to link out to a separate document that contains this information rather than including it directly in the finding aid.

---

[21] See https://directory-list-print-pro.en.softonic.com/

- Ex. *"For a complete inventory of files, please see: [link]"*.

However, just because inventories may be easy to generate automatically does not mean that they should be relied upon for description. It is generally worthwhile to attach a file inventory to a finding aid or collection record, but *it should not serve as the only description*, especially without review and possibly editing for clarification. Filenames may be inaccurate, misleading, or too vague to be helpful. In addition, the inclusion of a large and complex file list can have implications for public service staff, who may have to spend more time explaining these lists to researchers than they would with a standard finding aid.

It is recommended that file lists always be generated as part of processing born-digital records, but not that it be solely relied upon for description and access.

APPENDICES

Metadata Fields Crosswalk

UC Guidelines Term	DACS Second Edition (2013-15)	EAD 3	ArchivesSpace	MARC	RDA	ISAD(G)
Processor	Notes Element (7.1.8, Processing Information)	<processinfo> ; <control> (replaces <eadheader> from EAD 2002)	Processing Information Note ; Collection Management (inside Processors element) ; Finding Aid Author (if author of finding aid is same as processor)	N/A	N/A	3.7.1
Physical Description and Extent	Extent Element (2.5)	<physdesc> ; <physdescstructured> (replaces <extent> from EAD 2002)	Physical Description Note ; Extents	300	Extent, 3.4 ; File Size, 3.19.4	3.1.5
Abstract	Scope and Content Element (3.1)		Abstract Note	520	Summarization of Content, 7.10	N/A

Conditions Governing Access and Restrictions	Conditions Governing Access Element (4.1)	<accessrestrict>	Conditions Governing Access Note	355 ; 506	Restriction on Access, 4.4	3.4.1
Physical Characteristics and Technical Requirements	Physical Access Element (4.2) ; Technical Access Element (4.3)	<phystech>	Physical Characteristics and Technical Requirements Note	336-338; 340 ; 538	Media Type, 3.2 ; Carrier Type, 3.3 ; Digital File Characteristic, 3.19 ; Equipment or System Requirement, 3.20	3.4.4
Conditions Governing Reproduction and Use	Conditions Governing Reproduction and Use Element (4.4)	<userestrict>	Conditions Governing Use Note	540	Restriction on Use, 4.5	3.4.2
Immediate Source of Acquisition	Immediate Source of Acquisition Element (5.2)	<acqinfo>	Immediate Source of Acquisition Note	541	Immediate Source of Acquisition of Item, 2.19	3.2.4
Appraisal Information	Appraisal, Destruction, and Scheduling Information Element (5.3)	<appraisal>	Appraisal Note	583	N/A	3.3.2
Processing Information	Notes Element (7.1.8, Processing Information)	<processinfo>	Processing Information Note	583	N/A	3.7.1
Scope and Content	Scope and Content Element (3.1)	<scopecontent>	Scope and Content Note	520	Summarization of Content, 7.10	3.3.1
Organization and Arrangement	System of Arrangement Element (3.2)	<arrangement>	Arrangement Note	351	System of Organization, 7.8	3.3.4

Controlled Vocabulary

Background

The authors of the UC Guidelines for Born-Digital Description were aware of the need for a robust controlled vocabulary for born-digital materials throughout the writing of the guidelines. Born-digital terminology can vary greatly in spelling, term, or phrase, even amongst existing library thesauri. Efforts were already underway at UCLA to standardize term usage within their finding aids. In order to accelerate and leverage this work, the UCLA archivists were deputized to form a Lightning Team and tasked with developing a controlled vocabulary for born-digital source media and related terms.

It should be noted that the team chose to draw a line in the sand regarding analog audiovisual (AV) carriers. It was agreed that these were out of scope for this particular document, and so only digital AV carriers were included, as appropriate. However, a controlled vocabulary for AV terms independent of these guidelines is currently in the works. The born-digital controlled vocabulary developed by the lightning team was designed for the explicit purpose of describing born-digital materials in finding aids. It is acknowledged that catalog records will continue to adhere to other, existing standards.

Included in the controlled vocabulary list are "related terms" which consist of carriers and other digital terms that often appear in disparate ways. While more style-based, for example "Internet" vs. "internet," the standardization of these related terms will further aid in consistent archival description. We also recommend adhering to the Chicago Manual of Style, 17th Edition for general style guidance.

Methodology

The lightning team reached out informally to colleagues at institutions throughout the UC system, California, and the US and inquired as to which standards institutions were using for describing born-digital materials, and whether they had an internal controlled vocabulary. The responses of that informal inquiry revealed what many within the UC system have experienced -- a lack of agreement on terms. The Getty's Art and Architecture Thesaurus (AAT) was the most commonly used resource, but no one felt that any single authority was comprehensive. Furthermore, choosing a broad thesaurus such as AAT also leaves the door open to have a variety of terms used (albeit, still standardized). The desired controlled vocabulary was a narrower, specialized list, which would allow processors and users to easily search collections for those preferred terms and have a shared understanding of what those terms mean.

