

# EAD (Entire Archives Digitized): Development of a Working Model for Large Scale Digitization

SAA Research Forum,  
*Foundations and  
Innovations*

August 11, 2009

[weissk@si.edu](mailto:weissk@si.edu)

[www.aaa.si.edu](http://www.aaa.si.edu)



Smithsonian  
Archives of American Art

Home About Us Research Collections Exhibitions Publications

[Research Collections](#) >

### Collections Online

Primary Sources on American Art and Artists

Recently Added All Collections

- [Caffin, Charles Henry](#)
- [Blumenschein, Ernest Leonard](#)
- [Soyer, Raphael](#)
- [Stillman, James](#)
- [Mills Family](#)
- [Rossiter, Thomas Prichard](#)
- [Dreier, Dorothea A. \(Dorothea Adelheid\)](#)
- [Page, William](#)
- [Pollock, Jackson](#)

[See More Collections](#)

The screenshot displays a grid of nine small images representing various collections. These include a portrait of a man, a Native American figure, a person in a doorway, a landscape drawing, a handwritten letter, a portrait of a woman, a portrait of a man, a portrait of a woman, and a portrait of a man.

## Collections Online: Scaling Up Digitization of Special Collections

- No item level access
- Database driven; uses EAD finding aid data as sole source of metadata
- Flexible interface design
- Primarily grayscale, 300 dpi
- Easy to operate equipment
- Networked digital file storage and mgmt
- Automated, web-based workflow, notifications
- Effective and efficient integration of existing descriptive practices



The screenshot displays the Smithsonian Archives of American Art website. The header features the Smithsonian logo and the text "Smithsonian Archives of American Art". Below the header is a navigation bar with links: Home, About Us, Research Collections, Exhibitions, and Publications. The main content area is titled "Collections Online" with the subtitle "Primary Sources on American Art and Artists". There are two tabs: "Recently Added" and "All Collections". Under "Recently Added", a list of names is shown with links: Caffin, Charles Henry; Blumenschein, Ernest Leonard; Soyer, Raphael; Stillman, James; Mills Family; Rossiter, Thomas Prichard; Dreier, Dorothea A. (Dorothea Adelheid); Page, William; and Pollock, Jackson. To the right of the list is a grid of nine small images, including portraits and documents. At the bottom, there is a link "See More Collections".

