

Describing web archives: a standard with an identity crisis?

Jessica Cebra
Metadata Management Librarian



Stanford | LIBRARIES

SearchWorks catalog

All fields

books & media



Library services

Advanced search Course reserves Selections (1)

Books, media, & more

Find materials by...

Access	
At the Library	6,968,215
Online	2,974,208
On order	12,756

Resource type

Library

Language

Featured resources



Digital collections

Images, maps, data, & more from the Stanford Digital Repository.



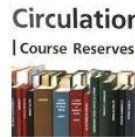
Government documents

Stanford is a depository library for state, federal, UN, and EU government documents.



Theses & dissertations

All student work held in the Stanford Libraries & Digital Repository. Show Stanford work only.



Circulation Course reserves

Course Reserves

Find books, media, and e-resources set aside for classes.



Databases

A-Z list of topic-specific databases. Not sure where to start? Try these selected databases.



IIF resources

Find IIF-compatible items in the Stanford Digital Repository.

What is IIF?

Library services ▾

Books, media, & more

Find materials by...

Access ▾











At the Library	6,968,215
Online	2,974,208
On order	12,756

Resource type >

Library >

Language >

Resource type ▾

 Archive/Manuscript	40,045
 Archived website	552
 Book	7,985,684
 Database	2,525
 Dataset	1,412
 Equipment	335
 Image	330,885
 Journal/Periodical	466,546
 Map	184,006
 Music recording	230,572
 Music score	106,457
 Newspaper	8,577
 Object	1,738
 Software/Multimedia	9,142
 Sound recording	12,599
 Video	138,584

Course reserves Selections (1) ▾

 **Government documents**

Stanford is a depository library for state, federal, UN, and EU government documents.

 **Course reserves**

Find books, media, and e-resources set aside for classes.

 **IIF resources**

Find IIF-compatible items in the Stanford Digital Repository.

What is IIF?

Basic descriptive metadata requirements for web archives in
SearchWorks

- Title of website
- Name of website creator(s) and/or publishers
- Language(s) used in the site
- Brief abstract describing the site
- Subject terms for the site's content
- website collector name
- type of resource (text)
- genre (archived website)
- date of website capture
- form (electronic)
- digital origin (born digital)
- internet media type (text/html)
- original site URL
- "archived by" note
- web archiving service note
- repository (Stanford University. Libraries)
- SWAP URL
- PURL
- record administrative info

- Title of website
- Name of website creator(s) and/or publishers
- Language(s) used in the site
- Brief abstract describing the site
- Subject terms for the site's content
- website collector name

- type of resource (text)
- genre (archived website)
- date of website capture
- form (electronic)
- digital origin (born digital)
- internet media type (text/html)
- original site URL
- "archived by" note
- web archiving service note
- repository (Stanford University. Libraries)
- SWAP URL
- PURL
- record administrative info

Metadata Object Description Schema

(MODS)

Descriptive Metadata for Web Archiving: Recommendations of the OCLC Research Library Partnership Web Archiving Metadata Working Group

DATA ELEMENTS AND USAGE GUIDELINES

Collector	Extent	Source of description
Contributor	Genre/Form	Subject
Creator	Language	Title
Date	Relation	URL
Description	Rights	

Descriptive Metadata for Web Archiving: Recommendations of the OCLC Research Library Partnership Web Archiving Metadata Working Group

DATA ELEMENTS AND USAGE GUIDELINES

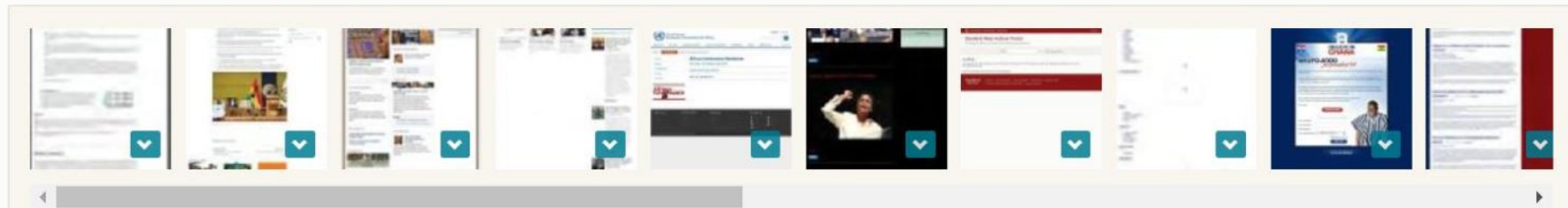
 Collector	 Extent	 Source of description
 Contributor	 Genre/Form	 Subject
 Creator	 Language	 Title
 Date	 Relation	 URL
 Description	 Rights	