To begin with, the team created a standards crosswalk which included common AV and born-digital terms, pulling from Public Broadcasting Metadata Dictionary Project's PBCore; AAT; Resource Description and Access' (RDA) content, media type, and carrier; and what were dubbed "terms in the wild." The crosswalk was referred to repeatedly throughout the process, which helped in two ways. It allowed for easy identification of gaps in the terms of the those authorities, such as inadequate or missing terminology, and illustrated when there was a lack of consensus, for example DAT vs. digital audio tapes. The crosswalk also served as a reminder when appropriate terms had been identified by an authority, such as "flash drive." The team created local terms where gaps were identified, however, existing authorized terms were chosen whenever possible.

Consideration was also given to colloquial language and usability needs of end users. As these are guidelines for description, it is important that familiar and recognizable language and terms are employed. A conscious choice was made to use proprietary names (Zip, Jaz, MiniDisc, etc.) as this information can be helpful to researchers in dating the contents of these carriers, especially because many of these formats were often short-lived. Many terms were not included in PBCore, as it focuses primarily on AV materials, and the terminology included in both AAT and PBCore often conflicted. In many cases the AAT term was chosen since it is the more commonly used resource, and also more comprehensive, but in some cases PBCore's detailed language was found to be the better option.

Software and system formatting concerns played a role, as well. For example, when deciding the language of floppy disks, i.e., "3.5" floppy disk," "3 ½ floppy disk," and "3.5 inch floppy disk," the latter was chosen in part because it does not employ fractions, which display differently depending on software, or quotes, which can be auto-changed to smart quotations depending on software. Furthermore, both RDA and Describing Archives: A Content Standard (DACS) encourage users away from abbreviations, resulting in the choice to spell out the word "inch." The final choice, "3.5 inch floppy disk," will both avoid common systems errors and be easily understood by end users.

In line with other UC guidelines, this list was created with the flexibility to be applied to efficient processing work. For example, when describing content at a series level, description could simply be limited to "memory cards." Alternatively, if more granular description is warranted, the processor has an option to include information about specific types of memory cards in parentheses after the term. For example, "Collection contains 2 memory cards (2 mircoSD, 1 CompactFlash II)."

This list is merely a starting point and will be updated on an ongoing basis, especially as new technologies, standards and best practices arise.

See the Controlled Vocabulary for Born-Digital Source Media below, for a comprehensive list of both born-digital carriers and related terms. The standards crosswalk, as well as the lists, can be [viewed online here](#).

Controlled Vocabulary for Born-Digital Source Media

Term	Use for	Source	Notes
3.5 inch floppy disk(s)	3.5", 3.5', three and a half inch, 3 1/2	local	
5.25 inch floppy disk(s)	5.25", 5.25', five and one quarter, 5 1/4	local	
8 inch floppy disk(s)	8", eight inch	local	
Betacam SX		local	
cellular telephone(s)	cell phone	AAT	
compact disc(s)	cd, CD, cd-r, CD-R, cd-rom, CD-ROM, cdrom, CDROM, data disk, cd-rw, CD-RW, cdr, CDR, cdrw, CDRW	AAT	Use for any type of compact disc. Optionally, include information about the specific type of compact disc in parentheses after the term, such as: CD-ROM, data disk, CD-R, CD-RW, CD-W. Example: Collection contains 15 compact discs (4 CD-ROMs, 9 CD-Rs).
computer(s)	computer tower, pc, PC, Mac, Macintosh, Apple computer	AAT	
D1		PBCore	
D2		PBCore	
D3		PBCore	
D4		PBCore	
D5		PBCore	
D6		PBCore	
D9		PBCore	
DCT		PBCore	
digital audio tape(s)	DAT (PBCore), DAT 72, DAT 160, DAT 320, R-DAT	AAT	
Digital Betacam(s)	DigiBeta, digi beta, digi-beta	PBCore	
digital compact cassette(s)	DCC	OED	