Smithsonian  
Archives of American Art

Home About Us Research Collections Exhibitions Publications

[Research Collections](#) >

### Collections Online

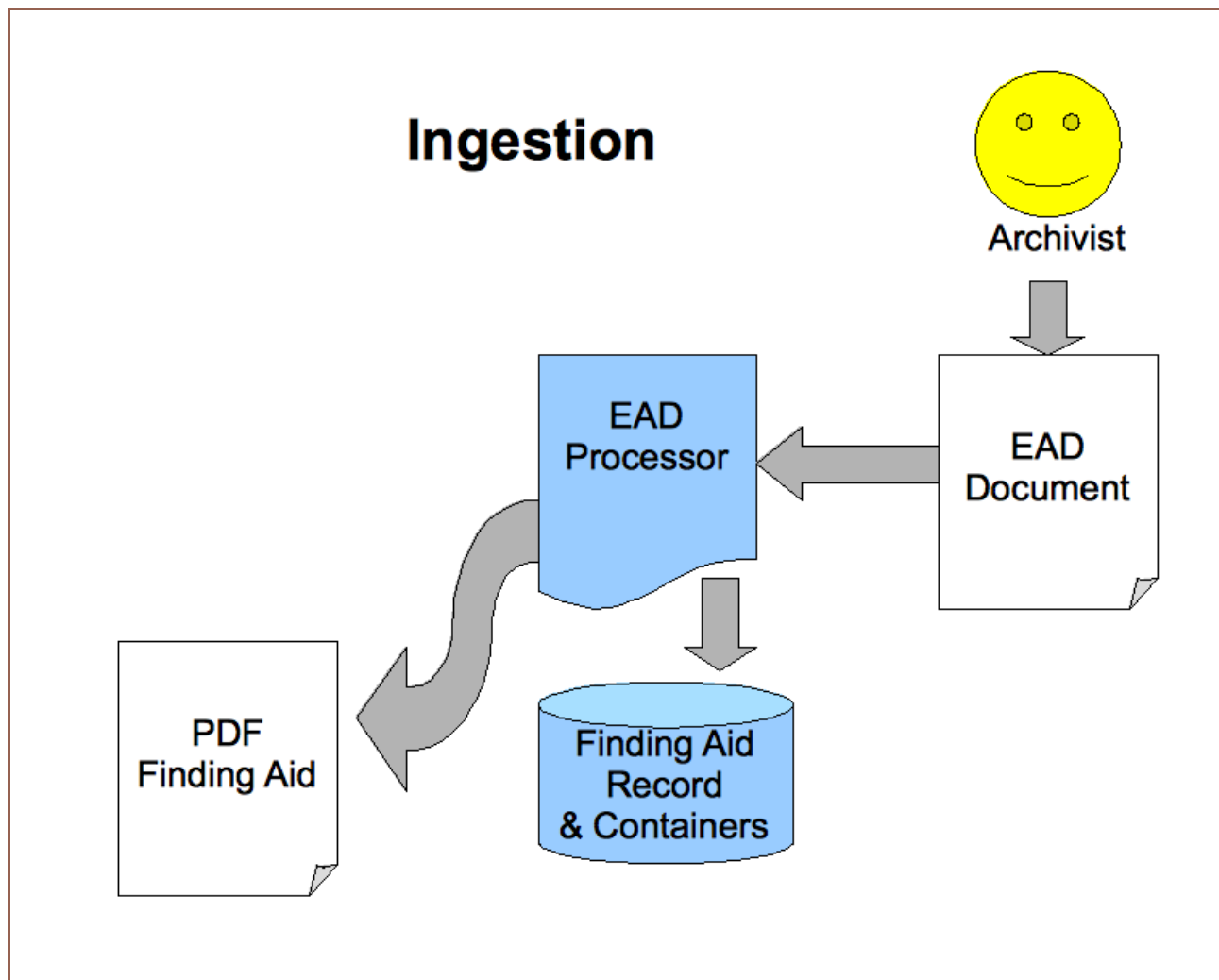
*Primary Sources on American Art and Artists*

Recently Added All Collections

- [Caffin, Charles Henry](#)
- [Blumenschein, Ernest Leonard](#)
- [Soyer, Raphael](#)
- [Stillman, James](#)
- [Mills Family](#)
- [Rossiter, Thomas Prichard](#)
- [Dreier, Dorothea A. \(Dorothea Adelheid\)](#)
- [Page, William](#)
- [Pollock, Jackson](#)

[See More Collections](#)

# Overview: The EAD “Engine”



# Database Tables

tblFindingAid		tblFindingAidComponent			tblContainer		
FindingAidID	int identity	PK	FindingAidComponentID	int identity	PK	ContainerID	int identity
FindingAidProperTitle	varchar(255)		FindingAidComponentType	nvarchar(50)	11	ContainerSeries	int
FindingAidSubTitle	varchar(255)		FindingAidComponentUnitId	nvarchar(50)		ContainerSeriesReference	varchar(100)
FindingAidPublisher	varchar(255)		FindingAidComponentUnitTitle	ntext		ContainerBoxLabel	varchar(100)
FindingAidAuthor	varchar(255)		FindingAidComponentUnitDate	nvarchar(255)		ContainerBoxNumber	int
FindingAidCreationStatement	ntext		FindingAidComponentPhysDesc	ntext		ContainerBoxType	varchar(25)
FindingAidCreationDate	varchar(255)		FindingAidComponentScopeContents	ntext		ContainerFolderLabel	varchar(100)
FindingAidRepository	varchar(255)		FindingAidComponentSeriesParent	varchar(50)		ContainerFolderNumber	int
FindingAidCreator	varchar(255)		FindingAidComponentOrder	varchar(100)		ContainerTitle	nvarchar(500)
FindingAidTitle	varchar(255)		fkFindingAidID	int		ContainerDisplayDate	varchar(100)
FindingAidDates	varchar(255)		DateRecordCreated	datetime		ContainerPhysicalDescription	ntext
FindingAidExtent	varchar(255)		DateEdited	datetime		ContainerSummary	ntext
FindingAidAbstract	ntext		FindingAidComponentBoxLabel	varchar(100)		fkFindingAidID	int
FindingAidBiography	ntext		FindingAidComponentReelLabel	varchar(100)		fkCollectionID	int
FindingAidScopeContent	ntext		FindingAidComponentFrameLabel	varchar(100)		DateRecordCreated	datetime
FindingAidOrganization	ntext		FindingAidComponentFolderLabel	varchar(100)		DateEdited	datetime
FindingAidAcquisitionInformation	ntext		FindingAidComponentHierarchyLevel	int		ContainerReelNumber	int
FindingAidAccessRestrictions	ntext		FindingAidComponentArrangement	ntext		ContainerReelLabel	nvarchar(100)
FindingAidUseRestrictions	ntext		fkCollectionID	int		ContainerFrameNumber	int
FindingAidPreferredCitation	ntext		FindingAidComponentParent	varchar(50)		ContainerOrder	int
FindingAidProcessingInformation	ntext		FindingAidComponentItemLabel	varchar(100)		ContainerFrameLabel	nvarchar(100)
FindingAidXMLURL	varchar(255)					ContainerComponent	varchar(50)
FindingAidURL	varchar(255)					ContainerParent	int
fkCollectionID	int					ContainerItemLabel	varchar(100)
SIRISBibNumber	int					ContainerItemNumber	int
DateRecordCreated	datetime					ContainerHeadingType	varchar(100)
DateEdited	datetime						
FindingAidPublicationDate	varchar(100)						
FindingAidOtherDescriptiveData	ntext						
FindingAidCreatorRole	varchar(20)						
FindingAidLanguageMaterial	varchar(255)						
FindingAidLanguageUsage	varchar(255)						
FindingAidRelatedMaterial	ntext						
FindingAidSeparatedMaterial	ntext						
FindingAidOtherFindingAid	ntext						
FindingAidIndex	ntext						
FindingAidBibliography	ntext						
FindingAidAccruals	ntext						
FindingAidPublicationAddress	ntext						
FindingAidArrangement	ntext						

