View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by DigitalCommons@USU

Journal of Western Archives

Volume 5 | Issue 1 Article 1

2014

From Accession to Access: A Born-Digital Materials Case Study

Cyndi Shein
J. Paul Getty Trust, c.shein@yahoo.com

Follow this and additional works at: https://digitalcommons.usu.edu/westernarchives

Part of the Archival Science Commons

Recommended Citation
Shein, Cyndi (2014) "From Accession to Access: A Born-Digital Materials Case Study," *Journal of Western Archives*: Vol. 5 : Iss. 1 , Article 1.
Available at: https://digitalcommons.usu.edu/westernarchives/vol5/iss1/1

This Case Study is brought to you for free and open access by the Journals at DigitalCommons@USU. It has been accepted for inclusion in Journal of Western Archives by an authorized administrator of DigitalCommons@USU. For more information, please contact rebecca.nelson@usu.edu.

Footer Logo

From Accession to Access: A Born-Digital Materials Case Study

Cyndi Shein

ABSTRACT

Between 2011 and 2013 the Getty Institutional Records and Archives made its first foray into the comprehensive ingest, arrangement, description, and delivery of unique born-digital material when it received oral history interviews generated by some of the *Pacific Standard Time: Art in L.A.* project partners. This case study touches upon the challenges and affordances inherent to this hybrid collection of audiovisual recordings, digital mixed-media files, and analog transcripts. It describes the Archives' efforts to develop a basic processing workflow that applies the resource-management strategy commonly known as "MPLP" in a digital environment, while striving to safeguard the integrity and authenticity of the files, adhere to professional standards, and uphold fundamental archival principles. The study describes the resulting workflow and highlights a few of the inexpensive technologies that were successfully employed to automate or expedite steps in the processing of content that was transferred via easily-accessible media and consisted of current file formats.

Introduction

Modern society creates and stores its personal histories and professional records in bits and bytes, chiefly rendering the documentation of contemporary culture in born-digital form. In spite of the growing prevalence and importance of unique born-digital resources in contemporary archives, many archival repositories have yet to responsibly address their born-digital holdings, citing the lack of funding, time, and expertise as the main impediments.[1] While contemplating the comprehensive stewardship of born-digital resources can be overwhelming, implementing incremental steps toward their management is within reach for most repositories. This case study traces the J. Paul Getty Trust Institutional Archives' first effort to manage an incoming born-digital collection from the time of transfer to the time of public dissemination. The paper discusses some of the challenges encountered, decisions made, and workflows developed by archival staff while handling this hybrid

1. Jackie M. Dooley and Katherine Luce, "Taking Our Pulse: The OCLC Research Survey of Special Collections and Archives," Dublin, Ohio: OCLC Research (2010), 60, http://www.oclc.org/research/publications/library/2010/2010-11.pdf (accessed May 1, 2013).

collection of audiovisual recordings, born-digital mixed-media files, and printed transcripts. Since the Archives embarked on this endeavor with no substantial experience processing born-digital materials and completed it primarily using free user-friendly tools, this paper offers a reasonable starting point for repositories on the threshold of digital curation—even those with inexperienced staff and limited means.

Literature Review

Over the years The Getty Research Institute working groups have examined standards, projects, and professional literature to guide the Institute as it moves toward a sustainable program for digital stewardship—much of that research is outside the scope of this paper, which focuses primarily on developing a pragmatic approach to accessioning, processing, and delivering a current hybrid collection. Although literature relevant to managing born-digital collections has been published since the 1990s,[2] more recent publications have the advantage of presenting case studies and strategies using the latest technologies. Between 2008 and 2009 information professionals began sharing intensive high-end approaches for the treatment of born-digital special collections materials. These publications were followed by papers presenting less intricate approaches that were more feasible in environments with financial or technological constraints, yet still appropriate for special collections materials. As the field of born-digital stewardship is maturing, practices appropriate to a wider spectrum of situations are emerging. Although the growing body of literature now documents a range of behind-the-scenes processes, there are still few studies addressing the nuts and bolts of providing access to born-digital special collections and archives, and still fewer presenting strategies for enabling online public access.

The most prominent findings on born-digital archives published between 2008 and 2011 are centered on complex hybrid manuscript collections comprised of legacy data on obsolete media, such as projects completed by The British Library,[3] Bodleian

2. Among the notable literature of the 1990s are: Adrian Cunningham, "The Archival Management of Personal Records in Electronic Form: Some Suggestions," *Archives and Manuscripts* 22, no. 1 (May 1994): 94-105; Tom Hyry and Rachel Onuf, "The Personality of Electronic Records: The Impact of New Information Technology on Personal Papers," *Archival Issues* 22, no. 1 (1997): 37-44; and Jeremy Leighton John, "Adapting Existing Technologies for Digitally Archiving Personal Lives: Digital Forensics, Ancestral Computing, and Evolutionary Perspectives and Tools" (paper presented at iPRES 2008: The Fifth International Conference on Preservation of Digital Objects, London, UK, September 29-30, 2008), http://www.bl.uk/ipres2008/presentations_day1/09_John.pdf (accessed March 15, 2011).

3. Leighton John.

Library at Oxford,[4] Beinecke Rare Book and Manuscript Library at Yale,[5] Emory University Library, Harry Ransom Center at the University of Texas at Austin, and Maryland Institute for Technology in the Humanities (MITH) at the University of Maryland.[6] While published reports on these projects remain relevant and convey useful information on everything from pre-custodial donor relations to researcher expectations for access, descriptions of their tools and processes lean heavily toward digital forensics. These groundbreaking projects provide an invaluable foundation for the development of born-digital stewardship, but they come from a very similar and limited perspective—that of large institutions with solid funding and expert technical support, working on high-profile humanities collections that merit the emulation of the creators' computing environments and/or granular (often file-level) description of the content. Until recently, perspectives from and processes applicable to smaller repositories, under-funded programs, and novice born-digital archivists have not been well represented in the literature.

During a 2010 Rare Books and Manuscripts Section (RBMS) presentation Emory University Library impressed the audience with the accomplishments of its multidivisional working groups, followed by Stanford University's awe-inspiring presentation on FRED (Forensic Recovery of Evidence Device)—and then a modest voice from the American Heritage Center (AHC) offered hope to repositories that felt overwhelmed by the very thought of managing born-digital materials. Ben Goldman (AHC) recognized the reality that, in the absence of dedicated funding and technical expertise, many repositories would be doing the best they could with the resources they had, and he encouraged them to do just that. Goldman's RBMS presentation and subsequent writings[7] are important in the development of born-digital stewardship because they are among the first to address the realities and constraints common to

4. Susan Thomas, "Curating the I Digital: Experiences at the Bodleian Library" in I *Digital: Personal Collections in the Digital Era*, ed. Christopher A. Lee (Chicago: Society of American Archivists, 2011), 280-305.

5. Michael Forstrom, "Managing Electronic Records in Manuscript Collections: A Case Study from the Beinecke Rare Book and Manuscript Library," *American Archivist* 72. no. 2 (Fall/Winter 2009): 460-477.

6. Matthew Kirschenbaum, Richard Ovenden, and Gabriela Redwine, *Digital Forensics and Born-Digital Content in Cultural Heritage Collections* (Washington D.C.: Council on Library and Information Resources, December 2010), http://www.clir.org/pubs/reports/pub149/pub149.pdf (accessed March 7, 2011); and Matthew Kirschenbaum, et al., "Approaches to Managing and Collecting Born-Digital Literary Materials for Scholarly Use," May 2009, http://drum.lib.umd.edu/bitstream/1903/9797/1/Born-Digital%20White%20Paper.pdf (accessed March 30, 2011).

7. Ben Goldman, "Moving Forward with Born-Digital Manuscripts" (paper presented at the Rare Books and Manuscripts Section of the American Library Association meeting, Philadelphia, 2010, http://www.rbms.info/conferences/preconfdocs/2010/SeminarIGoldman.pdf (accessed April 10, 2011); Goldman, "Bridging the Gap: Taking Practical Steps Toward Managing Born-Digital Collections in Manuscript Repositories," *RBM: A Journal of Rare Books, Manuscripts, and Cultural Heritage* 12, no. 1 (2011): 11-24; and Goldman, "Using What Works: A Practical Approach to Accessioning Born-Digital Archives" (guest post on Chris Prom's *Practical E-Records*, June 23, 2011), http://e-records.chrisprom.com/guest-post-ben-goldman/ (accessed February 15, 2012).

many repositories. In contrast to the intensive processes previously represented in the literature, Goldman suggests a "humble" process listing fundamental steps that can serve as a general structure for a born-digital workflow. Along those same lines, OCLC compiled a report in 2013 for repositories "wondering where to begin" and outlined basic steps for transferring born-digital content "from media you can read in -house."[8] The strength of the OCLC report lies in its step-by-step format and the explanation of specific hardware and software options related to each function/step, as well as a helpful list of resources and workflows.

While the work of Goldman and OCLC are of tremendous value in guiding repositories as they learn to gain control over their born-digital holdings, the literature also holds ample guidance for more mature programs. Concurrent with the work being performed at The British Library, Emory University Library, Harry Ransom Center, Maryland Institute for Technology in the Humanities (MITH), Bodleian Library, and Beinecke Rare Book and Manuscript Library, a collaboration known as the AIMS Work Group (2009-2011) was tackling the born-digital challenge from a wider perspective. The AIMS project partners, made up of University of Virginia Library, Stanford University, University of Hull, and Yale University, created a framework for digital stewardship that is not institution- or project-specific, but instead considers the broader archival community. The AIMS paper[9] asserts that conventional archival principles and standards are still relevant in the digital age. The Society of American Archivists' (SAA) Digital Archives Specialist (DAS) course materials[10] and SAA's 2013 publication *Archival Arrangement and Description*[11] support this assertion by proposing born-digital workflows that, while making adaptations and accommodations for digital materials, are still fundamentally built upon existing workflows for physical archives.

In the interest of efficiency, both existing archival workflows and their digital workflow offspring allow for nuanced approaches to arrangement and description as well as the incorporation of automated processes. The AIMS paper, the book *Archival*

8. Julianna Barrerea-Gomez and Ricky Erway, "Walk This Way: Detailed Steps for Transferring Born-Digital Content from Media You Can Read In-House" (Dublin, Ohio: OCLC Research, 2013), http://www.oclc.org/content/dam/research/publications/library/2013/2013-02.pdf (accessed August 16, 2013).

9. AIMS Work Group, "AIMS Born-Digital Collections: An Inter-Institutional Model for Stewardship," (2012), 41, http://www.digitalcurationservices.org/files/2013/02/AIMS_final.pdf (accessed May 18, 2013).

10. Among the DAS courses clearly illustrating this point are: *Managing Electronic Records in Archives and Special Collections* (as taught by Seth Shaw and Nancy Deromedi, December 2011) and *Arrangement and Description of Electronic Records Part I & II* (as taught by Christopher J. Prom, March 2013).

11. J. Gordon Daines III, "Module 2: Processing Digital Records and Manuscripts," in *Archival Arrangement and Description*, ed. Christopher J. Prom (Chicago: Society of American Archivists, 2013), 87-144. Pages 100-110 outline essential born-digital accessioning and processing steps that strongly echo traditional workflows, while pages 111-125 provide details on how these workflows can be adapted to accommodate born-digital materials.

Arrangement and Description, and SAA's DAS curriculum all acknowledge the validity of different levels of description and arrangement for different materials (endorsing the application of the MPLP approach)[12] and accept accessioning as a form of baseline processing in reference to born-digital materials.[13] Accepting "minimal processing" and/or "accessioning as processing" as viable options in the handling of born-digital materials meets a documented need for flexible and scalable workflows. The need for more efficient and extensible workflows was identified as early as 2008 by the National Library of Australia (NLA) when it found its existing workflows were unable to address the volume and risks associated with born-digital materials at the necessary pace and realized they "had to move away from hand-crafting to a more industrial way of processing this material."[14] NLA's suggestion was to automate processes to relieve staff of repetitive or labor-intensive actions. The Library of Congress National Information Infrastructure and Preservation Program (NIIPP) Digital Preservation Sustainability Group made similar recommendations in 2009.[15] Today automation is a commonly held goal, and many publications and presentations discuss ways archivists are achieving that goal with commercial software (Forensics Toolkit and FRED) and open-source utilities (Duke DataAccessioner, BitCurator, Archivematica, and Curator's Workbench).[16]

While publications on technologies and methodologies facilitating the management of born-digital materials are more prevalent than they were five years ago, a significant gap in the literature remains in reference to repositories providing access to born-digital special collections and archives, particularly in regard to providing online public access. While providing access has long been the driving force behind archival processing and preservation, enabling access to born-digital

12. For more details on the resource management strategy commonly known as "MPLP" see: Mark A. Greene and Dennis Meissner, "More Product, Less Process: Revamping Traditional Archival Processing," *American Archivist* 68, no. 2 (Fall/Winter 2005): 208-263; Dennis Meissner and Mark A. Greene, "More Application while Less Appreciation: The Adopters and Antagonists of MPLP," *Journal of Archival Organization* 8, no. 3/4 (2010): 174-226.

13. The AIMS Work Group paper "AIMS Born-Digital Collections: An Inter-Institutional Model for Stewardship" discusses baseline processing (pages 17-18) and factors determining levels of description (41). One of the steps in Daines' Sample Processing Workflow is to "Identify the appropriate level of description" (110). DAS workbooks for *Arrangement and Description of Electronic Records I & II* workflows mention using "the appropriate level of description" several times. DAS handouts for *Managing Electronic Records in Special Collections and Archives'* Basic Workflow's first point under Arrange/Describe is "Consider series, depth of description…" (Section 11 of December 2011 handout).

