

**The Art of Access:
The Development of Collection Information Access :
at the Archives of American Art**

**Karen B. Weiss
Head of Digital Operations
June 18, 2016**



Title Slide-Slide 1

The Art of Access: The Development of Collection Information Access at the Archives of American Art

Good afternoon. Thank you Mr. Mizukami for extending a most welcome invitation to address your working group's study on collaboration between museums and research institutions to integrate and increase access to information on art and artists. I am honored to be speaking to such an esteemed group of archivists, librarians, curators and information scientists about these important issues. As one of 12 archives, 19 art, history and culture museums, 20 libraries, and 8 other research centers that make up the Smithsonian Institution, the Archives has experienced the challenges and opportunities involved in working in a complex structure. I know we have much to learn from each other.

Slide 2

Advancing Access Through the Decades

Dr. Kirwin's talk took you on a wonderful alphabetically guided tour of the Archives. My presentation will move chronologically, starting in 1954 with the Archives' mission to locate and facilitate unprecedented access to primary resources through microfilm. Along the way, each decade brought advances in access. Technology was of course the major factor, but I'll share some of the more subtle characteristics of archival access that played an important role in the development of our digitization methodology that today offers free, online access to collections digitized in their entirety. I'll close by describing our involvement today in initiatives that push the envelope of accessibility and collaboration using linked open data, crowdsourcing and new standards.

Slide 3

Excerpt from a 1970 oral history interview of Archives' founder Lawrence Fleischman

The Archives of American Art was founded in 1954 as a central repository of historical documentation made accessible on microfilm through interlibrary loan to support research in the newly emerging field of American art history. Founders Ted (Edgar) Richardson, art historian and director of the Detroit Institute of Arts, and art collector and businessman Lawrence A. Fleischman's innovative plan combined Richardson's frustration with the lack of access to historic documentation as he was writing his book, *Painting in America: The Story of 450 Years*, and Fleischman's personal connections to American artists as a collector, patron and friend.

In an excerpt of a 1970 interview in the Archives Oral History Program, Lawrence Fleischman recalled the circumstances that led to the founding of the Archives

I quote: Ted Richardson had difficulty publishing books. ...Also, he had to spend ten years gathering facts. He had to go to Boston to collect a fact, he had to go to Philadelphia to collect a fact, to Washington to collect a fact. Lloyd Goodrich used to tell me about the amount of time and trouble it took to gather information. As I was thinking about these things the idea came to me: ... Why isn't there a central place where records could be saved? This was about the time that Macbeth started talking about going out of business. I began thinking of all the valuable letters between dealers and artists that were being lost. One day ... [I] went to see Ted Richardson. I started talking to him about my idea. I remember his eyes lit up and he said, "This would be a great thing to do." I had the idea of doing it on photostats. He said 'Microfilm is such a wonderful tool—now with microfilm it would be possible to go, for instance, to Philadelphia.' In fact, Philadelphia was one of the first places we did sweep to get records on microfilm. We talked about this for two hours. I suggested that we call it "The Archives of American Art."

End of quote

In the Archives' is a letter from Richardson to Fleischman and his wife Barbara, dated August 22, 1954, in which he discusses his work on the final chapter of his book, and closes by posing a question: "Do you realize what a big thing you have done in starting the Archives? I know you do. But do you? It is enormous in its implications; enormous!"

Slide 4

Founding Mission: Collaboration with Repositories to Increase Access

And "sweep" Philadelphia they did. Collaborating repositories included the Pennsylvania Academy of the Fine Arts, the oldest art academy in the United States, founded in 1805; the Historical Society of Pennsylvania; the Philadelphia Museum of Art; the Philosophical Society, the Free Library of Philadelphia, the Library Company of Philadelphia, and others. The project yielded over 100 reels of microfilm and unprecedented access to the papers of artists, collectors, art historians, patrons, sitters, and records of art societies, art schools, and art publications, for example the records of the Artists' Fund Society and the Peale family and Sartain family papers.