digital data storage	DDS, DDS-1, DDS-2, DDS-3, DDS-4	local	
Digital8		PBCore	
DV		PBCore	
DVCAM		PBCore	
DVCPRO		PBCore	
DVD(s)	dvd, Blu-ray, digital video disk, digital versatile disk, EVD, Super Video CD, DVD-ROM	AAT	
flash drive(s)	thumb drive, jump drive, USB stick, USB drive, memory stick	AAT	
floppy disk(s)	computer diskette, Floppy disk, floppy-disk, floppy disc	AAT	
hard disk(s)	hard disc, internal hard drive, internal drive	AAT	
hard drive(s)	hard-drive, harddrive, external drive, external hard drive, external hard disk drive	AAT	
HDCAM(s)		PBCore	
HDV		PBCore	
Jaz drive(s)	Jaz disk, Iomega storage disk, iomega	local	
laptop computer(s)	laptop	AAT	
LaserDisc(s)	laser disk, Laserdisk	local	
magnetic tape data storage	data tape, magnetic recording, Linear Tape-Open	local	Optionally include information about the specific type of magnetic tape data storage in parentheses after the term, such as: LTO (3 LTO-6).
memory card(s)	compact flash cards, CompactFlash, flash card, memory stick, microcard	OED	Use for any type of memory card(s) found in digital cameras, mobile phones, portable media players, etc. Optionally, include information about the specific type of memory card in parentheses after the term, such as: microSD. Example: Collection contains 3 memory cards (2 microSD, 1 CompactFlash II).

	minidisc, mini disc, Mini disk, Sony Mini Disc, MD	local	
MiniDisc(s)			
MiniDV(s)	Mini-DV (AAT), mini dv	PBCore	
MiniDVD(s)	mini dvd, mini-dvd	local	
personal digital assistant(s)	PDA, handheld PC, Palm Pilot, BlackBerry, Sidekick, personal handheld device	AAT	Use for any type of mobile device functioning as a personal digital information manager that predated smartphones. Optionally, include information about the specific type of Palm Pilot, BlackBerry, Sidekick, etc. Example: Collection includes 1 personal digital assistant (1 BlackBerry).
punched card(s)	punch cards	AAT	
smartphone(s)	iPhone, android	AAT	Use for any type of mobile device functioning as a personal digital information manager capable of operating computer applications. Smartphones typically feature cameras, touch screen interface, and ability to access the internet. Optionally, include information about the specific type of Palm Pilot, BlackBerry, Sidekick, etc. Example: Collection includes 1 personal digital assistant (1 BlackBerry).
SyQuest disk(s)	Syquest, syquest, syquest drive	local	
tablet computer(s)	iPad, Kindle	AAT	
Zip disk(s)	Iomega storage disk, iomega	local	

Controlled Vocabulary for Born-Digital Related Terms

Term	Use for	Source	Notes
artificial intelligence	AI	AAT	
audiovisual	audio-visual, a/v, A/V, av	AMIA	Spell out in full on first usage; use 'AV' thereafter.
augmented reality	AR	AAT	

born-digital	born digital (AAT)	local	
checksum	check sum	OED	
computer game(s)	PC games	AAT	
dataset	data set	OED	
disk image	Disk Image, image, disk-image	local	
electronic mailing list(s)	listserv, LISTSERV, mailing list	AAT	
email	electronic mail (AAT), E-mail, Email	OED	
emulation	emulating, emulated	AAT	
file directory		local	
filename		local	
filepath		local	
GB	Gigabyte, gig, gb, gigabytes (AAT)	SAA-ACRL/RBMS Joint Task Force on the Development of Standardized Holdings Counts and Measures for Archival Repositories and Special Collections Libraries	See 'Physical Description and Extent' section for more detail.
geographic information system	GIS	AAT	
geospatial data	GPS, geo-spatial data	AAT	
hardware	Hardware, hard-ware	AAT	
hash value	hash (OED)	local	
instant message(s)	IMs, chat, iMessage	local	
internet	Web, World Wide Web (AAT), web, Internet	OED	
intranet(s)	Intranet	AAT	
metadata	meta-data, Metadata, meta data	AAT	
open source	opensource, open-source	AAT	
personally identifiable information	PII, Personally Identifiable Information, sensitive personal information	local	Spell out in full on first usage; use 'PII' thereafter.

		US Department of Health and Human Services, summary of the HIPAA Privacy Rule	Spell out in full on first usage; use 'PHI' thereafter.
protected health information	PHI, personal health information		
server(s)		AAT	
social media	participatory media	AAT	
software	Software, soft-ware	AAT	
source media	storage media, removable media	local	
text message(s)	SMS, short message service, MMS, multimedia messaging service, text-message, iMessage, texts	AAT	
video game(s)		AAT	
virtual reality	VR	AAT	
web page(s)	webpage, page	OED	
web site(s)	Website, Web site (AAT), web-site, web page, online resource	OED	
webinar		AAT	
wiki(s)	Wiki, wikipedia	AAT	

Attaching Files to Finding Aids on the Online Archive of California

File lists, directory lists, and other supplementary files can be linked from the finding aid in such a way that they will display on the Online Archive of California (OAC). This can be a helpful way for institutions who use California Digital Library's (CDL) services to share these documents with researchers in cases where they have been generated as part of the description process and the institution wishes to provide access to them. File lists are attached using the EAD <extref>[22] element, which links to a file hosted on another server. A document linked using the <extref> element can be hosted simply on another server, or by the OAC/CDL itself. In the case of documents hosted by OAC, pdf files attached using the instructions (linked below) will also be indexed for search.[23]

---

[22] <extref> is an obsolete element in EAD3 and is no longer part of the schema. The authors are awaiting guidance from OAC about which EAD3 element will be used as its replacement for this functionality going forward. At the time of this writing EAD2002 was still the standard in use by most of the contributing institutions.