14. Douglas Elford, et al., "Media Matters: Developing Processes for Preserving Digital Objects on Physical Carriers at the National Library of Australia" (presented at the World Library and Information Congress: 74th IFLA General Conference and Council, Quebec, August 2008), 4, 6-11, http://archive.ifla.org/IV/ifla74/papers/084-Webb-en.pdf (accessed April 13, 2011).

15. William G. Lefurgy, "NDIIPP Partner Perspectives on Economic Sustainability," *Library Trends* 57, no. 3 (Winter 2009): 421.

16. See Barrerea-Gomez and Erway for more information on these technologies.

collections is still in a developmental phase. The AIMS Work Group devotes an entire section of its paper to discovery and access, but online delivery was still under development at the partner institutions at the time of publication.[17] Most of the publications on born-digital processing and delivery discuss providing limited access from dedicated workstations in reading rooms.[18] A 2012 Academic Research Library (ARL) survey shows that the main obstacle to born-digital collections access is the sensitivity of materials, closely followed by a lack of technical infrastructure. Although the survey reports that 66% of respondents provide "online access to a digital repository system" (remotely in an unmonitored space), the publication does not specify whether or not the access is public or restricted (on-campus or in-library only); nor does it provide examples of born-digital collections that are publically available online. [19]

The best-known examples of public online access to born-digital collections are found in presentations and publications related to Bentley Historical Library at the University of Michigan and University of California, Irvine (UCI) Special Collections and Archives. Through presentations and a case study, Nancy Deromedi has shared information about the Bentley's groundbreaking 1997-1998 processing of the hybrid papers of James J. Duderstadt.[20] Duderstadt's digital papers are described in a finding aid and are publically available online. Likewise, UCI has published on its processing of the Richard Rorty papers and the Mark Poster papers, which culminated in the daring implementation of a virtual reading room that allows online access to anyone who agrees to UCI's terms of use.[21] Michelle Light's 2013 presentation and

17. While the AIMS Work Group paper provides access models (pages 51-55) and discusses the Bodleian's "Publication Pathway" (56), no collection names or web addresses are given for examples of collections that are publically available online. Laura Wilsey, et al., "Capturing and Processing Born-Digital Files in the STOP AIDS Project Records: A Case Study," *Journal of Western Archives* 4, no. 1 (2013): 19, http://digitalcommons.usu.edu/westernarchives/vol4/iss1/1 (accessed December 2, 2013). Mentions that Hypatia, the AIMS project's program for preservation and access, was still under development at the time of publication.

18. Kirschenbaum, et al., Harry Ransom Center, Emory University and MITH mention providing onsite-only access to emulated computing environments. Thomas (pages 299-300) speaks of online public access only in the future tense. Wilsey, et al., (19) state that until Hypatia is ready, access from a dedicated workstation is the "short-term solution" at Stanford.

19. Naomi Nelson, et al., *SPEC Kit 329 Managing Born Digital Special Collections and Archival Materials* (Washington DC: Association of Research Libraries, 2012), 17, 76-82; when asked to name the biggest challenge to discovery and access, 50% of respondents cited the sensitivity of materials and 44% cited the lack of technical infrastructure.

20. Nancy Deromedi, "Case 1: Accessing, Processing, and Making Available a Born-Digital Personal Records Collection at the University of Michigan," University of Michigan 2006, http://bentley.umich.edu/academic/france/inp/docs/case1.pdf (accessed November 28, 2011); Deromedi and Shaw, *Managing Electronic Records in Archives and Special Collections*.

21. Dawn Schmitz, "The Born-Digital Manuscript as Cultural Form and Intellectual Record" (presented at "Time Will Tell, But Epistemology Won't: In Memory of Richard Rorty," Irvine, California, May 14, 2010, http://www.escholarship.org/uc/item/5ss5696t (accessed June 1, 2010).

forthcoming case study discuss UCI mitigating the risks of opening the papers online via a virtual reading room and how zipping files or packaging them as complex digital objects prevented (intentionally or not) discovery of file contents by search engines.[22] Efforts at UCI and Bentley demonstrate means for publically opening collections online without overexposing potentially sensitive materials.

When considering how to effectively and ethically process and deliver born-digital materials, J. Paul Getty Trust Institutional Archives was informed and inspired by the activities described in the literature. We decided to make this study available, in spite of the draft state of our current workflows, because what we've learned so far could be helpful to fellow archivists. We believe the strength of the existing literature lies not only in its common motivation to exchange ideas relative to preserving and providing access to born-digital archives, but also in the diversity of the perspectives and strategies offered toward the accomplishment of that goal. By critically reviewing existing publications and monitoring current activities in the field, a repository can selectively adopt tools and adapt methods from a variety of sources to formulate the approach that best suits it. By sharing our successes and failures, professionals can build upon the collective work of others to further advance not only our individual programs, but the profession as a whole.

Institutional Context and Collection Background

"The J. Paul Getty Trust is a cultural and philanthropic institution dedicated to critical thinking in the presentation, conservation, and interpretation of the world's artistic legacy."[23] The Trust accomplishes its mission through the collective work of its programs: the Museum, Conservation Institute, Research Institute, and Foundation. The Getty Institutional Archives, a department of The Getty Research Institute, supports the mission of the Trust by selecting, preserving, and making available permanently valuable institutional records, in all media, of past, current, and future programmatic units of the Trust. It was therefore the Archives' responsibility to assemble, secure, and disseminate documentation on The Getty's project, *Pacific Standard Time: Art in L.A. 1945-1980*.

Pacific Standard Time: Art in L.A. 1945-1980 was an unparalleled collaboration of more than sixty of Southern California's cultural institutions working together to tell

22. Michelle Light, "Born Difficult?" (presented at "Past Forward! Meeting Stakeholder Needs in 21st Century Special Collections," Yale University, New Haven, Connecticut, June 4, 3013), slides and video, http://www.oclc.org/research/events/2013/06-03.html (accessed July 23, 2013); Michelle Light, "Managing Risk with a Virtual Reading Room: Two Born-Digital Projects," forthcoming in *Innovative Practices in Archives and Special Collections: Reference and Access* (Lanham, Maryland: Scarecrow Press, 2014).

23. The J. Paul Getty Trust, "About The J. Paul Getty Trust," http://www.getty.edu/about/trust.html (accessed May 1, 2013).

the story of the birth of the vibrant Los Angeles art scene.[24] The project was the culmination of a long-term Getty initiative focusing on postwar (1945-1980) art in L.A. From 2011 to 2012, during the peak of the project, The Getty and its partner institutions communicated their findings through a multitude of simultaneous public exhibitions and events. Project partners conducted interviews with many of Los Angeles' key artists, filmmakers, curators, collectors, and critics. These oral histories were central to the project's research and were featured in various exhibitions and publications.

Many of the participating organizations were awarded funding by The Getty Foundation. As the exhibitions began to open across California, The Getty Research Institute (GRI) requested that copies of the recordings and transcripts of oral history interviews conducted by grant recipients be added to The Getty's archival record of the project. The deposit of oral history interviews became part of the grant requirement after most of the interviews and documentation had already been produced, so it was too late for the Foundation to impose conditions regarding creator-generated metadata, documentation, file formats, or preferred transfer methods. As such, The Getty accepted the interviews in whatever forms they were offered. The nature of the records reflects the varying levels of expertise and/or commitment each institution brought to the project.

Between April 2011 and March 2013 the Getty Institutional Archives received over 200 oral histories generated by nineteen of the *Pacific Standard Time: Art in L.A. 1945-1980* project partners. The interviews were transferred to the Institutional Archives with the expectation that we would provide broad access to the resources in a timely manner. Although the Institutional Archives had been accepting and providing access to born-digital files on demand via ad-hoc methods for some time, we had only recently begun to formally develop policies and procedures to govern the comprehensive stewardship of born-digital resources. A few months prior to the arrival of the interviews, The Getty Research Institute (GRI) had appointed a Born-Digital Materials Steering Committee "to define, recommend, and, where possible, implement the policies, procedures, and technological structures/systems required to govern the process of managing, from acquisition to permanent preservation, the born digital materials of the GRI, particularly Special Collections and Institutional Archives."[25] The first wave of *Pacific Standard Time* oral histories arrived before the committee had even finalized its recommendations. The Getty Institutional Archives had no funding or personnel allocated for born-digital collections/ records management, but the literature supported proceeding rather than waiting until we had all the details worked out. We needed to think big (consider scalable, extensible

24. The J. Paul Getty Trust, "Pacific Standard Time: Art in L.A. 1945-1980," http://past.pacificstandardtime.org/ (accessed May 2, 2013).

25. The Getty Research Institute Born-Digital Materials Steering Committee, "Revised Charge," (March 2011).

models for the future), but start small (do something now). Though our ultimate goal is to establish a trusted digital repository and a sustainable program for the stewardship of born-digital material, it was immediately apparent that we could not leap from nothing to a fully-developed program without taking some incremental steps. We took the approach expressed by Ben Goldman and began "moving forward with practical and achievable steps, and in developing the institutional framework to tackle the issue with greater complexity at some future point."[26]

Objectives and Priorities

The *Pacific Standard Time: Art in L.A. 1945-1980* project was an institutional priority at The Getty and the result of more than a decade of work. Aside from the underlying goal to ingest and preserve the content, the Archives' main objective was to enable discovery of and access to the materials promptly while the related exhibitions were still on display across Southern California. Given the Institutional Archives' broad responsibility to care for the records of the entire Trust, combined with the varied quality and completeness of the incoming materials, several questions immediately arose. While addressing these questions, our primary consideration was the Archives' and the Foundation's commitment to honor the interviewees' intentions to share their stories with the community, but we were also very aware of our limited time and resources. The following questions and decisions largely determined the extent of our archival management of the materials:

In light of our department's small staff and competing priorities, what resources should the Archives allocate to this endeavor? Since the Archives had no advance notification of its role in the management of these materials, there was no opportunity to write a proposal requesting funding for this endeavor—we knew we would have to do what we could with the resources we had at hand. GRI's Institutional Archives has only one full-time professional dedicated exclusively to archival management (versus staff who are more focused on records management). Her primary responsibilities include accessioning, processing, MARC cataloging, managing interns, reference, and general collection management at two different sites. It was decided that she should commit no more than 20 percent of her time (calculated on a monthly basis) to the *Pacific Standard Time* transfers as they arrived. The prioritization of timely access to the materials precluded the development of a customized delivery system; the Archives would have to use the existing tools and technologies available through our parent institution—no new expensive software, no dedicated interface, no bells or whistles. Materials received in digital format would be made available through our digital repository; however, in the interest of time, materials deposited only in print would be made available on site, but not digitized during our normal workflow. Per existing local policy, digitization could be performed later upon request.

26. Goldman, "Bridging the Gap."

Should one missing release in an accession hold up the dissemination of the entire accession? To honor the interviewees' wishes to share their stories, we decided to make publically available all of the interviews for which we had releases, regardless of the rights limitations of their archival siblings.

Should poor audiovisual quality prevent or delay access? Although we received a final edit of some of the interviews, some of the audio and video arrived in a raw form that would generally not be considered ready for an audience. Performing post-production editing or basic quality control (normalizing sound, adjusting brightness, stitching together files stored on separate media carriers, etc.) would require time we simply did not have. Accordingly, we decided to make the interviews available in the state in which they were received.

Should incomplete documentation prevent or delay access? Based on the lack of information accompanying the sets of interviews we initially received, it was clear that documentation of the materials would in many cases fall short of our usual standards. Given that The Getty asked contributing institutions for interview recordings and transcripts after the interviews had been conducted, we could not expect the contributors to supply information they had not collected during the interview process. With our primary goal being access, we decided that as long as the interviewee was identified, we would make the interview available (rights permitting).

What level of description and arrangement (virtual and physical) could we afford to perform on these records? Creating item-level description would be too labor-intensive for one person (devoting less than 20 percent of her time) to complete the collection in a timely manner. Other than addressing any conservation concerns, we decided to perform only the work that was necessary to make the interviews discoverable and usable, and we determined that aggregating and describing the electronic and printed materials together at the accession level (rather than the item level) would serve that end. To compensate for the absence of MARC records, each interviewee's name would be listed in the collection-level EAD record, and in the accession-level MARC and Dublin Core (DC) records to aid in the discovery of individual interviews.

Affordances and Challenges

Nature of the Material

As with any collection, the *Pacific Standard Time: Art in L.A. 1945-1980* oral history records came with innate affordances and challenges: the very nature of the files facilitated processing, but the indirect transfer method and creators' lack of documentation of their own records presented some difficulties. The materials are primarily comprised of audiovisual interview recordings, textual transcripts, and some form of signed agreement permitting each interview to be transformed and disseminated. Unexpectedly the material also includes a number of symposia

recordings, documentary videos, digital still images, exhibition-planning material, and (in one case) accompanying project documentation. Although the Archives would have preferred preservation-quality files, The Getty did not generally receive the original or archival master recordings, but more often a derivative format that we consider a service copy. Ultimately we received printed material, 1 hard drive, 26 videotapes, and 276 optical discs—totaling 1.52 terabytes and 8.5 linear feet. Electronic file formats received include, but are not limited to, MOV, VIDEO_TS (.vob, .ifo, .bup), MP4, MP3, CDA, WAV, PNG, TIFF, JPEG, DOC, DOCX, PDF, XLS, and XLSX.