Two 1950s-era catalog cards from this project on the right are typical of the type of cataloging treatment given to the microfilmed materials. Note that while the descriptive details are brief, the level of description was often at an item level. Each card had a reference to the microfilm reel and frame numbers to assist researchers in finding the documents on the microfilm reel.

Slide 5

1954 (through 1995) Microfilmed Papers Held in Private Hands

From the start, in addition to collaborating with other institutions, the Archives also sought out and microfilmed materials in private hands often with the agreement that the originals be returned to their owners after microfilming. The practice had its immediate benefits in preserving and making available the historical record despite the reluctance of the artists, the art dealers, or their heirs to part with the personal and tangible evidence of their legacies, and in

establishing long term relationships with lenders that in subsequent years often yielded donations to the Archives of the original documents. The catalog cards on the right show an example of the loaned papers of sculptor Paul Manship as described in 1959.

Over time, problems began to surface with this approach as well as loans from repositories that impacted the development of the Archives' access strategies and which drained resources required to provide access to donated collections. In the 1990s, the Archives stopped microfilming papers lent by individuals, and today when asked, we decline requests to digitize collections we do not own.

Slide 6

Early Acquisitions - Individual Items, Small Letter Collections

As word of the Archives' ambitious endeavor spread, it didn't take long for individuals in possession of manuscript materials to look to the Archives as a permanent home. Collections received before 1960 were mainly what are referred today as 'artificial' or 'assembled' artists' letter collections or single items. Notable examples include letters penned by Frederick Stuart Church, Winslow Homer, Eastman Johnson, Albert Bierstadt, John Trumbull and others. Pictured in the left slide is a May 3, 1847 letter from the painter Thomas Cole (1801-1848) to Asher Brown Durand (1796-1886), from the 1954 acquisition of the Charles Henry Hart Autograph Collection, which contains over 300 letters compiled by Hart written by dozens of artists.

These small collections, like the papers from other repositories, were copied on to microfilm. Their catalog card descriptions described each individual letter, sometimes in detail, and referenced microfilm reel and frame numbers, as shown in the catalog card pictured in the slide on the right.

Slide 7

1958-Oral History Program Launched

In 1958, the Archives conducted two oral history interviews, and could likely not imagine that it was embarking on what would become one of the most prolific and well established oral history programs in the United States. The model they followed was the program established in 1948 at Columbia University in New York City, credited with being the first to launch oral history as a historical method. The approach taken by the Archives was to supplement its microfilming and collecting activities with tape-recorded interviews conducted by trained art historians and make transcripts of the interviews accessible through interlibrary loan.

The earliest interviews were conducted on reel-to-reel tape, replaced by cassette tape in the 1970s, and Sony minidisc recorders in the 1990s. Secure digital memory card recorders have been in use since 2009. In 2010, the entire analog audio repository was digitally reformatted as part of a two year major preservation grant.

Slide 8

1960s - Moves Headquarters to New York City, The Pace Picks Up

In 1960, the Archives moved its headquarters to New York City to leverage the proximity to artists, galleries, historical repositories for microfilming and acquiring historical records on

American art, as well as establishing its now pivotal role offering free and accessible resources for graduate art history students, curators, collectors, dealers and the general public.

The “sweep” that started in Philadelphia now targeted art-related manuscript collections and artists’ files at the New York Public Library, the New York Historical Society, and the Whitney Museum of American Art, among others. Over 2,000 reels of microfilm were produced during this decade, including 700 between 1961 and 1966 through collaborations with dozens of libraries, museums and historical societies comprising well over 30,000 mainly American art auction and exhibition catalogs.

Slide 9

Collections Acquired Incrementally

Out of necessity, donations to the Archives were often collected incrementally. This practice is not uncommon even today at the Archives of American Art or in most manuscript repositories, but during the era when acquisitions were microfilmed upon receipt, this collecting method resulted in a steady build-up of multiple and distinct groups of microfilm reels for the same collection.