tblSeries		
PK	SeriesID	int identity
14	SeriesParent	int
	SeriesLevel	varchar(100)
	SeriesReference	varchar(100)
	SeriesTitle	varchar(255)
	SeriesDisplayDate	varchar(100)
	SeriesPhysicalDescription	ntext
	SeriesSummary	ntext
	13	fkFindingAidID
fkCollectionID		int
11	DateRecordCreated	datetime
	DateEdited	datetime

# Tbls Container and Image

tblContainer		
PK	<u>ContainerID</u>	int identity
11	ContainerSeries	int
	ContainerSeriesReference	varchar(100)
	ContainerBoxLabel	varchar(100)
	ContainerBoxNumber	int
	ContainerBoxType	varchar(25)
	ContainerFolderLabel	varchar(100)
	ContainerFolderNumber	int
	ContainerTitle	nvarchar(500)
	ContainerDisplayDate	varchar(100)
	ContainerPhysicalDescription	ntext
	ContainerSummary	ntext
	fkFindingAidID	int
	fkCollectionID	int
	DateRecordCreated	datetime
	DateEdited	datetime
	ContainerReelNumber	int
	ContainerReelLabel	nvarchar(100)
	ContainerFrameNumber	int
	ContainerOrder	int
	ContainerFrameLabel	nvarchar(100)
	ContainerComponent	varchar(50)
	ContainerParent	int
	ContainerItemLabel	varchar(100)
	ContainerItemNumber	int
	ContainerHeadingType	varchar(100)

tblCollectionsOnlineImage		
PK	<u>CollectionsOnlineImageID</u>	int identity
	fkItemID	int
	fkItemComponentID	int
	fkCollectionID	int
	DigResSummary	ntext
	MasterFile	nvarchar(50)
	RefFile	nvarchar(50)
	ThumbExcerptFile	nvarchar(50)
	PrintFile	nvarchar(50)
	fkPartTypeID	int
	PartNum	int
	OrderNum	int
	fkProjectID	int
	fkAAAPersonID	int
	fkDeviceID	int
	DigitizedDate	datetime
	ArchivedDate	datetime
	Quality	nvarchar(50)
	SuppressedDigRes	bit
	DateEdited	datetime
	BoxNum	int
	FolderNum	int
	fkContainerID	int
	microfile	nvarchar(100)
	ReelNum	varchar(25)
	FrameNum	int

[Research Collections](#) > [Collections Online](#) >

## Moses Soyer Papers

[View Collection](#)[About this Collection](#)[About Collections Online](#)

Moses Soyer to his son David at summer camp, ca. 1937

This site provides access to the papers of **Moses Soyer** in the Archives of American Art that were digitized. The papers have been scanned in their entirety.