Affordances

This being our first attempt to widely disseminate a hybrid collection, it was to our advantage that it was a strong example of the proverbial "low-hanging fruit." In spite of its challenges, several characteristics inherent to the collection indicated that successful processing and delivery was within our reach. *The Pacific Standard Time: Art in L.A. 1945-1980* materials offered many affordances that made it a good candidate for our first hybrid processing attempt:

- The collection is of modest size by digital standards.
- The content is current and from trusted sources—we took a calculated risk and decided not to quarantine the content during ingest.
- The hard disk drive and the optical discs were specifically created for the transfer of target files, so we saw no reason to capture the computing environment or image the discs.[27]
- The content is comprised of current file formats (with recognizable file extensions) on easily accessible media, requiring no forensics work.
- The files contain no known personally identifiable information.
- Most creators kept the originals, so our lack of expertise would not endanger the one-and-only version of the content.
- Most interviews were accompanied by some form of a signed release that permits editing, transformation, and broad dissemination of the content.
- There was no hard deadline, giving us some room for investigation and experimentation.

27. Approximately six months after the processing of this collection was completed we discovered a compelling reason to image optical discs that contain TS_VIDEO file formats. Details are given later in this study in the section on accessioning and ingesting optical discs.

Challenges

Although we took courage from the advantages afforded by the nature of the records, we still faced some formidable obstacles, including:

- Absence of local policies, procedures, and technical infrastructure.
- Incomplete submissions.
- Clarification of rights.
- Expectation for timely online access.

Policies, Procedures, and Infrastructure

As mentioned above, when the records began to arrive in early 2011, the Institutional Archives and Special Collections departments were in the earliest phase of developing formal policies for handling born-digital materials, and we had no technical infrastructure or procedures in place for their management. In order to meet the challenges ahead, we first needed to establish a computing environment in which we could begin to carry out our mandate to steward electronic records. Using this collection, which was generated by a high profile Getty-wide initiative, to demonstrate urgent need,[28] the Head of Institutional Records and Digital Stewardship prevailed upon Information Technology Services (ITS) to create a networked "Locked" server with 6 terabytes (TB) of space for Institutional Archives that is fully accessible to only two staff members. More space is allocated as needed. (At time of publication it is at 13TB.) ITS simultaneously created a separate networked space (2 TB) that serves as a temporary processing space or "Workbench" for archival staff.

Submissions

The main challenges associated with the management of this collection spring from the variety of the nineteen record creators (museums, galleries, universities, and other cultural organizations) and the Archives' lack of influence over the submissions we received. The Foundation administered the grants, collected the interviews from the creators, and then transferred the interviews to the Archives. The initial sets of interview materials arrived in the Archives without warning, and the Archives staff had no opportunity for pre-custodial conversations with the creators. Consequently,

28. For nearly a year the Head, Institutional Records and Digital Stewardship had been building a case regarding server space—incoming digital transfers had been saved to external hard drives. Unprocessed digital transfers were not backed up and had reached critical mass. The need for server space for the oral history files raised the awareness of decision makers since the *Pacific Standard Time* project was the center of institutional attention at that time.

the Archives devised a list of preferred file formats so the Foundation may communicate submission standards to contributing institutions in the future. The most significant challenge related to the submission of materials was the absence of accompanying metadata and the scant documentation provided by creators. Because the file formats and accompanying documentation differed wildly from one originating institution to the next, no single programmatic strategy could harvest metadata from the existing documentation, compelling us to enter descriptive metadata manually to the files submitted to discovery systems.

Rights

Most submissions included some form of agreement meant to grant The Getty permission to provide access to the materials; however, the variety of contracts posed a challenge. We determined rights for each interview by classifying them into three categories: signed Getty contract on file; signed non-Getty contract on file; and no contract or release received. Due to the potentially sensitive nature of oral histories, decisions about access had to be made carefully at the accession or item level. The Getty contracts transfer full rights, allowing us to transform and disseminate those interviews online.[29] The non-Getty contracts were less explicit; after reviewing the contracts, Legal Counsel determined that the intent of the non-Getty releases was to provide online public access to the interviews within the context of the *Pacific Standard Time* project and it was therefore essential that the Archives present the interviews within that context. We accomplished this by nesting the interviews within MARC, DC/METS, and EAD records that clearly explain the context of the interviews in a summary, abstract, or scope and contents note. Thus we were able to provide public access to all interviews for which we received some form of signed release. If no release was received for an interview, that interview was not made accessible in any way.

The discovery of unexpected content in the interview submissions raised further rights-related issues. While the Foundation requested the submission of oral history interviews, it received additional unexpected (and valuable) content, such as symposia recordings, still images, short documentary videos, and documentation of exhibition planning, all of which presented different sets of rights and access issues. Given the intellectual property rights inherent to works of art, we do not plan to post images of works of art or the associated exhibition planning files online, but will provide access to them in our reading room upon request. The symposia and artists' talks were public events—although no contracts were received, there were no privacy issues involved. Accordingly we are able to provide access to recordings of public

29. The standard release form used for these oral history interviews includes the following phrase: "I hereby grant and assign to the Institution all right, title, and interest in and to the Interview Materials, including, without limitation, the rights to reproduce, edit, publish, distribute and display the Interview Materials publicly in all media now known and hereafter devised, and prepare derivative works based on the Interview Materials."

events onsite, but lack permission to disseminate them globally. The documentary videos were accompanied by permissions and we are free to share them online. The mixture of rights/access restrictions within a single accession led to technical difficulties—the inability of our digital repository to deliver restricted and non-restricted files within the same digital object—an issue that is addressed later in this study.

Expectation for Online Access

The final challenge presented by this collection was the expectation that the Archives provide widespread access to it quickly, which was unprecedented for our department. The Getty is a comparatively young institution; much of the Institutional Archives' 7000+ linear feet of holdings are relatively recent institutional records that are not open to the public for 35 years from the date of creation. Although the Institutional Archives regularly services internal requests and occasionally provides public access for all manner of analog records, the Pacific Standard Time interviews represented the first time we were expected to provide public online access to current born-digital materials. To fulfill this responsibility, we had to incorporate new steps into our workflow and seek out new technologies to automate some of those steps.

Workflow

Fundamental Principles

There is no one-size-fits-all workflow for archival processing, whether the resource is analog or born-digital. Existing professional archival practices, ethics, and standards can (and should) be adapted and applied to the development of local policies and procedures for managing born-digital collections. At the Getty Institutional Archives we consistently apply the resource management concepts articulated in "More Product, Less Process" (MPLP), resulting in processing plans that attempt to balance our commitment to equitable access with the limited resources at our disposal. We embrace the efficiency of processing while accessioning[30] and leverage the knowledge gained during initial familiarization with a collection to simultaneously accession and process materials whenever feasible. During accessioning, while the details of the content and context are fresh in our minds, we either formulate a processing plan (if processing is imminent) or record processing ideas and suggestions for future reference.

We began formulating a processing plan for the *Pacific Standard Time* interviews as we ingested the first few accessions and became familiar with the content and file formats in the collection. Our objectives were similar to those for analog records: to

30. For more information about performing processing during accessioning, see Christine Weideman, "Accessioning as Processing," *American Archivist* 69, no. 2 (Fall/Winter 2006): 274-283.

gain intellectual control over the files, confirm rights, prepare material for access, create access points to facilitate discovery, and deliver the content. In developing our approach to this hybrid collection, we consulted several published born-digital workflows,[31] some of which account for myriad variables and are more complex than schematics for the Space Shuttle. Variables impacting a workflow might include: the desired outcome for the collection, the presence/absence of a digital repository with/without automated preservation features, available hardware and software systems, storage options and desired amount of redundancy, available methods of delivery, rights considerations, the presence/absence of a pre-custodial donor interview, whether or not appraisal is appropriate, the presence of potentially sensitive information, whether or not the institution can read the carriers in-house or if it needs to outsource them, and whether or not the file formats are easily accessible or require forensic recovery. In the interest of simplicity, we created a workflow that focused only on the situation and material at hand, but provided enough detail to guide staff that had never before worked with born-digital collections.

The workflow does not specifically address preservation steps because we're still investigating the most sensible strategy for ensuring the persistence of records in the digital preservation system to which we are currently migrating. Our workflow does consider the future viability and security of the content data—we save the data as received on a networked server, create a manifest (including a baseline checksum) for each accession, add contextual metadata, and create a unique persistent identifier for each digital object that is ingested into the repository.[32] By transforming the less stable file formats into more trusted formats (CDA into WAV, DOC into PDF/A, etc.) we improve the chances of survival for the files themselves. We embed metadata (descriptive and administrative) into transformed files to enable their identification should they be separated from externally-stored metadata. We create external descriptive and administrative metadata (DC/METS) that provides fundamental information about the content of the files as well as their original contexts. We create structural metadata (METS) that protects vital connections between archival objects, showing part-to-whole relationships and associating related versions of the same content. For example, the structural metadata preserves: the relationship between the preservation master, modified master, and access copy of a given file; the relationship of different formats of the same or similar content, such as a video and its corresponding transcript; and the relationship between numerous digital objects that comprise a single collection. (The technical metadata was automatically generated by the software that was used to create the files and existed prior to file transfer.)

31. We consulted the OAIS model, the AIMS Work Group paper, and sources mentioned in the literature review. We also conducted a Google image search for "Born-Digital Workflow."

32. For an introduction to the Preservation Description Information of the OAIS model, see Brian F. Lavoie, "The Open Archival Information System Reference Model: Introductory Guide," OCLC Online Computer Library Center, Inc. and Digital Preservation Coalition (2004), 12, www.dpconline.org/docs/lavoie_OAIS.pdf (accessed October 21, 2012).

Since the Institutional Archives did not have the opportunity for pre-custodial intervention, the workflow begins at the time of transfer and proceeds through the delivery of the content. Although the workflow is presented in a somewhat linear fashion, it is common for steps to overlap or occur simultaneously. For example, the archivist may enter descriptive information into the finding aid while ingesting data, or may create METS records while file format conversions are running in the background. See Appendix A for diagrams of draft workflows.

Transfer and Appraisal

The accessions that comprise this collection were transferred to the Archives internally from the Foundation; transfers were made periodically as the materials were received by the Foundation from outside organizations. Most of the oral history interview recordings and transcripts were transferred to the Foundation in digital format via optical discs or portable hard disk drives. Some transcripts were sent via email, while others arrived in hardcopy (with no digital representation). Because the material was solicited by The Getty, the only form of appraisal that was appropriate was the disposition of obvious duplicates.

Accessioning

Establishing basic physical and intellectual control

Although the materials were transferred to the Archives from a single source (the Foundation as collector) the Archives assigned a separate accession number to each contributing institution to facilitate rights and collection management.[33] Accessioning each institution's records as discrete units also enabled us to process them and make them available as they arrived rather than waiting until all expected transfers had been received. The Institutional Archives views ingest as a primary and critical action of the accessioning process. The Archives uses Archivists' Toolkit (AT) for collection management.[34] We created an accession record in AT that describes the traditional accession information (creator, dates, intellectual content, physical extent, rights, etc.) as well as the digital information. We recorded descriptive and contextual information about the content in the Description field under Accession Notes. We used the External Documents field under Accession Notes to record the path (HREF) to the "original" content data (received version) on the server. We customized some of the user-defined fields (Figure 1) to quantify the accession's digital volume in megabytes (MB). Using a single unit of measurement (MB) for all accessions (even if an accession might be more efficiently measured in KB, GB, or TB) allows us to automate reports that estimate the extent of the digital holdings or calculate the

33. This approach also worked well for the Foundation's ongoing administration of the grants because it provided a one-to-one match between a grant number and its corresponding accession number.

34. The Institutional Archives is planning to migrate from Archivists' Toolkit (http://www.archiviststoolkit.org/) to ArchivesSpace (http://www.archivesspace.org/about/).

combined digital volume of selected accessions. Within the user-defined fields we also designated a field for free text entry of file formats, transfer media, ingest or data failure, and actions performed during the curation of the digital files. We labeled this field "Digital File Management" and use it to record events and actions performed on the files.

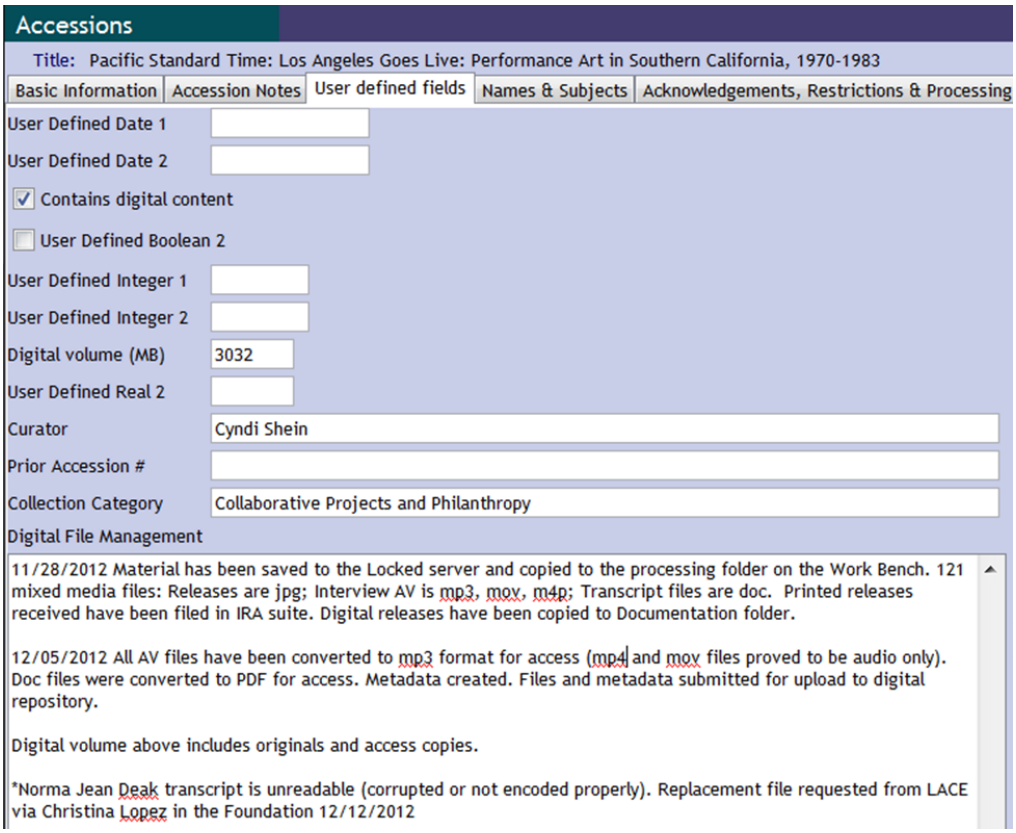


Figure 1. Screen capture of customized user-defined fields in Archivists' Toolkit

Transferring original content data from media carriers to networked server

Getty Institutional Archives' policy allows for known content from trusted sources transferred via the network, email, optical discs, and portable hard disk drives to be directly ingested into the Locked server during accessioning. With regard to the *Pacific Standard Time* oral histories, we were interested in and responsible for only the sets of files intentionally transferred to us—there was therefore no reason to look for hidden or deleted files. We ingested only the target files and did not create disc images or run the content through our forensics workstation. At the time of transfer

we were able to easily identify file formats and extents, noting disc failures and file formats that may prove problematic during processing.