The increased complications due to the segmented sets of reels, sometimes in combination with additions that were not yet available on microfilm requiring use of the original papers, or with portions that had been received as loans, complicated and inhibited research and presented challenges to describe and manage these collections effectively. While impossible to limit collecting to complete collections, our policy today of digitizing collections only when they are complete, at least to the best of our ability, reflects the lessons learned from the earlier, disjointed microfilm practice.

Slide 10

1960s- The Pace (And Size) Picks Up

The decade of the 1960s also saw the growth of acquisitions corresponding to our modern definitions of personal papers and organizational records, characterized as being organically created or accumulated, comprised of diverse formats, typically a mix of biographical documents, correspondence, photographs, writings, diaries, sketches and sketchbooks, and printed materials that were markedly larger in size than those that trickled in during the earlier years.

Among the collections acquired during this active decade were the records of the Macbeth Gallery, the first art gallery devoted to the sale of American art, totaling 103 linear feet; and over 36 feet of papers of Walt Kuhn, a painter, whose collection also holds the records he kept as secretary of the Association of Painters and Sculptures, organizer of the seminal 1913 International Exhibition of Modern Art known as the Armory Show.

Slide 11

1960s-New Approaches to Description and Arrangement

With the influx of new acquisitions and the commitment to microfilming for access, the cataloging treatment began shifting to a more minimal level, where an overall description of the

collection and reel numbers replaced specific identification of items and microfilm frame numbers. Interestingly, the Archives was practicing an approach for archival arrangement and description that didn't find a name until forty years later when in 2007 the infamous "More Product, Less Product" approach was advocated in the journal of the Society of American Archivists that has empowered archivists in the United States to embrace techniques that accelerate the process from acquisition to access and particularly to erase backlogs of unprocessed collections.

Slide 12

1970s-The Archives Joins the Smithsonian Institution

In 1970, the Archives became part of the Smithsonian Institution, and moved its headquarters and microfilming operation to Washington, D.C. where it shared space with the library of the National Collection of Fine Arts (now the Smithsonian American Art Museum) and the National Portrait Gallery as shown in the right slide. The Archives' New York office continued as a microfilm research and collecting center.

By becoming part of the Smithsonian, founded in 1846 to uphold a mission to "increase and diffuse knowledge," the Archives was able to strengthen its funding base, its commitment to fostering open and free access policies in line with those of the Smithsonian museums, and its stature as a part of the esteemed national institution. Furthermore, as new library, archives and web and digital technologies emerged in the coming decades, support from the Smithsonian's central information technology division was key to the Archives' ability to increase access and collaboration.

Slide 13

1970s -Expansion of Regional Network of Research and Collecting Centers

The Archives expanded geographically in the 1970s and 1980s with the opening of three additional collecting and microfilm research centers in Boston, Massachusetts, Detroit, Michigan and San Francisco, California, and a number of geographically focused projects. As the volume of new acquisitions increased, and agreements to microfilm loaned collections carried on, a smaller percentage of collections were being made available on microfilm. At the same time, academic programs and popular interest in American art history were on the rise, resulting in a steady increase of in-person visits to the Archives' Washington, D.C. headquarters to use original papers, and with it the increase of staff serving on the reference desk and the development of access policies and reading room appointments.

Slide 14

1980s-Standards for Archival Description: USMARC Format and Content Standards

By the mid-1980s, technological developments in library and archival bibliographic systems and revisions to descriptive content standards to support the characteristics of archival description ushered in a new era marked by adherence to standards, sharing of resources, and collaboration at the Smithsonian and beyond.

A main driver for the innovation of the decade was the introduction in 1983 of the US MARC Archival, Manuscripts and Control format supporting archival description. The AMC format was designed to remedy the deficiencies in the library-based USMARC format developed in the 1960s and widely used by librarians throughout the United States to copy-catalog published books, periodicals and other materials but which failed miserably at description of unique, unpublished aggregations of historical records.