The papers of realist painter Moses Soyer date from circa 1905 to 1974 and measure 3.8 linear feet. Within the papers are scattered biographical materials, general correspondence with friends, artists, galleries, art organizations, and others, as well as correspondence with his family, including his son David, writings and notes by Soyer and others, exhibition materials, news clippings, and other printed material, numerous photographs of Soyer, his family, fellow artists, friends and his artwork. Also found is artwork including book illustrations and loose sketches by Soyer. Read more [About this Collection](#).

*Funding for the processing and digitization of this collection was provided by the Terra Foundation for American Art.*

### [Series 1: Biographical Material, 1925-1974](#)



(not scanned)

### [Series 2: Correspondence, 1928-1974](#)



(not scanned)

### [Series 3: Writings & Notes, circa 1930s-1968, undated](#)



(not scanned)

### [Series 4: Printed Material, 1926-1974](#)



(not scanned)

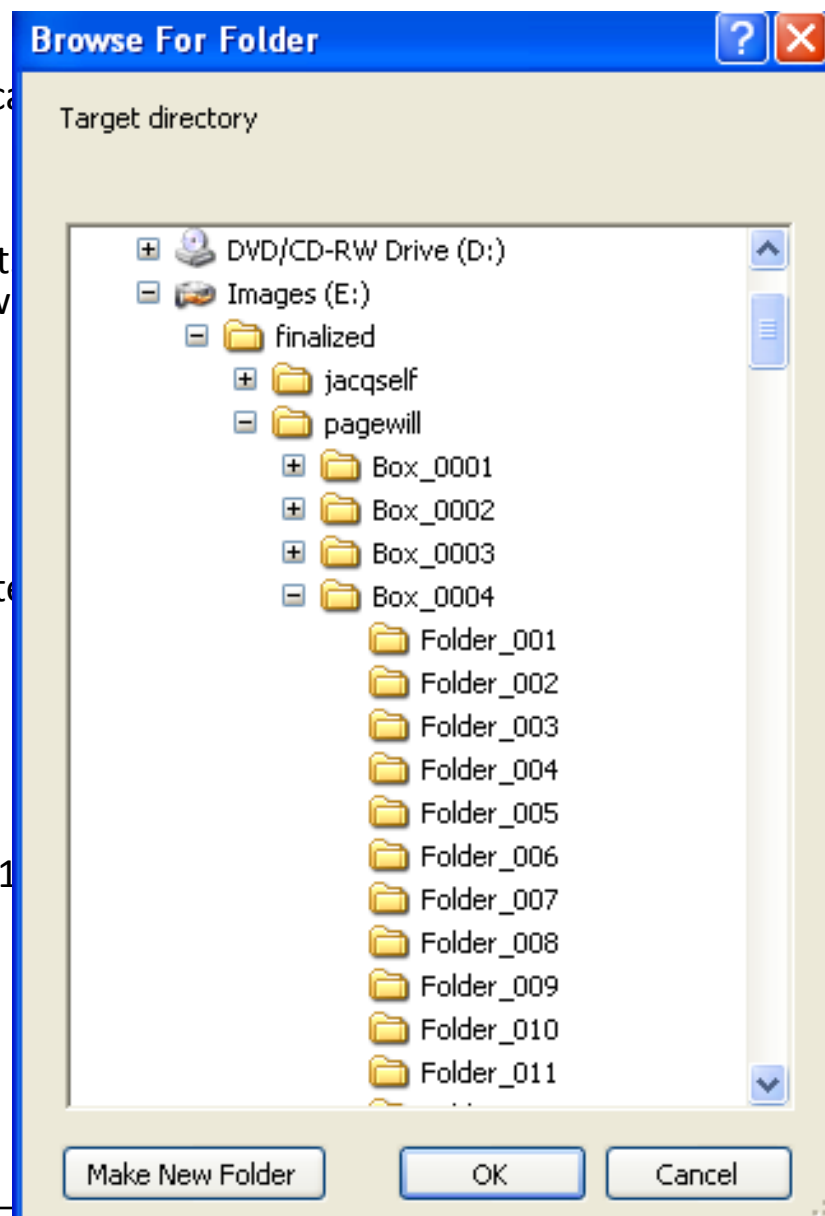


# Sample XML Encoding for Collection to be Scanned

```
<c01 level="series">
<did><unitid>Series 1: </unitid><unittitle>Biographical
</unitdate></unittitle>
<physdesc>(Box 1; 8 folders)</physdesc></did>
<scopecontent>
<p>Biographical Material includes various address list
notes, a catalog with handwritten prices, and other w
and other ephemera.</p>
</scopecontent>
```

```
<c02><did>
<container type="box">1</container>
<container type="folder">1</container>
<unittitle>Address Lists, <unitdate>undated</unitdate>
<physdesc></physdesc>
</did></c02>
```

```
<c02><did>
<container type="box">1</container>
<container type="folder">2</container>
<unittitle>Annotated Catalog with Prices, <unitdate>1
<physdesc></physdesc>
</did></c02>
```



# Collection Online Home Page: Data derived from EAD finding aid

William Page and Page Family Papers - Collections Online

Smithsonian Archives of American Art

Home About Us Research Collections Exhibitions Publications News & Events Support the Archives Ask Us

Research Collections > Collections Online >

## William Page And Page Family Papers, 1815-1947 bulk 1843-1892

[View Collection](#) [More about this Collection](#) [About Collections Online](#)



William Page, 187-?

The papers of **William Page** in the Archives of American Art were digitized in 2007. The bulk of the papers have been scanned and total 15,517 images.