During the ingest phase of accessioning, we reviewed the content of unlabeled discs, identified mystery files,[35] and compared the electronic files received with the contracts received to detect missing contracts, confirm legal custody, and determine our rights to manage, transform, and disseminate the content.

Optical discs (CD, DVD, and miniDVD). Simple folders/files transferred on optical discs were unceremoniously copied to the Locked server. We transferred the content from discs directly to the server from a networked PC, running a virus scan (McAfee) before opening each disc. Since optical discs are generally "safe to read" with minimal chance of accidentally writing to the disc, we did not employ a write blocker. Each disc generally contained a small number of large files and we summarily verified complete transfer by comparing the properties on the disc to the properties of the data transferred to the server. Once all the content data for a given accession was ingested we created a manifest (directory index, file characterizations, check sums, etc.) at the accession level[36] using Karen's Directory Printer.

About six months after we completed processing the collection, we began experimenting with some additional technologies and discovered that imaging the playable DVDs would have been a good idea. The Digital Library Assistant found he could create a preservation-quality version of a TS_VIDEO file (comprised of VOB, BUP, and IFO files) by imaging the optical disc using IsoBuster. He then used a program called FFmpeg to convert the disc image to an MPG2 file. He wrote a script that prompted FFmpeg to stitch together the VOB files copied from the ISO image and convert this concatenated file to the MPG2 format. The disc image (ISO file) can serve as a preservation file for TS_VIDEOs and the MPG2 can serve as an uncompressed working copy. Although we did not image the digital videodiscs for the collection discussed in this study, we plan to incorporate that practice into our workflow going forward.

Hard disk drives. We received one hard disk drive containing nearly 1 TB of content that was created in the Macintosh environment. We connected it to a networked Macintosh computer via a write blocker. Unfortunately the write blocker completely blocked the transfer of data, so we took a risk, connected the drive directly to the computer, ran a virus scan, and transferred the data to the Locked server. Our local procedure for hard drives includes creating a pre-ingest manifest and a post-transfer manifest, following which we verify the integrity of the transfer

35. Files from one institution were so poorly titled that the interviewees were largely unidentifiable. Pending requested clarification from the creator, that set of interviews remains unprocessed and inaccessible.

36. Although we have successfully employed the Duke DataAccessioner to generate checksums and metadata for discs in other collections, creating disc-level metadata for huge audiovisual files proved too time-consuming.

using Beyond Compare. Months after the hard drive had been ingested we realized that, due to an oversight or a technical glitch, a pre-transfer manifest had never been generated. Once we detected the failure, we promptly created a baseline manifest of the data on the Locked server for future reference.

For all transferred data, the archival originals (as received by the archives) were kept intact in the Locked server and the data was copied onto the Workbench server space before any work was performed on them. We left the content of the original files untouched during ingest, with the exception that we did rename some discs/files as needed. Upon ingest we discovered that some discs had no names or that an institution had named all their discs exactly alike (such as the ever-popular title "Oral History"). In such cases the discs were assigned consecutive numbers upon ingest to minimize the potential for them to overwrite one another as the data was transferred to the server. We physically numbered the discs with an archival pen during accessioning.

During accessioning we arranged materials physically and intellectually by accession; accessions were not intermingled. If there was a discernible order to the discs (such as alphabetical by interviewee), we physically arranged the discs and then assigned sequential numbers to discs. This facilitated matching the content of the files on the server to the discs in the boxes. The sequential numbers were helpful in tracking disc errors/failures and also aided in disc retrieval when we created courtesy copies for interviewees or access copies for Interlibrary Loan.[37] Since there was very little printed material in the collection, it was quickly and easily processed (foldered, boxed, and labeled) during accessioning. We created an accession folder on the locked server with the following sub-folders:

- Originals (data as received by the archives)
- Access (files transformed and renamed for dissemination)
- Documentation (transfer forms, correspondence with creators, manifests, metadata, contracts, readme, etc.)

Transforming digital files for dissemination

We copied files to the Workbench server space for processing and performed the following actions on copies of the files (not on the received version). We renamed

37. We retained the original DVDs for use as duplication masters, which has proven the most efficient way to create universally accessible, high-quality DVDs as courtesy copies for interviewees and their families. We also fulfill requests for DVDs from educational institutions in regions with unreliable Internet service.

files according to local protocol prior to ingesting them into our digital repository.[38] We used Better File Rename to normalize file names by the batch—to add the accession number as a prefix, replace capital letters with lowercase, delete or replace spaces and symbols with underscores, etc.

	Received Version		Access Copy
Text	.doc .docx .xls .xlsx	⟹	.pdf
Audio	.cda .wav	⟹	.mp3
Moving Image	.mov VIDEO_TS (.vob, .ifo, .bup)	⟹	.mp4 (320 x 240 Web optimized)
Still Image	.png .tif .jpg	⟹	No access copies created

Table 1. File format transformations/conversions

We did not create access copies for still images.[39] We used Adobe Acrobat X Pro to convert transcripts (by the batch) and to convert one spreadsheet to PDF. We used Sony Sound Forge to convert CDA audio to WAV (for preservation) and to MP3 (for access) and embed descriptive metadata. After experimenting with StreamClip, Adobe Premier, and HandBrake, we opted to use HandBrake to transform large, complex moving image files to smaller, Web-optimized files.[40] HandBrake won us over with its ability to automatically transcode dozens of complex files from a queue without human intervention. Bonus features include its ease of use, its variety of outputs, and the fact that in the few videos we spot-checked, HandBrake did not appear to drop any frames. In HandBrake we created a template and set output rules (file naming conventions, file type, file size, etc.). We then selected a number of large video files and set HandBrake to automatically transcode all the files in the job and save them as MP4 files to a designated folder on the server.

38. Renaming files is not a scalable practice and was only performed to accommodate local system requirements.

39. The PNG images are "cover art" for AV files, which we did not use; the TIFF and JPEG files are images of artworks for which we will provide onsite access only (due to our conservative interpretation of rights).

40. The access copies we created for the *Pacific Standard Time: Art in L.A. 1945-1980* videos are inconsistent in size, reflecting our various stages in the discovery of and experimentation with different technologies and file output specifications.

Description, Discovery, and Delivery

Background

The primary deliverable for this resource is a digital collection in our local repository[41] that is publically accessible through The Getty's institutional website (getty.edu). In addition to creating access points within our local systems, we intentionally pushed descriptions of the collection to outside sources: MARC and EAD records point to the collection and series-level digital objects from WorldCat, ArchiveGrid, and the Online Archive of California. The digital collection is made up of digital objects that include the archival context of the electronic files as well as their content. Fourteen of the accessions are available online, with only three of the accessions having an electronic component that (due to rights restrictions) is accessible only from computers onsite at The Getty. Transcripts for two of the accessions were received in print and are currently only available onsite in the GRI Reading Room. (Local policy for analog material is to "digitize upon demand" as justified.) Two of the accessions are unavailable due to the complete absence of contracts. Please note that because one accession is discoverable and publically accessible through the creator's YouTube channel, we opted to point to the YouTube channel from the MARC and EAD records rather than providing access through our digital repository. Although we did not create access copies for this accession, we did ingest the data and complete all the other steps in the workflow to meet anticipated preservation requirements.

The records most frequently requested from Institutional Archives by external researchers are audiovisual recordings with art historical content (lectures, symposia, interviews, etc.). Although the majority of our holdings receive minimal processing, audiovisual recordings of this nature qualify for more granular arrangement and description under our local policy. We anticipate that providing direct online access to these potentially popular oral history interviews is freeing staff time that would otherwise be committed to reference inquiries, box retrieval and re-shelving, reading room preparation and supervision, interlibrary loan efforts, etc.

Because of the popularity of our AV holdings, the Institutional Archives has traditionally created item-level MARC records for individual oral histories and recordings of events produced by The Getty to maximize their discoverability; however, as the volume and pace of production of electronic records accelerates with no parallel growth in staffing, it is increasingly challenging for us to create these item-level records. As the first big wave of *Pacific Standard Time* interviews crested, it was obvious that meeting the expectation for timely access would be impossible if the

41. As of the writing of this paper, The Getty Research Institute is migrating from Ex Libris' digital asset management system, DigiTool, to Ex Libris' Rosetta. Resources represented in the screenshots of this study will appear differently in the new system.

collection was treated at the item level. The AIMS project findings confirmed our instinct that the level of processing should be determined according to local policies and procedures and that the "overall aims and objectives of this step remain unchanged by format."[42] With that in mind, we applied our customary "MPLP" strategy to these mixed-media files. We developed a processing model somewhere between minimal and highly-intensive processing to accomplish fairly granular discovery and access while economizing our staff's time. Again in line with the AIMS project findings we viewed the "process as a whole" knowing that "decisions made at the beginning of the process [would] have direct impact on later outcomes."[43]

Before determining a processing strategy we first considered the dissemination mechanism—how would we provide access to the electronic components of the resource? Naturally our existing delivery vehicle influenced our decisions because our arrangement and description had to be compatible with the system through which the resource would be served. Eliminating item-level delivery and working within our current technological environment, our best option was to create descriptive units and digital objects at the accession level and hierarchically nest the individual interviews within each unit/object. Generally the interviews produced by each cultural institution were thematically related and the content of each accession focused on a particular topic such as ceramics, film, women artists, Japanese American artists, etc. This thematic cohesion logically supported grouping the interviews by provenance/ accession. Within each accession it was most practical to arrange and describe the content based on intellectual entity—we viewed each interview or symposium as a single entity, regardless of how many discs or files were associated with that interview or event. The only hiccup was that because of the way our digital repository was configured, we had to create separate digital objects for sets of files that are available only onsite (accessible only from a Getty IP address) versus sets of files that are globally accessible, meaning that accessions/series with mixed access rights must have two digital objects, one for open files and one for restricted files.

Digital Object

Prior to creating digital objects we created a collection-level parent record under which we could bring together the related digital objects. The local configuration of our digital repository allowed only for a brief abstract at the parent level—very little space was allotted for contextual information. The configuration did not support true hierarchical structures, but simply provided a way to gather designated digital objects as an itemized/logical set under a single parent node. In the Brief View (Figure 2) of this collection-level parent record each digital object (accession) is listed alphabetically by title, according to the system's default sorting. We therefore began

42. AIMS Work Group, 41.

43. AIMS Work Group, 1.

each digital object title with its exhibition name to ensure that digital objects belonging to the same accession would index adjacently even if they had been packaged separately due to rights restrictions.

Once the files were transformed we created a single digital object for each accession by packaging the digital interview recordings and transcripts together at the accession level (with the exception that an accession with mixed rights may have two digital objects). Depending on the nature of the accession there may be only a single file in the digital object or there may be fifty or sixty files packaged together in a more complex digital object. We created a structure map (Excel spreadsheet) in which we assigned each file a place in the digital object's hierarchy and specified its file name, label, and format. We then converted the structure map to TXT and encoded it as UTF-8. We concurrently created a Dublin Core record wrapped in METS describing the context and content of the digital object. We then ran a command-line Perl script that mashed together the DC/METS record with the TXT

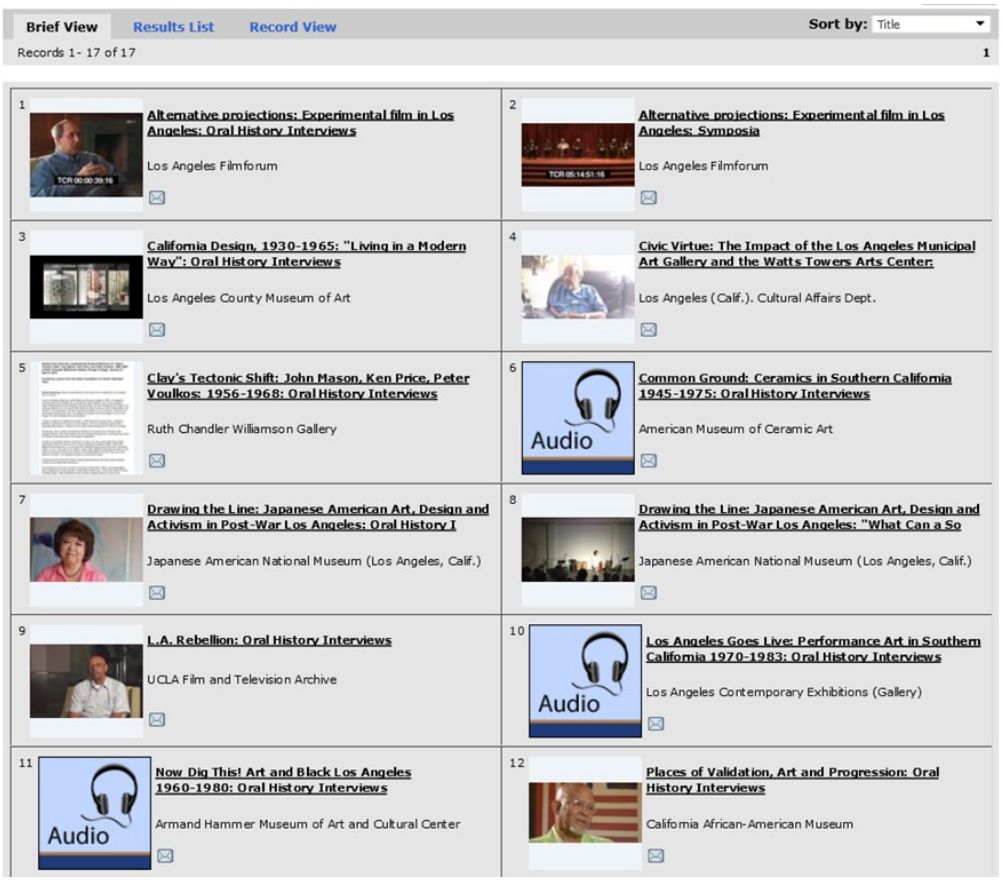


Figure 2. Screen capture of Brief View of digital collection: http://hdl.handle.net/10020/ia40011.

structure map to create a master METS record that includes the DC metadata and the nested hierarchy of the recordings and transcripts for each interview in the accession. See Appendix B for a sample METS record.