A second important development was the 1983 publication of *Archives, Personal Papers and Manuscripts* (APPM) that replaced the item-focused rules for description of manuscripts in the Anglo-American Cataloging Rules with one designed for the needs of archivists and special collections.

Slide 15-

1985 - Smithsonian Institution Bibliographic Information System

As accommodating as these new standards were for archival description, many archives in the 1980s were unable to participate due to the lag in development by the library systems industry to support them. The Smithsonian Institution's innovative solution to this challenge was a bold undertaking that sought to centralize bibliographic data from across the Smithsonian's libraries, archives and research centers. Launched in 1985 as the Smithsonian Institution Bibliographic System (SIBIS), the Archives enthusiastically participated where our records were integrated with those of the Smithsonian's libraries, among them the Smithsonian American Art Museum, the Hirshhorn Museum and Sculpture Garden, the National Portrait Gallery, the Cooper-Hewitt Museum of Design, and the Freer Gallery of Art, and databases such as the Inventory of American Painting before 1913 and the Art Exhibition Catalog Index, as well as the archival collections at various art, history and culture museums.

The technological advances allowed by the Archives' use of the shared system was accompanied by a shift in working with colleagues from across the Smithsonian, and a new era of sharing and collaboration, particularly in the process of developing shared vocabularies and authority files.

As a robust library system and powerful database, the Archives was also able to gain collection management functionality such as barcoding and reporting.

The system, renamed SIRIS for the Smithsonian Institution Research Information System in 1996, has been continually maintained and data migrated through several system upgrades by the central Smithsonian.

Slide 16

1980s-1990s-Integrating Information Within and Beyond the Smithsonian Institution

In 1988, with grant funding from the J. Paul Getty Trust and the Andrew Mellon Foundation, the Archives began a three-year effort to convert its card catalog to the USMARC format which gave the Archives the opportunity to identify, organize and merge our catalog card entries into a more cohesive, aggregated description that correlated with the new standards. Among the most notable developments coming out of this initiative was the contingency that our MARC records be contributed to the Research Library Information Network (RLIN), a national MARC

based catalog and research library support organization, which required that access points for names of creators and subjects adhere to the Library of Congress Name Authority File (LCNAF) and to rules for formulating names. Our records also needed to include at least one authorized topical subject heading. The creation of an artists' and other art related authority file structured in USMARC provided an enormous advantage in subsequent years in our ability to expand the reach of our metadata and to collaborate with other libraries and museums.

Shown on the right is a MARC authority record, using as an example Yasuo Kuniyoshi whose papers were donated to the Archives in 1963 and 2015.

Slide 17

1980s-1990s - Processing, Finding Aids and Encoded Archival Description (EAD)

For all its benefits, the MARC AMC format was never designed to accommodate archival finding aids, a tool that facilitates the discovery of information about archival collections' complex structure, arrangement, and contents to encompass the overall as well as the series and folder levels. Although routinely used in archives for decades, identified variously as calendars, inventories, or finding guides, the Archives of American Art's adoption of finding aids as an access tool came late due to its 40 year history using detailed microfilm reel labels and microfilm targets as its chief source of information. Through a concerted effort beginning in the late 1980s the Archives began preparing and in some cases, publishing finding aids for some of its most voluminous, complex, and significant collections, among them the records of Jacques Seligmann & Co., the National Arts Club, the Downtown Gallery, and the papers of Rockwell Kent. When a new XML standard known as Encoded Archival Description (EAD) was released by the Library of Congress in 1992, the Archives was eager to embrace it.

EAD was designed by and for archivists as a solution to electronically structuring and sharing finding aids. Following the pattern in the previous decade of systems lagging behind standards, it wasn't immediately adopted until open source and XML editing software that facilitated encoding and sharing were made available. By the end of the 1990s, with the software tools in place, the Archives converted its small number of finding aids to EAD. Along with the encoding was the introduction of revised processing guidelines and reconciling finding aid descriptions with those in the USMARC catalog that served to improve access. Descriptions pointing to specific boxes and folders simplified the difficult process of conducting research and improved the ability for reference staff to furnish researchers with specific boxes and folders upon request.