The papers of the portraitist and art theorist William Page and the Page family measure 8.5 linear feet and date from 1815 to 1947, bulk 1843-1892. In addition to the papers of William Page, the papers include documents related to Page's wife's career as a writer and records documenting their personal lives and the lives of their family members. Types of documents found include personal documents and artifacts, correspondence, essays, lectures, diaries, poems, notes and notebooks, financial records, legal records, published works, clippings, catalogs, photographs, and artwork. Read more [About this Collection](#).

*Funding for the processing and digitization of this collection was provided by the Terra Foundation for American Art*

**Series 1: Biographical Materials and Artifacts, 1847-1917, undated**



**Series 2: Correspondence, 1815-1942, undated**



**Series 3: Notes and Writings, 1839-1888, 1949, undated**



**Series 4: Personal Business Records, 1848-1932, undated**



[Image for Series 040000]

<http://www.aaa.si.edu/collectiononline/pagewill/series4.htm>

start PowerPointPresentati... Shifting\_Gears\_RBMS... Workflows for Large ... William Page and Pag... William Page and Pag...



Funding for the processing and digitization of this collection was provided by the Terra Foundation for American Art

William Page, 187-?

**Series 1: Biographical Materials and Artifacts, 1847-1917, undated**



**Series 2: Correspondence, 1815-1942, undated**



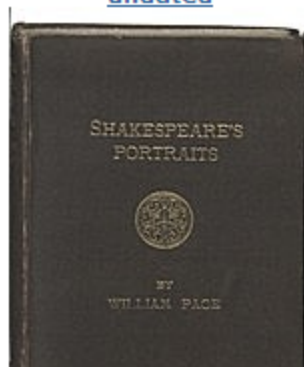
**Series 3: Notes and Writings, 1839-1888, 1949, undated**



**Series 4: Personal Business Records, 1848-1932, undated**



**Series 5: Printed Materials, 1845-1938, undated**



**Series 6: Photographs, 1845-1947, undated**



**Series 7: Artwork, 1856-1874, undated**



## William Page And Page Family Papers, 1815-1947, bulk 1843-1892

View Collection

More about this Collection

About Collections Online

### Series 6: Photographs, 1845-1947, undated (Box 7-9, 10-12; 1.4 linear feet)

This series includes photographic portraits, snapshots, photographs of homes and other buildings, landscape photographs, and photographs of works of art. Most portraits are mounted cabinet cards or cartes-des-visites, but a small number of tintypes, daguerreotypes, and unmounted prints are also found. Portraits depict William Page, his family, descendants, and friends, as well as subjects for his painted portraits. Photographs after 1900 depict primarily the family of George S. Page, William Page's son, his wife and children, and their families. [READ MORE](#)

[PREVIOUS SERIES](#) | [NEXT SERIES](#)