The outcome is a digital object (Figure 3) that provides access to the content and also describes the content within its archival context. Local policy requires that we include fields for title, creator, dates, abstract ("summary"), subject headings, conditions governing rights and access, and, in an effort to increase the discoverability of all the persons and corporate bodies that were oral history subjects and/or interviewees, we constructed standardized access points for each name according to LCNAF or ULAN authorities or AACR2 rules (preferenced in that order). Using the Handle System, we assigned persistent identifiers at the digital object level rather than at the file level.[44] Breadcrumbs at the bottom of the digital object "Record View" lead back to its parent (the digital collection), which acts as a gateway to all its siblings.

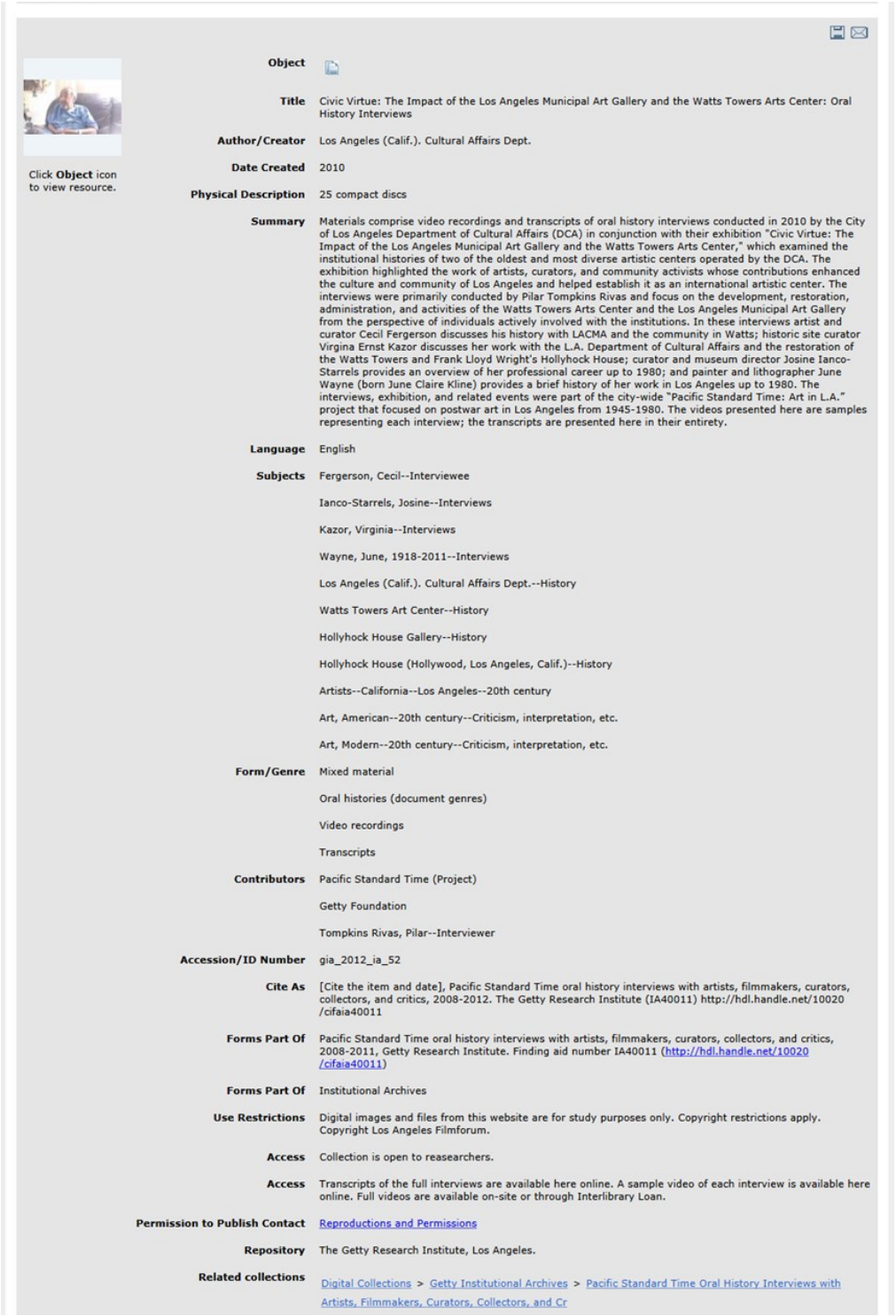
Clicking on the digital object opens the Object Viewer (Figure 4). Within the Object Viewer the components are listed in a navigation pane on the left. From there the user can select an item, which opens in a window on the right. The example below illustrates an object comprised of transcripts and audiovisual recordings, which are nested beneath each interviewee's name. Where a complete digital transcript is available for interviews that are several hours long, the local decision was to post only the content of the first TS_VIDEO or MOV file and make the remaining hours of video content available upon request.[45] The rationale behind the decision is two-fold: 1. Posting up to ten hours of video for a single interview uses a great deal of server space, and 2. Since the PDF of the transcript is complete and made keyword-searchable through OCR, the content of the interview is most easily discovered and studied via the transcript. Most researchers will be satisfied with a sample of the video, through which they can experience the interviewees' physical appearance, mannerisms, facial expressions, and voice. We will make the full video available for research that requires clarification of content or more detailed study (such as interviewees' intonation or countenance).

Finding Aid

The first (admittedly heretical) question we asked was: if the most granular/ detailed description of the interviews will be the DC/METS package in the digital repository and we create access points for each interviewee and subject in an accession-level MARC record, what purpose is served by the traditional finding aid?

44. Assigning identifiers at an aggregate level is supported by the findings of the AIMS Work Group, page 26.

45. This practice applies to traditional interviews in which the camera is focused exclusively on the interviewee. If the footage includes works of art, galleries, or other visual components, the entire video is posted online.



Click **Object** icon to view resource.

Object	
Title	Civic Virtue: The Impact of the Los Angeles Municipal Art Gallery and the Watts Towers Arts Center: Oral History Interviews
Author/Creator	Los Angeles (Calif.). Cultural Affairs Dept.
Date Created	2010
Physical Description	25 compact discs
Summary	Materials comprise video recordings and transcripts of oral history interviews conducted in 2010 by the City of Los Angeles Department of Cultural Affairs (DCA) in conjunction with their exhibition "Civic Virtue: The Impact of the Los Angeles Municipal Art Gallery and the Watts Towers Arts Center," which examined the institutional histories of two of the oldest and most diverse artistic centers operated by the DCA. The exhibition highlighted the work of artists, curators, and community activists whose contributions enhanced the culture and community of Los Angeles and helped establish it as an international artistic center. The interviews were primarily conducted by Pilar Tompkins Rivas and focus on the development, restoration, administration, and activities of the Watts Towers Arts Center and the Los Angeles Municipal Art Gallery from the perspective of individuals actively involved with the institutions. In these interviews artist and curator Cecil Fergerson discusses his history with LACMA and the community in Watts; historic site curator Virgina Ernst Kazor discusses her work with the L.A. Department of Cultural Affairs and the restoration of the Watts Towers and Frank Lloyd Wright's Hollyhock House; curator and museum director Josine Ianco-Starrels provides an overview of her professional career up to 1980; and painter and lithographer June Wayne (born June Claire Kline) provides a brief history of her work in Los Angeles up to 1980. The interviews, exhibition, and related events were part of the city-wide "Pacific Standard Time: Art in L.A." project that focused on postwar art in Los Angeles from 1945-1980. The videos presented here are samples representing each interview; the transcripts are presented here in their entirety.
Language	English
Subjects	Fergerson, Cecil--Interviewee Ianco-Starrels, Josine--Interviews Kazor, Virginia--Interviews Wayne, June, 1918-2011--Interviews Los Angeles (Calif.). Cultural Affairs Dept.--History Watts Towers Art Center--History Hollyhock House Gallery--History Hollyhock House (Hollywood, Los Angeles, Calif.)--History Artists--California--Los Angeles--20th century Art, American--20th century--Criticism, interpretation, etc. Art, Modern--20th century--Criticism, interpretation, etc.
Form/Genre	Mixed material Oral histories (document genres) Video recordings Transcripts
Contributors	Pacific Standard Time (Project) Getty Foundation Tompkins Rivas, Pilar--Interviewer
Accession/ID Number	gia_2012_ia_52
Cite As	[Cite the item and date], Pacific Standard Time oral history interviews with artists, filmmakers, curators, collectors, and critics, 2008-2012. The Getty Research Institute (IA40011) http://hdl.handle.net/10020/cifaia40011
Forms Part Of	Pacific Standard Time oral history interviews with artists, filmmakers, curators, collectors, and critics, 2008-2011, Getty Research Institute. Finding aid number IA40011 (http://hdl.handle.net/10020/cifaia40011)
Forms Part Of	Institutional Archives
Use Restrictions	Digital images and files from this website are for study purposes only. Copyright restrictions apply. Copyright Los Angeles Filmforum.
Access	Collection is open to reasearchers.
Access	Transcripts of the full interviews are available here online. A sample video of each interview is available here online. Full videos are available on-site or through Interlibrary Loan.
Permission to Publish Contact	Reproductions and Permissions
Repository	The Getty Research Institute, Los Angeles.
Related collections	Digital Collections > Getty Institutional Archives > Pacific Standard Time Oral History Interviews with Artists, Filmmakers, Curators, Collectors, and Cr

Figure 3. Screen capture of Record View of digital object: http://hdl.handle.net/10020/2012ia52

While we might well debate the value of a traditional finding aid in respect to a purely digital collection, in the case of a hybrid collection the finding aid is still our primary means of gathering and describing all the materials in the collection. The

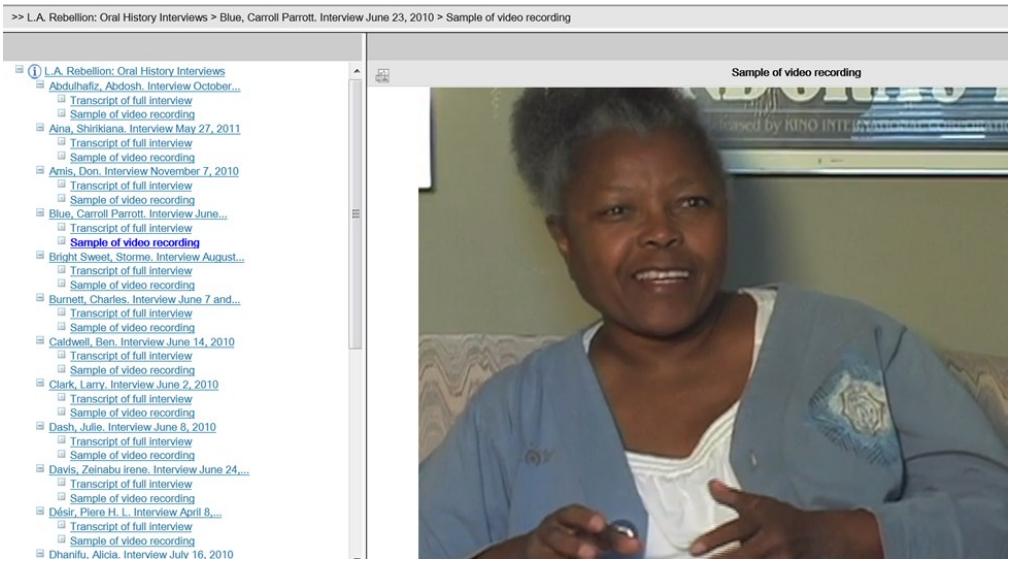


Figure 4. Screen capture of Object Viewer for digital object comprising complete transcripts and sample videos.

DC/METS records only describe the digital files that comprise the digital object to which they provide access. Materials that are available only in print and digital files with limited rights (closed or only available onsite)[46] are described in analytic MARC records. Although the analytic MARC records communicate the part-to-whole relationship between each accession and the parent collection, the finding aid still provides the most comprehensive view of the collection.