In 2011, the Archives began using Archivists' Toolkit, an open-source software in widespread use across the United States to facilitate authoring finding aids and exporting EAD XML files.

Slide 18

1996- Web Access to Catalog Records

1997- Archives' first website www.aaa.si.edu

The Archives of American Art's first website was launched in 1997 with support from the central Smithsonian's Office of Information Technology. This slide provides a glimpse through the Internet Archives' Wayback Machine of the home page circa 2001. In 1996, a year before our site launched, the Smithsonian's online SIRIS catalog was made available on the World Wide

Web. In 2001, working collaboratively with the SIRIS staff who managed the online catalog, the Archives made its first use of repurposing our catalog data for the website by exporting the approximately 6,000 names of artists and other names of each collection's main creators from our USMARC authority records, and presenting them in alphabetical HTML lists where users could browse or search the names. The systems were linked through programmatically exporting the names with the URL of the SIRIS record to take users from our website back to the online catalog for a more complete description.

Note: Wayback Machine, Internet Archive,
(<http://web.archive.org/web/20000816113900/http://artarchives.si.edu/>)

Slide 19

1999 –2004-Large-scale Digitization of Microfilm Sets the Stage

In 1999, the Archives' received a grant to process and microfilm one of its largest and most important collections, the records of the Downtown Gallery, comprising 109 linear feet. Given its size, the Archives' took the unusual step to outsource the microfilming to a contractor. Upon completion, the processing archivist replaced the box and folder encoding in the EAD finding aid with the microfilm reel and frame numbers. The microfilm vendor, having been eager to experiment with new technology to produce digital images of microfilm, also delivered to the Archives nearly 200,000 digital files on CDs along with the 167 reels of microfilm. Through the process of determining if, how, and when the digital files might be made available on its website, the Archives sought the help of a web programmer who ultimately recognized the structured XML's data for reels and frames mapped perfectly with the method of storing and naming the digital files by reel and frame number. By linking the directory file to the XML, the Archives had unwittingly discovered the basis for its subsequent digitization methodology where the archivists' finding aid serves as the only source of underlying metadata to the digital files.

Over the next five years, twenty reels of microfilm for two more collections totaling 21 linear feet were digitized in their entirety, and the finding aids' encoded to result in another 35,000 images linked to the respective box and folder titles.

Slide 20

2004-Describing Archives: A Content Standard (DACs) Published by the Society of American Archivists

In 2004, at the same time the Archives was ramping up its EAD XML finding aid output and recognizing its value in linking digital resources to folder descriptions, the Society of American Archivists published *Describing Archives: A Content Standard*, known as DACS. The new standard replaced *Archives, Personal Papers and Manuscripts* to better guide archivists in their use of Encoded Archival Description.

Slide 21

2005-2018-Terra Foundation for American Art Digitization Grant

In 2005, the Terra Foundation for American Art extended an invitation to the Archives to submit a grant proposal to increase access to its collections through digitization, indicating its intent to

support a major “what would it take to do it all” effort. The Foundation’s prescience in recognizing the alignment of digitization with the Archives’ pivotal role in the advancement of art scholarship fortuitously coincided with the Archives’ readiness to replace microfilm and confidence to undertake such an ambitious proposition.

The ensuing award provided funding for a six year effort, specifically digitization of over 100 archival collections in their entirety, comprising over 900 linear feet, by replacement of its microfilm production with a digital scanning operation; development of a redesigned website and robust technical infrastructure capable of supporting internal workflow, digital asset management, and publicly web accessible access to the digitized content. In 2011, the Foundation generously committed another seven years of funding.

Slide 22

Fully Digitized Collections

In 2005, the Terra Foundation for American Art extended an invitation to the Archives to submit a grant proposal to increase access to its collections through digitization, indicating its intent to support a major “what would it take to do it all” effort. The Foundation’s prescience in recognizing the alignment of digitization with the Archives’ pivotal role in the advancement of art scholarship fortuitously coincided with the Archives’ readiness to replace microfilm and confidence to undertake such an ambitious proposition.