African Politics Web Archive

Type of resource Text

Date captured 2007-2015

Digital collection preview



Online

Explore this collection (**172 digital items**)

Description

Creators/Contributors

Collector

Stanford University. Libraries. Humanities and Area Studies Resource Group


Subjects

Genre


Archived website

Digital content




05-Jan-2011 




12-Jan-2011 




19-Jan-2011 



26-Jan-2011 



02-Feb-2011 



New Search

Save Records

MARC View

Return to List

Modify Search

More Like This

Keyword

Arcade (e-resources)

Search

375 results found. Sorted by [relevance](#) | [date](#) | [title](#) .

(Search History)

Record: [Prev](#) [Next](#)

Title **Picturing the Americas : landscape painting from Tierra del Fuego to the Arctic.**

Publisher Chicago, Illinois : The Terra Foundation for American Art, [2017]

©2017

Description 1 online resource : color illustrations



[More research tools](#)

Connect to:

[Live site](#)

[Archived site](#)

Note Online supplement to exhibition held from June 20 - October 20th, 2015 at the Art Gallery of Ontario in Toronto, from November 7, 2015 - January 18, 2016 at the Crystal Bridges Museum of American Art in Bentonville, Arkansas, and from February 27 - May 29, 2016 at the Pinacoteca do Estado de São Paulo.

Title from caption (viewed on May 13, 2017).

Archived websites may not be exact copies of the original **website** at the time of the web crawl due to technical or privacy considerations.

Summary Online exhibition featuring an interactive map, timeline, artist biographies, and listing of themes, complementing the eponymous traveling exhibition and catalog.

Language note Parallel versions of site in English, French, Spanish, and Portuguese.

System details Mode of access: World Wide Web.

Description source Description of the resource based on live site viewed on May 13, 2017, and archived site; title from home page.

LC subject [Landscape painting -- 19th century -- Exhibitions.](#)
[America -- In art -- 19th century -- Exhibitions.](#)
[Landscape painting -- 20th century -- Exhibitions.](#)
[America -- In art -- 20th century -- Exhibitions.](#)

Genre/Form [Web sites.](#)

Related corporate bodies [Terra Foundation for American Art, issuing body.](#)

[Art Gallery of Ontario, host institution.](#)

[Crystal Bridges Museum of American Art, host institution.](#)

[São Paulo \(Brazil : State\), Pinacoteca do Estado, host institution.](#)

WorldCat no. 987459783

Permanent link to this page:

<https://arcade.nyarc.org:443/record=b1415935~S16>

Note Online supplement to exhibition held from June 20 - October 20th, 2015 at the Art Gallery of Ontario in Toronto, from November 7, 2015 - January 18, 2016 at the Crystal Bridges Museum of American Art in Bentonville, Arkansas, and from February 27 - May 29, 2016 at the Pinacoteca do Estado de São Paulo.

Title from caption (viewed on May 13, 2017).

Archived websites may not be exact copies of the original **website** at the time of the web crawl due to technical or privacy considerations.

Summary Online exhibition featuring an interactive map, timeline, artist biographies, and listing of themes, complementing the eponymous traveling exhibition and catalog.

Language note Parallel versions of site in English, French, Spanish, and Portuguese.

System details Mode of access: World Wide Web.

Description source Description of the resource based on live site viewed on May 13, 2017, and archived site; title from home page.

LC subject Landscape painting -- 19th century -- Exhibitions.
America -- In art -- 19th century -- Exhibitions.
Landscape painting -- 20th century -- Exhibitions.
America -- In art -- 20th century -- Exhibitions.

Genre/Form Web sites.

Related corporate bodies Terra Foundation for American Art, issuing body.

Art Gallery of Ontario, host institution.