Following existing professional standards,[47] we created a finding aid in the Resource Module of the Archivists' Toolkit in which the digital and analog content of each accession is integrated intellectually and described in aggregate at the series level. Since the nature of the content was largely homogenous (interview recordings and transcripts), and very little of it was analog, integrating the description of the electronic and analog components was more logical than separating the materials by format.[48] Although we encoded each accession as a separate series in the EAD

46. Although we did not transform or arrange accessions that are completely closed, local practice includes describing restricted material that is part of an otherwise open collection.

47. During the processing of this collection we referred to *Describing Archives: A Content Standard* (DACS), (Society of American Archivists, Chicago: 2007), which allows for describing of materials at any level of specificity.

48. We have been encouraged to find that publications released since our processing of the collection have affirmed the option to intellectually integrate analog and digital records in a finding aid (i.e., Daines, 98).

schema, we did not label the series components "Series I," "Series II," etc. This enabled us to integrate each accession into the finding aid alphabetically by creator as we received them over the two-year period without having to renumber all the series with each addition. We did not write an administrative history for each institution, but wrote a brief scope and contents note that includes a link to each institution's own description of their participation in the *Pacific Standard Time* project (usually in the form of a press release issued with the opening of their exhibition).

The intellectual entities in the finding aid are listed in a manner that facilitates linking to their counterparts (digital objects) in the digital repository, with separate components for onsite versus public sets of material within a series with mixed rights.[49] We created a link to each digital object in the Resource Module of the Archivists' Toolkit, using the Digital Object dropdown in the Instance Type field. We're still in a quandary about how to title/label digital objects—some of the object names are cumbersome and result in annoying repetition in current displays, but abbreviating them results in incomplete information when objects are harvested and displayed outside their original context. In an effort to render the object most understandable to the users of our current technological environment, our standard data entry and tagging procedures have been subjugated to the limitations of our systems and style sheets. In an effort to identify audiovisual files that might be intentionally disassociated with their parent records in future systems, we decided to embed some very basic descriptive metadata into the files.

Embedded Metadata

Since our descriptive records (MARC, DC/METS, EAD) primarily describe the material at the series level, and the individual file names are not descriptive, we embedded basic descriptive metadata in the files themselves. Using Mp3tag we entered the collection title, creator, project name, and accession number to the entire accession as a batch and then cut-and-pasted the interviewee name and interview date from the existing structure map into the file's "Title" fields. This essential description of each interview travels with the file behind the scenes. Furthermore, if our delivery method changes in the future and the interviews are accessed through a media player, this file description will display to the user. We customized the labels of the Mp3tag fields to suit local needs: we changed Artist to Creator, Album to Collection, Comment to Subject, Disc Number to Accession No, etc. (Figure 5). Mp3tag is also extremely useful for global metadata changes and quality control because one can open the metadata for all the files in a set in a single window and apply edits to the entire set at once. Though not evident in Figure 5, Mp3tag also opens and displays any technical metadata that is inherent to the files.

49. To view the finding aid, including examples of contents lists for series/accessions with mixed access rights see http://hdl.handle.net/10020/cifaia40011.

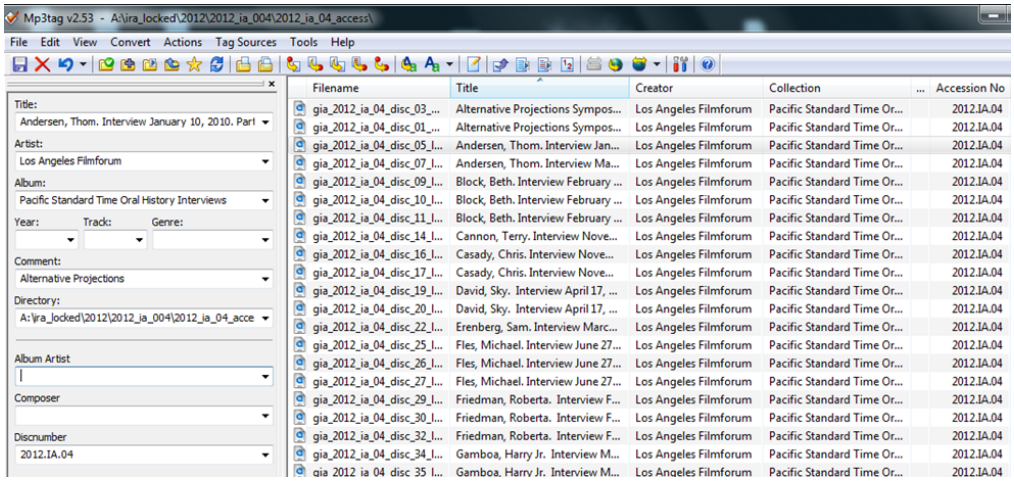


Figure 5. Screen capture of Mp3tag showing locally customized labels at the tops of columns

Bibliographic Records

We streamlined metadata creation by repurposing the descriptive metadata in the MARC, EAD, and DC records, but had no mechanism in place to automate a MARC to DC/METS conversion. We created a collection-level bibliographic record (MARC21) that largely parallels the digital object's DC/METS record. The bibliographic record points to the digital collection and also points to the finding aid for the more detailed collection description. We created analytic accession-level MARC records for each of the accessions that point to the finding aid and also point to their corresponding series-level digital objects (where they exist). As mentioned above, in an effort to expose interviewee names and subjects to search engines, we created access points for each interviewee name in the analytic MARC records, according to professional standards. The MARC records are very traditional and are available on WorldCat, ArchiveGrid, and our local website, getty.edu.

Conclusion

The J. Paul Getty Trust Institutional Archives has made notable progress over the past two years in the area of digital stewardship. While managing the *Pacific Standard Time: Art in L.A. 1945-1980* oral history interview materials, the Institutional Archives gained insight regarding the practicality of the theories that we encountered in professional literature; identified needs regarding our local infrastructure, technologies, and skills; and took initial steps to fill those gaps by formulating draft policies, procedures, and workflows. Of the perceived impediments to taking the first steps toward managing born-digital materials—lack of funding, time, and expertise—

the lack of time proved to be the most challenging for us. Before taking any action on the digital files we engaged in research and attended workshops, and while we have not yet achieved "expertise," we have certainly developed competencies in digital curation. Since we educated ourselves and employed technologies that were user-friendly, our lack of expertise proved inconsequential. With the exception of server space (which can be costly), we used free or inexpensive technologies to manage this collection, so our lack of funding was not an issue. Our biggest concern was a shortage of staff time, compelling us to adopt a moderate level of processing for this collection and inspiring us to continuously seek more extensible and less labor-intensive workflows. As we continue to adapt our strategies and workflows to fit the ever-changing world in which we operate, we are now able to balance the information we obtain through workshops and professional literature with the perspective gained while working with this hybrid collection. Upon reflection, of the lessons learned so far, the most helpful ideas were:

- All born-digital material is not created equal—it is prudent to determine the appropriate level of treatment for collections/records based on priorities and available resources.
- Increasing automation of actions is essential for streamlining processes to keep stride with the mass of inbound acquisitions/transfers.
- Fostering an atmosphere of cross-departmental collaboration and experimentation leads to innovation and process improvement.

The limited professional literature on born-digital curation that had been published by early 2011 was dominated by projects focusing on collections that received a high level of processing, including bit-for-bit disk imaging and the preservation and/or emulation of the creator's computing environment.[50] While the projects were extremely impressive, the level of processing performed on prominent collections isn't applicable in many situations: it isn't feasible for resource-constrained repositories, necessary for routine records, or scalable in repositories responsible for an exponentially increasing volume of electronic records. Although we tentatively identified actions that represented emerging best practices in digital stewardship from our 2011 research,[51] we did not strictly adhere to these practices during the processing of the *Pacific Standard Time* oral histories. Instead we proceeded under the assumption that, just as with analog records, not all digital

50. During our initial environmental scan we reviewed projects that were underway at Stanford University, Emory University, Harry Ransom Center, Bodleian Library, and others as noted in: The Getty Research Institute Born Digital Materials Working Group, "Task Force One Report on Best Practices and Sustainability" (2011), http://tinyurl.com/DigitalArchivesBestPractices1 (accessed December 2, 2013).

51. See The Getty Research Institute Born Digital Materials Working Group Task Force One Report for more detail.

records need "full-level" or "item-level" processing to render them viable and accessible into the future. In an age where backlogs of non-described archival collections/records are commonplace, the staggering volume of born-digital materials on our doorsteps demands that we take measures to accelerate and automate processes related to their stewardship.

While there are certain fundamental actions that should be taken to ensure the persistence of electronic files, the value and importance of a collection/records should determine the level of human and financial resources devoted to the material beyond those fundamental actions. We agree with the University of California's assertion that "'good-enough' processing can be quality processing,"[52] and, extending that maxim into the electronic environment, the Getty Institutional Archives has begun incorporating general guidelines for different levels of processing born-digital materials into our local archival processing manual. The workflow in this study is an example of a moderate level of treatment, reflecting thoughtful and deliberate decisions that were made to expedite processing, such as describing and delivering content in aggregate at the series level.

Describing and delivering the resources at the accession/series level saved time at nearly every step.[53] To begin with we saved time by creating one manifest for each of the 19 accessions rather than creating one for each of the 276 discs as we ingested them.[54] Rather than spending an average of 35 minutes creating an original MARC record for each of the 200-plus interviews/ symposia (which would have taken well over 120 hours), we spent an average of 65 minutes creating a detailed MARC record for each of the 19 accessions (which took about 21 hours). We also repurposed the metadata from the MARC record to create the DC/METS record.[55] Further study is needed to fully determine the implications of describing the interviews at the series/ accession-level (MARC, DC/METS) and ingesting them as an accession rather than as individual digital objects. One known consequence is that the individual interviews are not sortable or retrievable in the digital repository—the result is that researchers cannot directly or specifically target an individual recording, but must retrieve it as part of the series-level digital object. While the content is discoverable at the

52. University of California Next Generation Technical Services, "Guidelines for Efficient Archival Processing in the University of California Libraries" (University of California, September 18, 2011), 7-8, http://libraries.universityofcalifornia.edu/groups/files/hosc/docs/_Efficient_Archival_Processing_Guidelines_v3-1.pdf (accessed July 13, 2013).

53. Time saved during the creation of the finding aid is negligible because, although we described the content at the series level, we did take the time to list every interviewee by name.

54. We created individual manifests for only a handful of the 276 discs we ingested. We made no attempt to establish a norm for manifest creation time.

55. GRI Library Information Systems also wrote a very handy script to automatically convert a MARC record to METS for ingest into the digital repository. This script can only be applied when there is a one-to-one relationship between a MARC record and a digital object, which was not the case for all the materials in this collection.

interview level (within WorldCat, ArchiveGrid, the Online Archive of California, and our local systems), delivering the interviews in a "package" is less convenient (requires more mouse clicks) for the user than delivering them as individual digital objects. Additional study is also needed to determine how discoverable the individual interviews are from searches conducted outside the host systems (discovery directly through search engines such as Safari, Google, etc.). Nevertheless, unless other significant disadvantages to this approach become evident, it is likely that the model described in this study will become the norm rather than the exception in the Institutional Archives in order for us to keep pace with the rapid rate of incoming born-digital materials.

Throughout this study we referred to various inexpensive or free technologies that streamlined our procedures through automation or batch processing, such as Archivists' Toolkit to generate EAD, HandBrake to transcode video, Mp3tag to embed metadata, and Better File Rename to modify filenames. In addition to open-source and commercial tools, we also benefitted greatly from local tools that automate procedures and empower our archival staff to complete the many steps required to get a finding aid or digital collection across the finish line. In addition to the above-mentioned command-line Perl script that outputs a master METS record, GRI Library Information Systems (LIS) staff also wrote a command-line script enabling archival staff to clean up the EAD generated by Archivists' Toolkit in preparation for submission to the Online Archives of California and our local system. Additionally, they created a local tool (user-friendly GUI interface) that generates unique persistent identifiers for finding aids and digital objects upon demand. These in-house tools enable authorized archival staff to automatically complete the final steps associated with making collections accessible, streamlining the workflow by reducing the number of staff required to publish a collection and its description.

Given the technical nature of the workflow, our progress with this first hybrid collection was entirely dependent upon leadership that cultivated relationships between archivists and technologists and actively campaigned to build the infrastructure and fund the technology to support born-digital resource management. The Head of Institutional Records and Digital Stewardship created a productive working environment by facilitating communication between departments and encouraging staff to abandon perfection and espouse experimentation. Permission to fail was essential in ensuring uninhibited investigation and testing of the tools required to process and deliver the collection. Since our environment is steadily evolving—locally, professionally, and technologically[56]—a spirit of adventure and

56. At the time this case study was written, major changes include: local GRI systems upgrades/migrations (DigiTool to Rosetta, Voyager to Alma, and introduction of Primo search/discovery interface); professional standards are in transition (RDA is superseding AACR2, the second edition of DACS has been published, and EAD3 is forthcoming); born-digital best practices are still emerging; tools to help manage digital archives are still developing (ArchiveSpace is superseding Archivists' Toolkit; Archivematica is still in beta); and content carriers and file formats are constantly evolving. Our local changes are apparent in the links and screen captures in this study—the links are still valid, but the persistent identifiers now resolve to the new digital repository (Rosetta) as seen through the Primo interface.

exploration is necessary for success. As twenty-first century archivists we are (and ever will be) trying to hit a moving target—we must remain vigilant and resourceful.