The ensuing award provided funding for a six year effort, specifically digitization of over 100 archival collections in their entirety, comprising over 900 linear feet, by replacement of its microfilm production with a digital scanning operation; development of a redesigned website and robust technical infrastructure capable of supporting internal workflow, digital asset management, and publicly web accessible access to the digitized content. In 2011, the Foundation generously committed another seven years of funding.

Slide 23

Processing Archivists and EAD Finding Aids

The Archives’ approach takes advantage of the work already taking place by skilled archivists to make collections accessible – the well-organized, well described collection and the EAD XML finding aid. Our methodology is as much processing as it is digitization. All collections scheduled to be digitized in their entirety must first have an EAD finding aid.

Slide 24

EAD XML Finding Aids Uploaded to Archives’ Collection Information System

Customized software designed for the digitization workflow includes a suite of web-based processes all revolving around the finding aid XML formatted data that each archivist uploads into the Archives’ SQL database to supply all the metadata, structure, navigation cues, and scanning directories for each collection. The programming recognizes the XML encoding for hierarchical arrangement - from top level description, to series level description, leading to folder titles that serve as the linking mechanism to the digital files.

The slide here shows the Workflow for digitization starting with the upload of the finding aid to the database and ending with the deployment of the online collection to the website.

Slide 25

Digitization-Rapid Capture Techniques

One of the archival “More Product, Less Process” techniques that has made its way into our digitization approach is a simplified scanning process. The collection is digitized in the order in which the archivist arranged it and files saved to a directory structure matching that arrangement. Except for film formats, which are done in a second pass with a specialized lens, all contents are scanned sequentially even if the folder contains diverse formats and sizes. The files are named automatically by the scanning equipment’s hardware so that no time is wasted with conforming to a semantic file naming requirement. This straightforward approach is further supported by an investment in equipment that is both extremely fast and captures high quality color files. Capture One Cultural Heritage software and other automated workflow support an efficient post-processing workflow that ingests the digital files into the Archives’ database where they are associated to the appropriate box and folder container record.

Slide 26

EAD XML Provides Structure for Display

A web-based preview platform that mirrors the final product is used by the digitization specialist, the data asset manager, and the processing archivist for quality control. The EAD XML metadata in combination with the digital files allows the Archives to graphically present the collection’s overall arrangement at the series level.

Slide 27

Finding Aids Provide Content and Context

The finding aid’s metadata provides the description at overall, series, and folder levels, where its purpose in the online collection is recast as contextual and descriptive metadata for the digital files.

Slide 28

Digital Images Presented At the Folder Level

The link at the folder title level brings the user to the image viewer where the folder’s contents are displayed in the sequence in which the documents were arranged and scanned. The programming takes advantage of the EAD XML to provide navigation to the previous and next folder.

Slide 29

Digital Collection Via Finding Aid Supports a Virtual Reading Room

The More about the Collection tab is a user-friendly label to present the entire archival finding aid rendered in html, fully indexed, and available in PDF form for download.

Slide 30

2009- present Digital Asset Management

In 2009, after four years accruing digital files and storing them on network accessible storage, the Smithsonian launched an enterprise Digital Asset System. As with previous system implementations, the Archives' use of the DAMS required specialized programming and workflows to support the hierarchical and structural requirements of fully digitized archival collections.

In addition to the DAMS systems' critical role for preservation of our digital repository, it services as the central repository from which applications and services point to for delivery of assets to the web and other applications.

Slide 31

2016 Managing Collections and Collection Information

Today, the Archives of American Art holds nearly 5,000 collections totaling over 16,000 linear feet. Our oral history program now comprises nearly 2,300 interviews, most of them transcribed. Despite the technological advances of digitization and the professional developments in the archival profession, the challenges we face in making collection information available are at their core very similar from those in previous decades. Modern collections are still expanding in size, and comprised now of born digital media and at risk audio-visual formats, requiring new solutions for access and stewardship. Our users' expectations for online access and open data are ever-growing.