[23] Note that ONLY pdf files can be attached using the OAC-hosted method -- the architecture does not currently allow for any other file types.

OAC's guide to submitting supplemental pdf finding aids, directory, or files lists can be found here:
https://help.oac.cdlib.org/support/solutions/articles/9000093762-submitting-supplemental-pdf-finding-aids

Sample Finding Aid

Note: portions of this finding aid have been abbreviated as indicated by [...]
Integrated born-digital description appears in red.

Finding Aid to the Bruce Conner papers, 1940 - 2010, bulk 1960 - 2010

Collection Number: BANC MSS 2000/50 c

The Bancroft Library
University of California, Berkeley
Berkeley, California

Finding aid written by: Dean Smith, Nissa Nack, Kate Tasker
Date completed: 2017

Collection Summary

Collection Title: Bruce Conner papers
Date (inclusive): 1940 – 2010
Date (bulk): 1960 – 2010
Collection Number: BANC MSS 2000/50 c
Creators: Conner, Bruce
Extent: 30 linear feet (20 Cartons, 1 Box, 2 Oversize Boxes and 7 Oversize Folders) and 1.4 GB (595 files).
Repository: The Bancroft Library
University of California, Berkeley
Berkeley, California, 94720-6000
Phone: (510) 642-6481
Fax: (510) 642-7589
Email: bancref@library.berkeley.edu
URL: http://bancroft.berkeley.edu/

Abstract: The Bruce Conner papers, 1940s - 2010, form a voluminous, comprehensive overview of Conner's life as a visual artist and filmmaker spanning the length of his celebrated

career. The collection includes correspondence with individuals, galleries and museums, announcements, programs, articles, reviews, interviews, lectures, awards and grants, contracts, invoices, legal files, chronological files, digital images, text files, and audio recordings, and other sundry documentation all in great detail.

Languages Represented: Collection materials are in English and Spanish.

Physical Location: Many of the Bancroft Library collections are stored offsite and advance notice may be required for use. For current information on the location of these materials, please consult the Library's online catalog.

Access
Collection is open for research.

Physical Characteristics and Technical Requirements
Born-digital content is found in Series 10, Personal. These files must be accessed using the Library Digital Collections laptop in the Reading Room. Access copies are provided as PDF, JPG, and MP3 files. Advance notice is required for use. Online access is not available.

Publication Rights
Materials in this collection may be protected by the U.S. Copyright Law (Title 17, U.S.C.). In addition, the reproduction of some materials may be restricted by terms of University of California gift or purchase agreements, donor restrictions, privacy and publicity rights, licensing and trademarks. Transmission or reproduction of materials protected by copyright beyond that allowed by fair use requires the written permission of the copyright owners. Works not in the public domain cannot be commercially exploited without permission of the copyright owner. Responsibility for any use rests exclusively with the user. All requests to reproduce, publish, quote from, or otherwise use collection materials must be submitted in writing to the Head of Public Services, The Bancroft Library, University of California, Berkeley 94720-6000. See: http://bancroft.berkeley.edu/reference/permissions.html

Preferred Citation
[Identification of item], Bruce Conner Papers, BANC MSS 2000/50 c, The Bancroft Library, University of California, Berkeley.

Alternate Forms Available
The scrapbooks in this collection were partially digitized by the creator. Access copies of digital files are available as PDF, JPG, and MP3 files. See Series 10, Personal.

Related Collections
- Photographs from Bruce Conner correspondence [graphic], BANC PIC 1997.069

- Bruce Conner correspondence concerning Jay DeFeos "The Rose", circa 1930-1996, BANC MSS 98/32 c
- Steven Fama collection on Bruce Conner, circa 1960-2003, BANC MSS 2008/236
- Paula Kirkeby/Smith Andersen Gallery collection of Bruce Conner, 1970-2001, BANC MSS 2010/101
- Gallery Paule Anglim records, [ca. 1976-2001], BANC MSS 2005/162 c
- Richard Brautigan papers, 1942-2003, BANC MSS 87/173 c
- Michael McClure papers: additions, 1874-2003 (bulk 1949-2002), BANC MSS 2003/222 c
- Auerhahn Press records, 1959-1967, BANC MSS 71/85 c
- Serious Business Company records, BANC MSS 84/93 c
- Larry Keenan, Jr. photograph archive, BANC PIC 2009.050

Separated Material

Printed materials have been transferred to the book collection of The Bancroft Library. Photographs have been transferred to the Pictorial Collections of The Bancroft Library. Videotapes/sound recordings have been transferred to the Microforms Collection of The Bancroft Library. Digital media have been transferred to the Digital Collections Unit of The Bancroft Library.