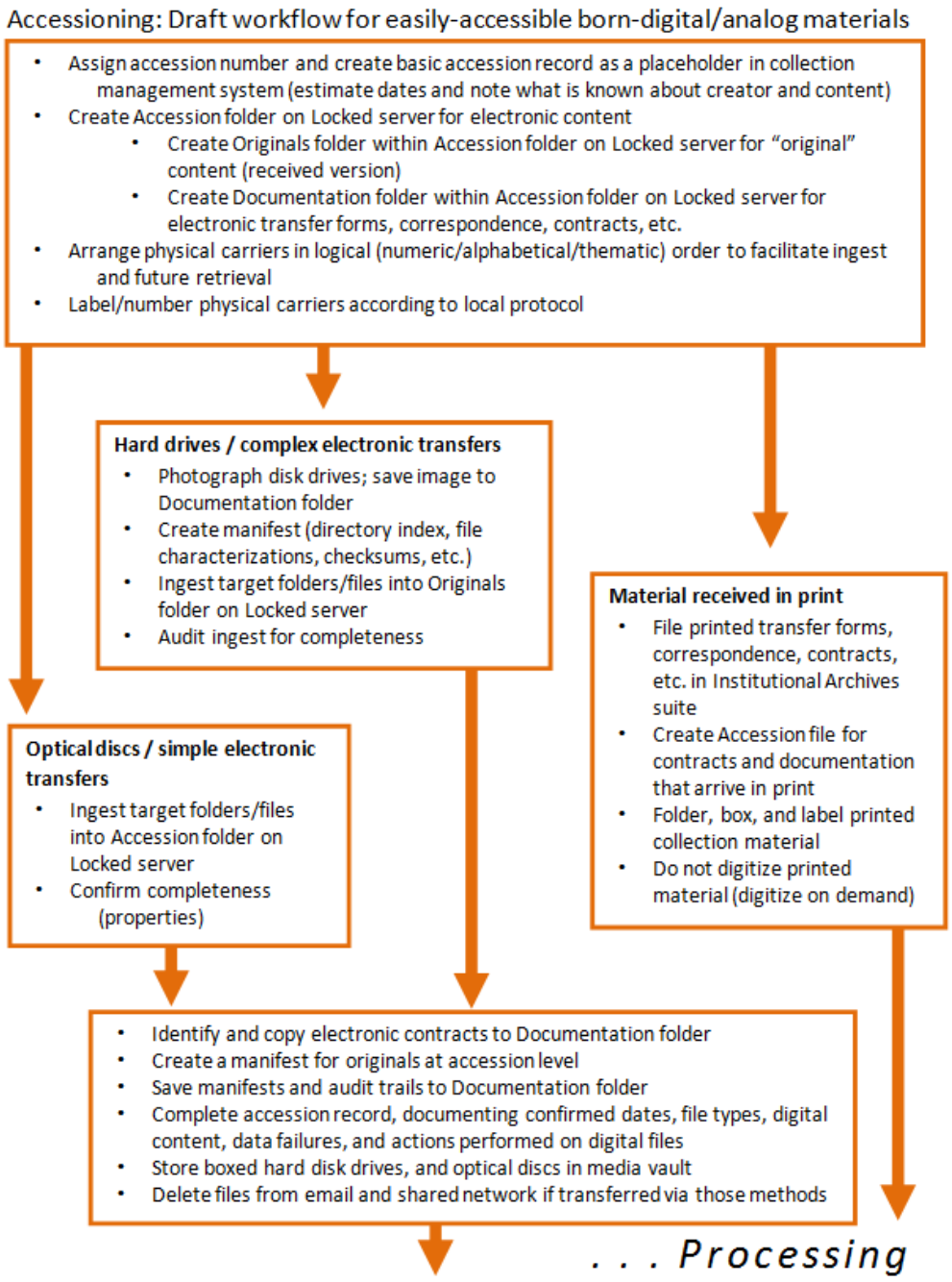
While the workflow presented in this study is a positive move away from item-level processing toward more aggregated description and delivery, it is not extensible for establishing control over the multiple terabytes of file directories that have since been transferred to the Archives. Presently, the Getty Institutional Archives is managing older, larger, more complex hybrid collections that make working with the *Pacific Standard Time: Art in L.A. 1945-1980* oral histories seem like a walk in the park; from our current vantage point we view the curation of this modest hybrid collection as a small victory, but a victory nonetheless. The management of the *Pacific Standard Time: Art in L.A. 1945-1980* oral histories was a solid first step toward the establishment of our department's digital infrastructure, the integration of digital accessioning and processing into our everyday workflow, the advancement of our electronic file management competencies, and the refinement of our strategies for responsibly managing unique born-digital resources.

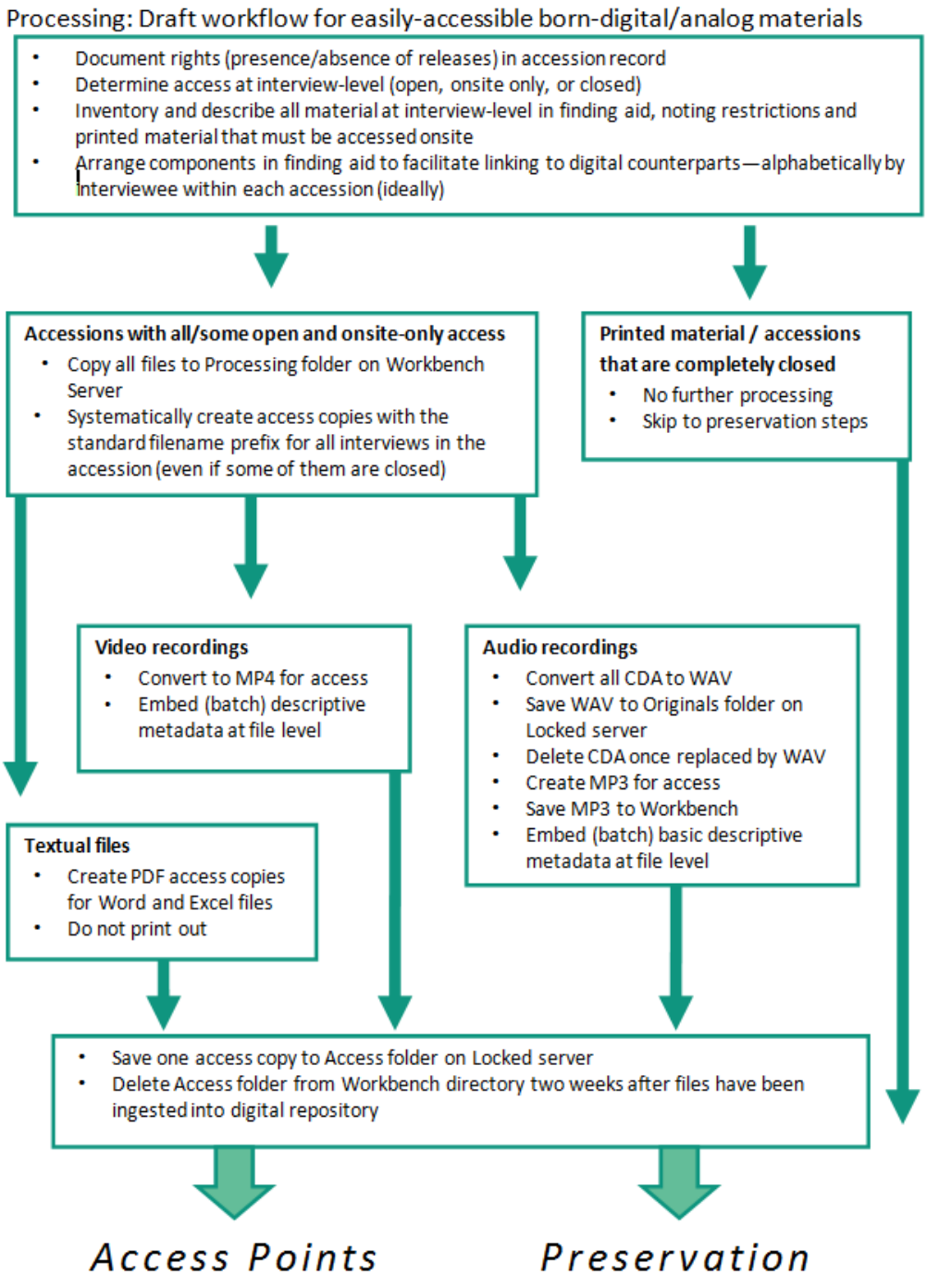
Acknowledgements

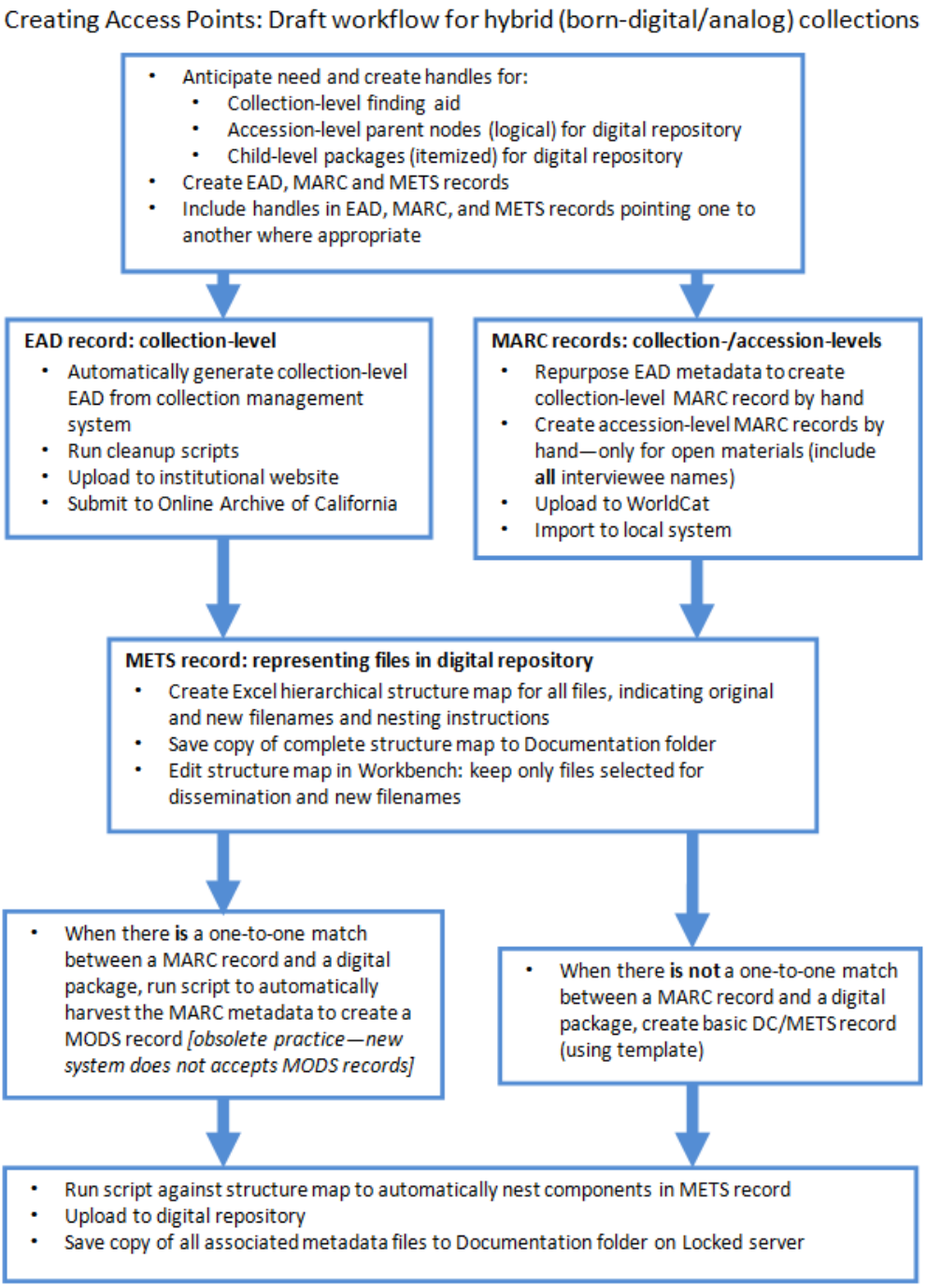
As is evident throughout this study, cross-departmental collaboration was absolutely essential in acquiring, managing, and disseminating the *Pacific Standard Time: Art in L.A.* oral histories. Management of this collection was accomplished through the joint efforts of staff from across the Getty, including:

- **Nancy Enneking**. Head of Institutional Records and Digital Stewardship, J. Paul Getty Trust Institutional Records and Archives. GRI Born-Digital Materials Steering Committee Co-Chair and Chair of Task Force on Ingest; archival administration and advocacy; customization of Archivists' Toolkit fields and analytics
- **Drew Krewer**. Digital Library Assistant, Getty Research Institute Digital Services. Digital repository data ingest, management, and preservation
- **Christina Lopez**. Senior Staff Assistant, Getty Foundation. Collecting and transferring materials; liaison with creators
- **Laney McGlohon**. Applications Systems Specialist, Getty Research Institute Library Information Systems. Getty Research Institute Born-Digital Materials Steering Committee Member at Large; systems, programming, and technical advisor
- **Lawrence Olliffe**. Applications Systems Analyst, Getty Research Institute Library Information Systems. Born-Digital Materials Working Group member of Task Force on Intake; programming and automation of critical actions
- **Cyndi Shein**. Assistant Archivist, J. Paul Getty Trust Institutional Records and Archives. GRI Born-Digital Materials Steering Committee member, Chair of Task Force on Best Practices and Sustainability, and member of Task Force on Metadata; data ingest; file management; audiovisual reformatting and file conversions; collection management, description, metadata, and access
- **Gene Tomilko**. Senior Systems Administrator, J. Paul Getty Trust Information Technology Services. Configuring workspace and secure server space for processing and storage of born-digital materials
- **Maureen Whalen**. Associate General Counsel, J. Paul Getty Trust General Counsel. Rights management advisor
- **Mary K. Woods**. Digital Library Assistant, Getty Research Institute Digital Services. Born-Digital Materials Working Group member of Task Force on Intake and Task Force on Metadata; digital repository data ingest, management, and preservation; and digital audiovisual advisor