Slide 32

Metadata Creation

The Archives of American Art creates and manages its collection information and management through a dedicated team of professional staff and their use of various systems and software. The Archives' metadata strategy has focused on the goal to integrate as many of the systems as possible in order to repurpose data for public consumption on the website as well as support internal tracking of collections, workflows for collection management, and digitization.

Our primary systems that all play a vital role are our own internal Collection Information System, the SIRIS MARC catalog, Archivists' Toolkit for our EAD finding aids, the Smithsonian Institution Digital Asset Management System, or DAMS, and Aeon, a system used for patron management and to support reading room and reproduction requests.

Slide 33

AAA CIS Delivers Collection Information to www.aaa.si.edu/collections

The Archives' internal web-based Collection Information System serves a critical purpose as the central source of collection information delivered to our website. Our earliest website integration with the SIRIS MARC catalog in 2001, where information was transformed as static HTML pages derived from our authority files, was overhauled in 2005 as a nightly import of collection and authority records into the collection database where metadata is integrated with EAD XML finding aid data uploaded by the archivist and with metadata associated with individually described and digitized documents entered directly into the system.

The presentation for the collection information is presented in the Research Collections section of the website, offering an alphabetical browse through our collections and oral histories, and through keyword access. Plans for a new website later this year will tap into additional structured metadata from our collection records to offer a faceted search and browse.

Slide 34

Oral History Online

Oral history interviews are each described individually, and leads users to an online transcript when available. The digital reformatting project completed in 2012 has made it possible to present audio excerpts and to plan for making full audio available in the future.

Slide 35

Crowdsourcing, Collaborations, and Crosswalks

The development of our information environment has been decades in the making, but as we reach one milestone, new opportunities and more challenges arise. Recently, the activity centers around cross-walking and sharing data to build exciting platforms and collaborations.

Slide 36

Smithsonian Collection Search Center

<http://collections.si.edu>

One such effort, starting in 2007 but continually being developed and enhanced today, is the Smithsonian's Collection Search Center, a major pan-Institutional initiative designed to bring together in one search platform collection information from over 50 data sources from the SIRIS MARC catalog, art, history and culture museums, and datasets from across the Smithsonian. The site leverages a Smithsonian-developed metadata index and web services that provides the Institution with a unified view across these multiple archives, libraries and museums. The Smithsonian developed metadata model incorporates elements and best practices from established standards such as Dublin Core, MARC, and METS.

The platform has not only brought together data, but has sparked new ideas about sharing data and nurturing a collaborative and access driven culture across the entire Institution.

Slide 37

Crowdsourcing Transcription

<http://transcription.si.edu>

The Smithsonian Transcription Center has been one of the most ambitious and transformative access initiatives undertaken by the Smithsonian in the past decade, relying on crowdsourcing to help the Smithsonian's archives, museums and libraries transcribe documents and collection information to result in expanding the repository of searchable text. For the Archives, the overwhelming amount of handwritten documents have long prevented taking advantage of optical character recognition software, and the daunting task involved in transcription was never entertained for more than a fraction of our collections. By collaborating with the crowd, the Smithsonian has been able to provide engaging opportunities for anyone with Internet

access to transcribe and review transcriptions of diaries, letters, writings, accession records, specimen sheets and other types of materials.

Slide 38

Expanding Access- The Smithsonian Online Virtual Archive (SOVA) and

Another example of cross walking data to expand access is the Smithsonian Online Virtual Archive, or SOVA, launched in 2015, that brings together the EAD XML finding aids from fifteen different repositories. When tallied, the holdings of the Smithsonian Institution's various archives represent one of the largest aggregations of archival collections in the world, totaling more than 150,000 cubic feet.