Indexing Terms

The following terms have been used to index the description of this collection in the library's online public access catalog.

Conner, Bruce--Archives
Artists--20th century
Artists--California--San Francisco
Experimental films--California--San Francisco
Assemblage (Art)--California--San Francisco
Conceptual art--California--San Francisco
Drawing--20th century
Drawing--California--San Francisco
Born digital[24]

Acquisition Information

The Bruce Conner papers were given to The Bancroft Library by Bruce Conner and, subsequently, The Conner Family Trust beginning in 1996 with additions continuing through 2010. The digital files were received on seven compact discs in 2008. Disk images of the

---

[24] Authorized term from the Art & Architecture Thesaurus (AAT). Adding the term "born digital" to a list of subject headings for a born-digital or hybrid collection can serve as a useful access point for researchers who are specifically searching for born-digital materials by format. Institutions may want to consider adding the term to records for all born-digital or hybrid collections in their holdings, to enable discovery of all cataloged born-digital materials via a subject search.

compact disks were created on November 21, 2013 to generate master preservation copies and to extract the content files.

Appraisal Information

Temporary digital files (.tmp file format) and system files were removed from this collection, according to the Library's digital preservation and privacy policies.

Accruals

Future additions are expected.

System of Arrangement

Papers are arranged to the folder level. See the specific arrangement notes for each series. Digital files are organized in their original file structure. Please see the attached file directory list for a complete inventory of all available files.

Processing Information

Papers processed by Dean Smith in 2009-2010. Digital materials processed by Nissa Nack in 2015 under the supervision of Kate Tasker.

The digital files in this collection were received on seven compact discs. The majority of the files were created by Bruce Conner as digital images of pages in his personal scrapbooks. Text files in .doc format and audio recordings in .mp3 format are also present. The digital files were originally saved to the compact discs using an Apple (Mac) computer with an HFS+ file system, type and age unknown.

Disk images of the seven compact discs were created by Library staff on November 21, 2013 using Forensic Toolkit (FTK) Imager, with the resulting seven .iso disk image files saved to preservation storage. MD5 checksums were assigned to the disk images at the point of creation. The compact discs were scanned for viruses before imaging, with no results.

The disk images were processed in July 2015. 936 files were extracted and analyzed in FTK version 5.0. Temporary files and system files were identified and removed, leaving 595 content files in .doc, .psd, .pct, .jpg, .tif, and .mp3 formats. All files were screened for personally identifiable information (PII), and no files were restricted. The processed files were exported from FTK in their original formats and in their original file structure and were saved to the Library's preservation server.

Copies of the original image files (.psd, .pct, .jpg, and .tif formats) were normalized to TIFF files for preservation and to JPG files for research access, using Adobe Bridge CS6 version 5.0.2.4 x64. Copies of the original text files (.doc format) were normalized to PDF files for preservation and for research access using Adobe Acrobat XI Pro version 11.0.17. Copies of the original audio files (.mp3 format) were normalized to WAV files for preservation and to MP3 files for research access using Audacity version 2.1.1.

Some filenames were truncated during processing due to different filename character limits in Mac and Windows file systems. A full list of the truncated and original filenames is available upon request.

Biographical Information
"Bruce Conner in twenty-five words or less: Bruce Conner is the best cereal in America. Tasty and nutritious. He never gets soggy. He's always crisp."
- Richard Brautigan (1975)

Sculptor, filmmaker, collagist, painter, draftsman, photographer, conceptual prankster, Bruce Conner defined the very qualities of artistic freedom and in doing so boldly defied categorization and mainstream co-option in pursuit of his visionary images and ideas.

Born in McPherson, Kansas, in 1933, Bruce Conner spent his childhood and young adulthood in nearby Wichita. Upon graduating from Wichita High School East, Conner went on to study art at Wichita University and University of Nebraska, where he met his wife-to-be, Jean Sandstedt. He continued art studies at the Brooklyn Art School and the University of Colorado. In 1957, at the urging of his childhood friend, the poet Michael McClure, and attracted by stories of a vibrant art and literary scene that included visual artists Jay DeFeo, Joan Brown, and Jess, and poets Lawrence Ferlinghetti, Philip Whalen, and Philip Lamantia, he and his wife, Jean, moved to San Francisco. Conner subsequently became a key figure in the City's legendary Beat community. After brief sojourns to Mexico City, 1961-1962, where his son, Robert was born, and Brookline, Massachusetts, 1963-1964, Conner resettled in San Francisco where he continued to live and work until his death.

[...]