Crystal Bridges Museum of American Art, host institution.

São Paulo (Brazil : State). Pinacoteca do Estado, host institution.

WorldCat no. 987459783



Permanent link to this page:

<https://arcade.nyarc.org:443/record=b1415935~S16>

bibliographic

archival

bibliographic

identification
discovery
context

archival

origin, authenticity
accountability
provenance, processing

- selecting what and what not to keep
- using a technological tool (which may have limitations of its own) to capture the content, and almost always results in an ‘incomplete’ copy of the original
- crawl configuration
- playback environment and its display behavior

UC Guidelines for Born-Digital Archival Description

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

Table of Contents:

[INTRODUCTION](#)

[Background and Methodology](#)

[Goals](#)

[How to Use These Guidelines](#)

[Systems of Record](#)

[Level of Description](#)

DESCRIPTIVE ELEMENTS

Processor

Physical Description and Extent

Abstract

Conditions Governing Access and Restrictions

Conditions Governing Reproduction and Use

Physical Characteristics and Technical Requirements

Immediate Source of Acquisition

Appraisal Information

Scope and Content

Organization and Arrangement

Processing Information

Container List & Inventory

Note to

request

Annalise

Table of

INTRODU

Backg

Goals

How to Use these Guidelines

Systems of Record

Level of Description

are pull

to:

DESCRIPTIVE ELEMENTS



Processor

Physical Description and Extent

Abstract

Conditions Governing Access and Restrictions

Conditions Governing Reproduction and Use

Physical Characteristics and Technical Requirements

Immediate Source of Acquisition



Appraisal Information

Scope and Content

Organization and Arrangement



Processing Information

Container List & Inventory

Note to

request

Annalise

Table of

INTRODU

Backg

Goals

How to Use these Guidelines

Systems of Record

Level of Description

are pull

to:

“Decisions made during processing can greatly affect who, what, where, when, why and how researchers access and understand the digital material within a given collection. For us, the Processing Information section is therefore one of the most important aspects of any finding aid that describes born-digital materials. In particular, processing legacy born-digital material can often involve changing the nature of the data to preserve it and make it accessible.”

[Berdini et al., “Describing Digital,” p.8](#)

Consider the following questions while describing a group of seeds:

What is this collection for? Who is it for?

What is the general scope and content?

Why these seeds and not others? What was the selection criteria?

A collection-level record is recommended to describe a curated group of web archive seeds. A descriptive overview to highlight a theme, topic(s), event, or other criteria can provide much needed context for a collection of individual websites. Recent literature about the perspective of researchers using web archives collections has identified the user-end need for transparency and understanding the curator's rationale for the selection of websites to be archived, contributing to an emerging concept of 'web archives provenance'.

Inspiring works

[NYARC Documentation, Metadata Application Profile for Description of Websites with Archived Versions Version 2 \(August 2018\)](#)

[UC Guidelines for Born-Digital Archival Description](#)

[Descriptive Metadata for Web Archiving. Recommendations of the OCLC Research Library Partnership Web Archiving Metadata Working Group](#)

Francis X. Blouin, William G. Rosenberg. *Processing the Past: Contesting Authorities in History and the Archives*. Oxford University Press, 2011

[Emily Maemura. "What's Cached is Prologue: Reviewing Recent Web Archives Research Towards Supporting Scholarly Use," University of Toronto, 2018](#)

[Yasmin AlNoamany. "Using Web Archives to Enrich the Live Web Experience through Storytelling," Old Dominion University dissertation \(2016\)](#)

[Amy Wickner. "Web Archiving and You.Web Archiving and Us," Code4Lib 2018](#)

[Gregory Wiedeman. "Describing Web Archives with the Partner Data API," Archive-It Partner Meeting, 2018](#)

"Ultimately web archiving is about capturing and recapturing aspects of the experience and performance of the live web, and it's up to collectors, users, and subjects to negotiate together what exactly that means."

[Amy Wickner, "Web archiving and you. Web archiving and us,"](#)
[Code4Lib 2018](#)

Describing web archives: a standard with an identity crisis?

Jessica Cebra, cebraj@stanford.edu
Metadata Management Librarian



Stanford | LIBRARIES

Thank you!