Appendix A. Detailed draft workflow for easily accessible file formats on current media




Processing: Draft workflow for easily-accessible born-digital/analog materials
Document rights (presence/absence of releases) in accession record
Determine access at interview-level (open, onsite only, or closed)
Inventory and describe all material at interview-level in finding aid, noting restrictions and printed material that must be accessed onsite
Arrange components in finding aid to facilitate linking to digital counterparts—alphabetically by interviewee within each accession (ideally)
Accessions with all/some open and onsite-only access
Copy all files to Processing folder on Workbench Server
Systematically create access copies with the standard filename prefix for all interviews in the accession (even if some of them are closed)
Printed material / accessions that are completely closed
No further processing
Skip to preservation steps
Video recordings
Convert to MP4 for access
Embed (batch) descriptive metadata at file level
Audio recordings
Convert all CDA to WAV
Save WAV to Originals folder on Locked server
Delete CDA once replaced by WAV
Create MP3 for access
Save MP3 to Workbench
Embed (batch) basic descriptive metadata at file level
Textual files
Create PDF access copies for Word and Excel files
Do not print out
Save one access copy to Access folder on Locked server
Delete Access folder from Workbench directory two weeks after files have been ingested into digital repository
Access Points
Preservation


Creating Access Points: Draft workflow for hybrid (born-digital/analog) collections
Anticipate need and create handles for:
Collection-level finding aid
Accession-level parent nodes (logical) for digital repository
Child-level packages (itemized) for digital repository
Create EAD, MARC and METS records
Include handles in EAD, MARC, and METS records pointing one to another where appropriate
EAD record: collection-level
Automatically generate collection-level EAD from collection management system
Run cleanup scripts
Upload to institutional website
Submit to Online Archive of California
MARC records: collection-/accession-levels
Repurpose EAD metadata to create collection-level MARC record by hand
Create accession-level MARC records by hand—only for open materials (include all interviewee names)
Upload to WorldCat
Import to local system
METS record: representing files in digital repository
Create Excel hierarchical structure map for all files, indicating original and new filenames and nesting instructions
Save copy of complete structure map to Documentation folder
Edit structure map in Workbench: keep only files selected for dissemination and new filenames
When there is a one-to-one match between a MARC record and a digital package, run script to automatically harvest the MARC metadata to create a MODS record [obsolete practice—new system does not accepts MODS records]
When there is not a one-to-one match between a MARC record and a digital package, create basic DC/METS record (using template)
Run script against structure map to automatically nest components in METS record
Upload to digital repository
Save copy of all associated metadata files to Documentation folder on Locked server

Appendix B. Sample DC/METS record for digital object (http://hdl.handle.net/10020/2013ia18)


```
<?xml version="1.0" encoding="UTF-8"?>
<mets:mets xmlns:mets="http://www.loc.gov/METS/" xmlns:xlink="http://
www.w3.org/1999/xlink"
        xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
        xsi:schemaLocation="http://www.loc.gov/METS/ http://www.loc.gov/mets/
mets.xsd http://www.w3.org/1999/xlink   http://www.loc.gov/standards/xlink/
xlink.xsd"
        LABEL="Places of Validation, Art and Progression: Oral History Interviews"
        TYPE="Collection" PROFILE="http://www.loc.gov/standards/mets/
profiles/00000021.xml"
        xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:dcterms="http://purl.org/
dc/terms/">
        <mets:metsHdr>
                <mets:agent ROLE="CREATOR" TYPE="ORGANIZATION">
                    <mets:name>Getty Research Institute</mets:name>
                </mets:agent>
        </mets:metsHdr>
        <mets:dmdSec ID="Metadata">
                <mets:mdWrap MDTYPE="DC">
                    <mets:xmlData>
                        <record>
                            <dc:identifier>gia_2013_ia_18</dc:identifier>
                            <dc:identifier type="URI">http://
hdl.handle.net/10020/2013ia18</dc:identifier>
                            <dc:title>Places of Validation, Art and Progression: Oral
History Interviews</dc:title>
                            <dc:creator>California African-American Museum</
dc:creator>
                            <dc:contributor>Pacific Standard Time (Project)</
dc:contributor><dc:contributor>Getty Foundation</dc:contributor>
                            <dcterms:created>circa 2011</dcterms:created>
                            <dc:format>5 compact discs</dc:format>
                            <dc:type>Mixed material</dc:type>
                            <dc:type>Oral histories (document genres)</dc:type>
                            <dc:type>Video recordings</dc:type>
                            <dc:language>English</dc:language>
                            <dcterms:abstract>Resource comprises five video
recordings created by the California African American Museum (CAAM) in relation to
its 2011-2012 exhibition, "Places of Validation, Art and Progression.'" The exhibition
and related interviews were part of Pacific Standard Time: Art in L.A., a city-wide
research project that focused on the postwar (1945-1980) art scene in Los Angeles.
According to a press release from CAAM, “The perspective of 'Places of Validation,
```

Art and Progression' is on the history and driving forces that made venues and opportunities possible for Black art to be seen while allowing the art to reflect the wide variety of artists, styles, venues and personalities that served the Black arts scene between 1940-1980." The video recordings each run from three to six minutes in length and feature excerpts of interviews with artists and their supporters, as well as images of art works and ephemera. The interviews are undated, but were probably conducted in 2011.</dcterms:abstract>

<dcterms:abstract>Disc 1. Brockman Gallery: Dale and Alonzo Davis -- disc 3. First Voices of Validation: 1940s: Bill Pajaud, Greg Pitts, Betye Saar, Samella Lewis (early arts scene in Los Angeles); 1950s and 1960s: Dale Davis, Samella Lewis, Joe Sims (creating their own opportunities); 1960s and 1970s: Donald Stinson and Cecil Fergerson (progression); 1960s and 1970s: Greg Pitts and Cecil Fergerson (expanding recognition of Black art); Roderick Sykes and Dr. Samella Lewis (true validation is progression) -- disc 4. Murals: Betye Saar, Joe Sims, Elliott Pinkney, and Roderick Sykes -- disc 5. Other Places of Validation: Roderick Sykes, Greg Pitts, Samella Lewis, Donald Stinson, and Bernie Casey.</dcterms:abstract>

<dc:subject>Davis, Alonzo--Interviewee</dc:subject>

<dc:subject>Davis, Dale B., 1945---Interviewee</dc:subject>

<dc:subject>Stinson, Donald--Interviewee</dc:subject>

<dc:subject>Pajaud, William E., 1925---Interviewee</dc:subject>

<dc:subject>Lewis, Samella S.--Interviewee</dc:subject>

<dc:subject>Pitts, Greg--Interviewee</dc:subject>

<dc:subject>Saar, Betye--Interviewee</dc:subject>

<dc:subject>Sims, Joe--Interviewee</dc:subject>

<dc:subject>Fergerson, Cecil--Interviewee</dc:subject>

<dc:subject>Sykes, Roderick--Interviewee</dc:subject>

<dc:subject>Pinkney, Elliott--Interviewee</dc:subject>

<dc:subject>Brockman Gallery (Los Angeles, Calif.)--History</dc:subject>

<dc:subject>Artists and community--California--Los Angeles--20th century</dc:subject>

<dc:subject>Artists and patrons--California--Los Angeles--20th century</dc:subject>

<dc:subject>African American art--California--Los Angeles--20th century</dc:subject>

<dc:subject>African American artists--California--Los Angeles--20th century</dc:subject>

<dc:subject>Art, American--20th century--Criticism, interpretation, etc.</dc:subject>

<dc:subject>Art, Modern--20th century--Criticism, interpretation, etc.</dc:subject>

<dcterms:isPartOf>Institutional Archives</

dcterms:isPartOf>
<dcterms:isPartOf>Pacific Standard Time oral history interviews with artists, filmmakers, curators, collectors, and critics, 2008-2011, Getty Research Institute. Finding aid number IA40011 (http://hdl.handle.net/10020/cifaia40011)</dcterms:isPartOf>
<dc:rights>Digital images and files from this website are for study purposes only. Copyright restrictions apply. Copyright Los Angeles Filmforum.</dc:rights>
<dcterms:accessRights>The recordings exist solely in digital format (electronic files that were transferred to the Archives on optical discs). The discs act as duplication masters and are therefore restricted. Recordings from discs 1, 3, 4, and 5 are open to researchers and are publically available online. The content of disc 2 is restricted until such time as the J. Paul Getty Trust receives signed permissions from all the interviewees. All recordings that are open to researchers are available online.</dcterms:accessRights>
<dcterms:mediator>The Getty Research Institute, Los Angeles.</dcterms:mediator>
<dcterms:license>http://hdl.handle.net/10020/repro_perm</dcterms:license>
<dcterms:bibliographicCitation>[Cite the item and date], Pacific Standard Time oral history interviews with artists, filmmakers, curators, collectors, and critics, 2008-2012. The Getty Research Institute (IA40011) http://hdl.handle.net/10020/cifaia40011</dcterms:bibliographicCitation>
</record>
</mets:xmlData>
</mets:mdWrap>
</mets:dmdSec>
<mets:fileSec>
<mets:fileGrp ID="GID8" USE="reference">
<mets:file ID="FILE00002mp4" MIMETYPE="video/mp4" GROUPID="gia_2013_ia_18_disc_1" SEQ="1">
<mets:FLocat LOCTYPE="URL" xlink:href="file://gia_2013_ia_18_disc_1.mp4"/>
</mets:file>
<mets:file ID="FILE00003mp4" MIMETYPE="video/mp4" GROUPID="gia_2013_ia_18_disc_3" SEQ="2">
<mets:FLocat LOCTYPE="URL" xlink:href="file://gia_2013_ia_18_disc_3.mp4"/>
</mets:file>
<mets:file ID="FILE00004mp4" MIMETYPE="video/mp4" GROUPID="gia_2013_ia_18_disc_4" SEQ="3">
<mets:FLocat LOCTYPE="URL" xlink:href="file://gia_2013_ia_18_disc_4.mp4"/>
</mets:file>
<mets:file ID="FILE00005mp4" MIMETYPE="video/mp4"

GROUPID="gia_2013_ia_18_disc_5" SEQ="4">
<mets:FLocat LOCTYPE="URL" xlink:href="file://
gia_2013_ia_18_disc_5.mp4"/>
</mets:file>
</mets:fileGrp>
</mets:fileSec>
<mets:structMap LABEL="List">
<mets:div ORDER="1" DMDID="Metadata" LABEL="Places of Validation, Art and Progression: Oral History Interviews">
<mets:div ORDER="2" LABEL="Brockman Gallery: Dale and Alonzo Davis ">
<mets:fptr FILEID="FILE00002mp4"/>
</mets:div>
<mets:div ORDER="3" LABEL="First Voices of Validation: Pajaud, Pitts, Saar, Lewis, Dale Davis, Sims, Stinson, and Sykes ">
<mets:fptr FILEID="FILE00003mp4"/>
</mets:div>
<mets:div ORDER="4" LABEL="Murals: Saar, Sims, Pinkney, and Sykes">
<mets:fptr FILEID="FILE00004mp4"/>
</mets:div>
<mets:div ORDER="5" LABEL="Other Places of Validation: Sykes, Pitts, Lewis, Stinson, and Casey">
<mets:fptr FILEID="FILE00005mp4"/>
</mets:div>
</mets:div>
</mets:structMap>
</mets:mets>

Appendix C. Tools used to expedite processes mentioned in the case study

Be aware that The Getty does not officially endorse any of the following tools and that downloading any freeware is accompanied by risk (viruses, malware, etc.). Also note that these tools have uses beyond those noted below; only functionalities relevant to this study are listed.

Archivists' Toolkit http://www.archiviststoolkit.org/

Collection management system. Used to generate accession records and finding aids (PDF and EAD). Used as place of record for summary of actions performed on files and pointers to more detailed records regarding digital file management of each accession. (Archivists' Toolkit will soon be superseded by ArchivesSpace http://www.archivesspace.org/).

Better File Rename http://www.publicspace.net/windows/BetterFileRename/

User-friendly tool for modifying file names by the batch. Used for normalizing file names, adding accession numbers to file names, and eliminating unacceptable characters in file names.

Beyond Compare http://www.scootersoftware.com/

User-friendly file comparison tool. Used for side-by-side comparisons of file directories/disk images/checksums (comparisons before and after ingest) to verify integrity of transfers.

*FFmpeg http://www.ffmpeg.org/

Command line tool (requires writing a script). Used to convert a digital video disc image (ISO file) to an MPG2 file, stitch together the VOB files copied from the ISO image, and convert this concatenated file to the MPG2 format. MPG2 can serve as an uncompressed working copy.

HandBrake http://handbrake.fr/

User-friendly tool for transcoding video. Used to create access copies (MP4) from MOV and digital video discs (TS_VIDEO: VOB, BUP, IFO).

*IsoBuster http://www.isobuster.com/

Data recovery software. Used to create disc images from optical discs. Used to image digital video discs (TS_VIDEO: VOB, BUP, IFO). The resulting disc image (ISO file) can serve as the preservation file for videos that were transferred in non-preservation formats.

Karen's Directory Printer http://www.karenware.com/powertools/ptdirprn.asp

User-friendly manifests/file directories generator. Used to characterize files (including checksums, dates, file extensions, etc.) before and after files were ingested.

Mp3tag http://www.mp3tag.de/en/

User-friendly tool for batch entering and viewing existing metadata. Used to enter and edit embedded metadata into MP3 and MP4 files.

*Tools that Institutional Archives only recently tested and is currently integrating into our workflow are mentioned for their value in creating preservation video, which was a gap in the workflow of this project.