Scope and Content
The Bruce Conner papers, 1940s - 2010, form a voluminous, comprehensive overview of Conner's life as a visual artist and filmmaker spanning the length of his celebrated career. The collection includes correspondence with individuals, galleries and museums, announcements, programs, articles, reviews, interviews, lectures, awards and grants, contracts, invoices, legal files, chronological files, digital images, text files, and audio recordings, and other sundry documentation all in great detail.

The digital files were created by Bruce Conner between 2000-2006 to document and preserve his life and art works. Many of the digital files were created as Conner digitized physical materials kept in 4 scrapbooks, which include news clippings, art show announcements, reviews and other memorabilia which document his life and career from grade school through the 1970s. The scrapbooks are included in the collection. Born-digital files, including MP3 audio recordings,

digital images, and text documents are also present. All of the digital files were created on a Mac computer using the HFS+ file system and burned to seven compact discs.

Original file formats include .doc, .psd, .jpg, .tif, .mp3, and .pct. There are 595 permanent digital files in the collection (1.4 GB). The digital files are dated from 2000-12-16 to 2006-03-23, according to their "Last Modified" dates.

Container List

Series 1 Correspondence 1961 - 2009; Undated

Physical Description: Cartons 1 - 3, Oversize Box 1, Folder 1

Arrangement: Hierarchical then alphabetical then chronological.

Scope and Content Note: Outgoing filed first, followed by Family and General.

> **1.1 Outgoing 1961 - 2007; Undated**
>
> **Physical Description:** Carton 1
>
> **Arrangement:** Chronological.
>
> **Scope and Content Note:** Correspondence to Michael McClure is followed by General outgoing which is arranged chronologically.
>
> Carton 1, Folder 1
> **McClure, Michael 1961 Sep. 19 -2000 June 19; Undated**
>
> Carton 1, Folder 2
> **General 1973 Feb. 4 - 2007 July 9**

[...]

Series 10 Personal [1943] - 2005; Undated

Physical Description: Carton 19, Oversize Volumes, 1 – 4; Digital Folders 1 - 7 (1.4 GB or 595 files).

Arrangement: Chronological. Scrapbooks and Miscellany organized at end of series. Digital files are arranged in seven digital folders, which correspond to the seven compact discs containing the original files. Files have been maintained in their original order.

Scope and Content Note: Various materials that are mainly related to events and/or activities outside Conner's art/filmmaking career with the exception of the Scrapbooks which meticulously

document his art/filmmaking career from grade school to the mid-1970s. The Scrapbooks are included in this series as they represent a private accounting of his life in the arts.

Conner digitized many of the items in the scrapbooks, and the resulting digital image files are included in this series. Digital content also include recordings of the artist jamming on harmonica, a recording of a radio interview, and written commentary by the artist.

Physical Characteristics and Technical Requirements: Series 10) Personal contains digital files. These files must be accessed using the Library Digital Collections laptop in the Reading Room. Access copies are provided in .pdf, .jpg, and .mp3 formats. Please see the attached file directory list for a complete inventory of all available files. Advance notice is required for use. Online access is not available.

 Carton 19, Folder 15 - 16

Wichita High School East Yearbooks 1950 - 1951

Carton 19, Folder 17
University of Wichita Yearbook 1952

Carton 19, Folder 18
Smithsonian Institution - Withholding Tax Statement 1954

Carton 19, Folder 19
Passports 1966 - 1993

Carton 19, Folder 20
Business Cards [1960s]

Carton 19, Folder 21
Household Expenses Ledger 1975 - 1980

Carton 19, Folder 22
Museum Memberships 1983; 1988; Undated

Carton 19, Folder 23
Christmas Chair 1985 Dec. - 1986

Carton 19, Folder 24
Graphic Arts Guild Membership 1997 Mar.

Carton 19, Folder 25
Commission/Composer Agreement - [Bruces Traveling Machine] 1999 June - Sep.

Carton 19, Folder 26
The Glen Canyon Park Neighborhood Carnivore Coalition 2002 Oct.

Carton 19, Folder 27
Bruce Conner Dinner Celebrating 50 Years in Show Business 2004 Dec.

Carton 19, Folder 28
Bruce Conner Recording Session - Performing on Harmonica and Vocal Acrobatics 2004 Dec.

Carton 19, Folder 29
Photographs of Bruce Conners Hands by J. John Priola 2005 Aug.

Carton 19, Folder 30
Address Book - Los Angeles Undated

Carton 19, Folder 31
Bottle Decorated with Toys, Walnuts, Shells, etc. Undated

Carton 19, Folder 32
Miscellany 1948 - 2005; Undated.