From Series 6:  
Photographs, 1845-1947,  
undated

#### Portrait Photographs

- 2 [William Page, undated](#)
- 3 [Henry Ward Beecher, 1874, undated](#)
- 4 [Reuben E. Fenton, 1860-1875](#) (Oversized items housed in Box 9, sol)
- 5 [Emma A. Fortuna, 1882, undated](#)
- 6 [Sophia Stevens Page \(daughter\), 1882, undated](#)
- 7 [William Page \(grandson\), 1889-1892, undated](#)
- 8 [Wendell Phillips, 1855-1865](#) (Oversized items housed in Box 9, sol)
- 9 [Winfield Scott, 1862](#)
- 10 [B.Frank and Charlotte Stevens, 1882](#)
- 11 [Henry Stevens, undated](#)

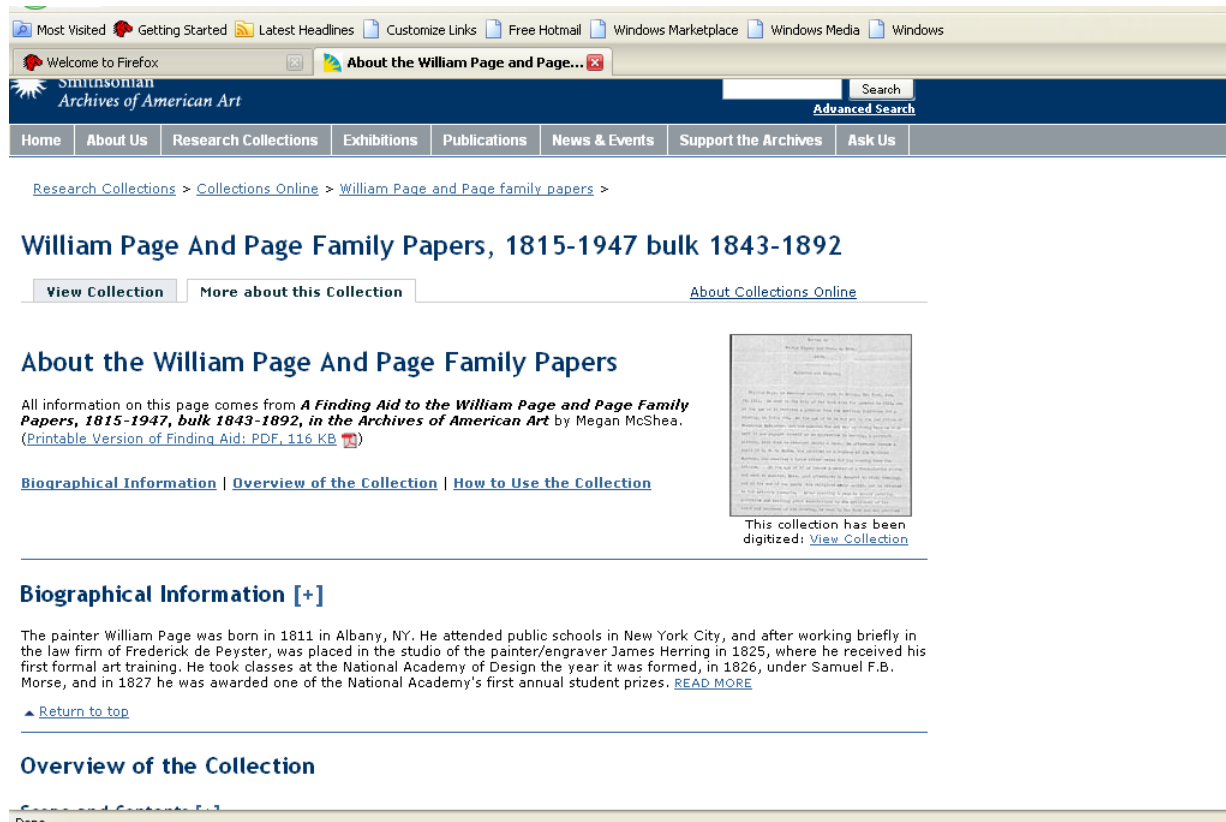
## Series 6: Photographs. Portrait Photographs. Henry Ward Beecher, 1874. (Box 7, Folder 3)

Box 7, Folder 3 << Previous Folder | Next Folder >>

Print



# More About This Collection = Entire Finding Aid



The screenshot shows a web browser window with the address bar displaying "About the William Page and Page...". The page is from the Smithsonian Archives of American Art. The navigation bar includes links for Home, About Us, Research Collections, Exhibitions, Publications, News & Events, Support the Archives, and Ask Us. The main content area is titled "William Page And Page Family Papers, 1815-1947 bulk 1843-1892". Below the title are buttons for "View Collection" and "More about this Collection", and a link for "About Collections Online". The "About the William Page And Page Family Papers" section contains a paragraph of text about the collection, a link to a printable version of the finding aid, and links for biographical information, overview of the collection, and how to use the collection. To the right of the text is a small image of a document page. Below the image is a note that the collection has been digitized and a link to view the collection. The "Biographical Information [+]" section is partially visible, showing the beginning of a paragraph about William Page.