Slide 39

Digital Public Library of America (DPLA)

(<http://dp.la>)

Another advancement in access comes from the Smithsonian's partnership in the Digital Public Library of America, where the data available via the SI Collection Search Center is repurposed for contributing to this portal that specializes in open access and welcomes application developers to harvest the data for whatever purpose they desire.

With planning dating as far back as the 1990s, the development was led by the Library of Congress, HathiTrust, the Internet Archive and others to make accessible to anyone with Internet access digitized books, images, historic records, and audiovisual materials. The DPLA also assists cultural heritage institutions, particularly those with limited resources such as public libraries and smaller historical societies, in their goals to digitize collections to ensure a wider and more diverse representation of the society's digitized cultural heritage.

Slide 40

Encoded Archival Context (EAC) and Social Networks in Archival Context (SNAC)

For the past decade archivists and technologists have been engaged in the development and application of a new standard for encoding and sharing records describing corporate bodies, persons and families using an XML schema. Named Encoded Archival Context – Corporate Bodies, Persons and Families, or EAC-CPF, the schema is designed as a companion to Encoded Archival Description (EAD).

The schema is the backbone of SNAC, which stands for Social Networks in Archival Context, a platform to support discovery of these entities as found in archival collections from across the United States. SNAC uses data harvested from EAD finding aids, particularly the finding aid's biographical note and access points for creators, subjects, places, occupations and events, as well as other sources such as OCLC's WorldCat catalog.

The Smithsonian Institution is a partner along with the National Archives, the Library of Congress, and many other archives, working with the developers to see the SNAC platform become an invaluable resource for research.

Slide 41

American Art Collaborative – Publishing Linked Open Data

For the past year, the Archives has been gaining experience and understanding of Linked Open Data through its partnership in the American Art Collaborative. The Collaborative is creating a diverse critical mass of Linked Open Data on the Web on the subject of American art by putting the collections of the 13 participating museums and the Archives of American Art in the cloud and tagging this data as LOD. The goals are to use technologies designed for the semantic web to enhance the accessing, linking and sharing of information about American art in a way that transcends what is currently possible with structured data. The project has been funded by the Institute of Library and Museum Services as well as the Mellon Foundation, and encompasses training in Linked Open Data, conversion of data, the development of a browse application to test the potential for online discovery.

Among the 14 partners are the Smithsonian American Art Museum and the National Portrait Gallery.

Slide 42

Making Connections Using Linked Open Data Sets

As the only archival repository among 13 museum partners, participation in the American Art Collaborative has been an interesting opportunity to understand the mapping of the data first into the CIDOC Conceptual Reference Model, chosen as the preferred ontology for the Collaborative. The Conceptual Reference Model (CRM) has been developed by the Documentation Committee of the International Council of Museums (CIDOC) as an ontology for art and culture.

Particularly interesting will be the outcome of linking the Archives of American Art's entities for names of artists and institutions with museum data from the American Art Collaborative's partners, and moreover, beyond to the thousands of other data sets published as Linked Open Data such as the New York Times, DBPedia, and the Getty Vocabularies.

Slide 43

"Anytime, Anywhere, Archives"

Thank you for your attention during the tour through the Archives' past seven decades of refining the art of access. The Archives of American Art will continue to grow, learn, and collaborate to fit its goal to become the "Anytime, Anywhere, Archives."

Brief biography:

Karen Weiss is the Head of Digital Operations at the Smithsonian's Archives of American Art, where she oversees implementation of the Archives' strategic goals to digitize collections and make them accessible to the public. She has overall responsibility for the Archives' collection information systems, website, and stewardship of digital assets. Since 2005, Ms. Weiss has directed a major grant from the Terra Foundation for American Art involving the work of

archivists, information professionals, imaging specialists, and web developers. Her article, "Collections Online: An Archival Approach to Digitization and Web Accessibility in the Archives of American Art" appeared in *Collections: A Journal for Museum and Archives Professionals* (Summer 2010). In 2013 she shared the Smithsonian Institution's Secretary's Award for Digital Enterprise. Ms. Weiss holds a Masters in Library and Information Science from the University of Maryland, College Park.