Oversize Volume 1
Scrapbook [1943] - 1963

Oversize Volume 2
Scrapbook 1963 - 1968

Oversize Volume 3
Scrapbook 1971 - 1974

Oversize Volume 4
Scrapbook 1975 – 1978

Digital Folder 1
Scrapbook I, digitized 2004

Physical Description: 0.201 GB (66 files)

Scope and Content Note: Original disc labeled "Bruce Conner Scrapbook I, 1944 - 1963 (1-61D)." Digital images of news clippings, art show announcements, reviews and other memorabilia pertaining to the life and works of Bruce Conner. Works covered include *The Music Master*, *The Child*, *Black Dahlia*, *Senorita* and *Rat Bastard* memorabilia. Items of a more personal nature include wedding and graduation announcements. Most of these images can be found in Bruce Conner's Scrapbook Vol 1, items 1 - 61 D. Some materials in Spanish.

Digital Folder 2
Scrapbook #1 and other images, digitized 2002 - 2004

Physical Description: 0.236 GB (154 files)

Scope and Content Note: Original disc labeled "Bruce Conner Scrapbook #I, 1963 - 1964-B (64 - 77c) Also: Clippings, etc. 1957. Photo of B.C. 2001 by ©Frank English. B.C. Tapestry

show 2004 at Michael Kohn Gallery, Los Angeles (digi photos). Movie stills: A movie, 10 second film, Looking for mushrooms, Cosmic ray, Marilyn times 5." Digital images of news clippings, art show announcements, reviews and other materials pertaining to the life and art of Bruce Conner, some of which map to Scrapbook Vol. 1, items 64 - 77c. In addition to materials from Conner's scrapbooks, this disc contains stills from the films *A Movie*, *10 Second Film*, *Looking for Mushrooms*, *Cosmic Ray*, and *Marilyn Times 5,* as well as other materials. Some materials in Spanish.

Digital Folder 3
Bruce Conner Art and Music, digitized 2003 - 2006

Physical Description: 0.230 GB (27 files)

Scope and Content Note: Original disc labeled "BC art + music." Digital images of art works produced by Bruce Conner in addition to photographs of the artist. Photographed art works include *Chou Rat*, *We Trussed*, and *Old Nobodaddy*. Also on this disk are Microsoft Word documents and .mp3 files. The Word documents contain comments by Bruce Conner regarding three of his works; *Abaddon*, *Knox*, and *The Last Supper*. The .mp3 files include a recording of a 2005 radio interview conducted by David Platzker for Recorded Matter and several recordings of Bruce Conner and Mel Lyman playing harmonica.

Digital Folder 4
BC Art, Scrapbooks, Images and Words, digitized 2004 - 2006

Physical Description: 0.199 GB (141 files)

Scope and Content Note: Original disc labeled "BC art, scrapbooks, images + words." Images on this disc include detailed photographs of *The Child* and a series of photographs that either was or is displayed by the San Francisco Museum of Modern Art (SFMOMA). A group of images in a folder titled "Prints, SF State" center around a collection of Conner 's fingerprints and the ownership of those prints. Also included are duplicates of some images found in Scrapbook 1 and translations of Spanish language articles pertaining to Conner's work. Word documents on this disc include an open letter "To Young Artists," advice on creating collages, and a discussion of Jay De Feo's *The Rose*. Some materials in Spanish.

Digital Folder 5
Part B of BC Scrapbook #1, digitized 2004

Physical Description: 0.196 GB (61 files)

Scope and Content Note: Original disc labeled "Part B of BC Scrapbook #1 1944 - 1963." Digital images of news clippings, art show announcements, reviews and other materials

pertaining to the life and art of Bruce Conner. Images can be found in Bruce Conner's Scrapbook Vol. 1, items 1-77.

Digital Folder 6
Bruce Conner Scrapbooks Part C, digitized 2000 - 2004

Physical Description: 0.266 GB (60 files)

Scope and Content Note: Original disc labeled "IVA. Part C of #1 = 1944 - 1963 scrapbook, #2 = 1964 - 1968 scrapbook, #3 = 1970's scraps." Digital images of news clippings, art show announcements, reviews and other materials pertaining to the life and art of Bruce Conner. Most of these images also appear in the Bruce Conner scrapbooks. In addition, the images on this disc include lithographs and various assemblages such as *The Chair*, *Looking Glass*, and *Diamond Lady Card Game*. A timeline of the artist's life is provided in several Word documents. Some materials in Spanish.

Digital Folder 7
B.C. Film and Art , digitized 2002 – 2006

Physical Description: 0.107 GB (86 files)

Scope and Content Note: Original disc labeled "B.C. Film + Art from 1950's + 1960's." Digital images of art by Bruce Conner and Jean Conner. Images of works by Bruce Conner include *Vivian*, *Resurrection*, *Oldnobody* and *The Child*. Works by Jean Conner are represented by *Floating Head* and *Young Woman and Skull*. Other files contain movie stills and filmstrips from Bruce Conner films *Toni Basil*, *10 Second Film*, *Report*, and *Cosmic Ray.* Also found on this disc are photographs that appear to have been taken during the filming of *Cool Hand Luke*, the 1967 film starring Paul Newman. Labels for art works, a bio of the artist, and commentary are contained in Word documents.