Research Collections > Collections Online > William Page and Page family papers >


## William Page And Page Family Papers, 1815-1947 bulk 1843-1892

[View Collection](#) [More about this Collection](#) [About Collections Online](#)

### About the William Page And Page Family Papers

All information on this page comes from *A Finding Aid to the William Page and Page Family Papers, 1815-1947, bulk 1843-1892, in the Archives of American Art* by Megan McShea.  
([Printable Version of Finding Aid: PDF, 116 KB](#))

[Biographical Information](#) | [Overview of the Collection](#) | [How to Use the Collection](#)



This collection has been digitized: [View Collection](#)

### Biographical Information [+]

The painter William Page was born in 1811 in Albany, NY. He attended public schools in New York City, and after working briefly in the law firm of Frederick de Peyster, was placed in the studio of the painter/engraver James Herring in 1825, where he received his first formal art training. He took classes at the National Academy of Design the year it was formed, in 1826, under Samuel F.B. Morse, and in 1827 he was awarded one of the National Academy's first annual student prizes. [READ MORE](#)

[Return to top](#)

### Overview of the Collection

[Scope and Contents: Full](#)

# Equipment and Scanning Details





# 50+ Years of Microfilming Supports Large-Scale Digitization

Microfilming with Zeutschel microfilm camera,  
2002



Scanning with Zeutschel 10000AI Scanner,  
2005





AAA Collections Online Workflow - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://aaastaff/Workflow/ead/controller/index.cfm?SearchField=Moses+Sc ☆ Google

[Finding Aid / Collections Online Workflow](#) Logged in as **Erin Corley** ([Logout](#))

### Moses Soyer papers

([Choose different collection](#))

Collection Progress Checklist Check-In (Upload) Check-Out (Download) Preview Image Processing

#### Collections Online Workflow Progress

Any tasks that have been ~~eressed out~~ have been completed. The resource is available to the public once all tasks are completed.

1. ~~Archivist uploads new XML File to Finding Aid Repository~~ (Completed by Erin Corley on 3/19/09 11:02 AM [Redo](#) )
2. ~~Supervisor approves XML Finding Aid~~ (Completed by Barbara Aikens on 1/22/09 09:00 AM [Undo](#) )
3. ~~Cataloger provides index terms; archivist adds to XML File~~ (Completed by Erin Corley on 3/18/09 03:01 PM [Undo](#) )
4. [Second supervisor approves XML Finding Aid](#)
5. [Create Image Directory / Begin Scanning](#)
6. [Finished Scanning](#)
7. [Process Images](#)
8. [Deploy Images to Staging server](#)
9. [Archivist Reviews Collection Online on aaasite.si.edu](#)
10. [Archivist reports errors to be corrected](#)
11. [Archivist confirms errors have been corrected and approves Collection Online](#)
12. [Supervisor deploys XML file, HTML file, and Collection Online to Web](#)
13. [Supervisor verifies pages on Web site \(Collections Online, Col. A-Z, & Finding Aids\), notifies staff](#)
14. [Cataloger Updates Smithsonian Institution Research Information System \(SIRIS\)](#)

# Findings/next steps

- Collections Online has increased no. of visits to the Archives' website and is the most popular link from the homepage
- Digitization funding is supporting processing and creation and management of finding aids
- Out-sourcing and/or scan on demand digitization is possible
- Flexible programming provides web 2.0 potential
- Requires web development team, new skills, Usability studies
- Standards based approach can lead to open-source development (Archivists Toolkit)

# For more information

- Karen Weiss – [weissk@si.edu](mailto:weissk@si.edu)
- Barbara Aikens – [aikensb@si.edu](mailto:aikensb@si.edu)
- AAA website: <http://www.aaa.si.edu>
- Collections Online:  
<http://www.aaa.si.edu/collectionsonline>