Digital Processing Resources

BitCurator, https://www.bitcurator.net/

Community Owned digital Preservation Tool Registry (COPTR), http://coptr.digipres.org

Demystifying Born Digital OCLC Reports,
http://www.oclc.org/research/themes/research-collections/borndigital.html

Digital Curation Exchange, http://digitalcurationexchange.org/

Digital Preservation Q & A, http://qanda.digipres.org/

Digital Preservation Handbook, 2nd Edition, http://handbook.dpconline.org/, Digital Preservation
Coalition © 2015.

DigiPres Commons, http://www.digipres.org/

DigiPres Commons: Tools by Function, http://www.digipres.org/tools/by-function/

ePADD (email Processing, Appraisal, Discovery and Delivery),
https://library.stanford.edu/projects/epadd

Paradigm, Workbook on Digital Private Papers, http://www.paradigm.ac.uk/workbook/index.html

Preserving (Digital) Objects With Restricted Resources (POWRR), http://digitalpowrr.niu.edu/

Works Cited

Arton, Carla. "So you want to be an Audiovisual Archivist." *AMIA Education Committee, WordPress*. May 1, 2015.
https://amiaeducomm.wordpress.com/about/so-you-want-to-be-an-av-archivist/.

"Processing Manual, Chapter 7.2, Elements of Description." *Beinecke Rare Book & Manuscript Library, Yale University.* 2015. http://beinecke.library.yale.edu/72-elements-description.

"Fixity and checksums." *Digital Preservation Handbook, Digital Preservation Coalition*. 2nd ed. 2015.
http://www.dpconline.org/handbook/technical-solutions-and-tools/fixity-and-checksums

"Encoded Archival Description Tag Library, Version 2002 and Version EAD 3." *EAD: Encoded Archival Description, The Library of Congress*. https://www.loc.gov/ead/index.html.

"Art & Architecture Thesaurus Online." *Getty Vocabulary Program, Getty Research Institute*. 2017. http://www.getty.edu/research/tools/vocabularies/aat/

Greene, Mark A., and Dennis Meissner. "More Product, Less Process: Revamping Traditional Archival Processing." *The American Archivist* 68, Fall/Winter (2005): 208-63.
http://www.archivists.org/prof-education/pre-readings/IMPLP/AA68.2.MeissnerGreene.pdf

ISAD(G): General International Standard Archival Description. 2nd ed. Ottawa: International Council on Archives, 2000.
http://www.ica.org/sites/default/files/CBPS_2000_Guidelines_ISAD%28G%29_Second-edition_EN.pdf

Matthews, Brian, Brian McIlwrath, David Giaretta, and Esther Conway. *The Significant Properties of Software.* Oxford: JISC, 2008.
https://www.webarchive.org.uk/wayback/archive/20100624233431/http://www.jisc.ac.uk/media/documents/programmes/preservation/spsoftware_report_redacted.pdf

Oxford English Dictionary. http://www.oed.com/

"PBCore Controlled Vocabularies." *Public Broadcasting Metadata Dictionary Project.*
http://pbcore.org/pbcore-controlled-vocabularies/

RDA Toolkit. https://access.rdatoolkit.org/

Guidelines for Standardized Holdings Counts and Measures for Archival Repositories and Special Collections Libraries. SAA-ACRL/RBMS Joint Task Force on the Development of

Standardized Holdings Counts and Measures for Archival Repositories and Special Collections Libraries (Affiliated Group), 2014-2016. https://www2.archivists.org/sites/all/files/SAA-RBMS_JTF-HCM01-04_GuidelinesFull.pdf

Describing Archives: A Content Standard. 2nd ed. Chicago: Society of American Archivists, 2013. http://files.archivists.org/pubs/DACS2E-2013_v0315.pdf

Software Preservation Network. http://www.softwarepreservationnetwork.org

Thornton, Katherine, and Euan Cochrane. "Wikidata as a digital preservation knowledge base." *Open Preservation Foundation*. 2016. http://openpreservation.org/blog/2016/09/30/wikidata-as-a-digital-preservation-knowledgebase/

"Best Guess Guidelines for Cataloguing Born Digital Material." *UK Archives & Records Association*. 2016. http://www.archives.org.uk/about/community/groups/viewgroup/50-best-guess-guidelines-for-cataloguing-born-digital-material.html

Guidelines for Efficient Archival Processing in the University of California Libraries, Version 3.2. University of California: Next Generation Technical Services POT 3 Lightning Team 2 with contributions from Lightning Team 2b, 2012. http://libraries.universityofcalifornia.edu/groups/files/hosc/docs/_Efficient_Archival_Processing_Guidelines_v3-1.pdf