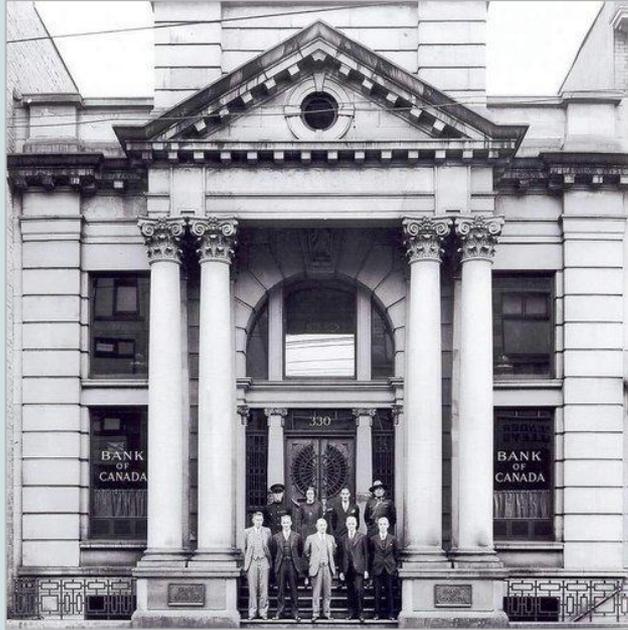


Collaborative Community Collecting with the Internet Archive

Sarah Beth Seymore
Program Officer
Archiving and Data Services
Internet Archive





The Internet Archive is a non-profit digital library of Internet sites and other cultural artifacts in digital form. Like a paper library, we provide free access to researchers, historians, scholars, people with perceptual disabilities, and the general public.

Archive.org - Free Storage and Access

The screenshot shows the Internet Archive website interface. At the top, there's a navigation bar with categories like WEB, BOOKS, VIDEO, AUDIO, SOFTWARE, and IMAGES. Below that, a secondary navigation bar includes links for ABOUT, BLOG, PROJECTS, HELP, DONATE, CONTACT, JOBS, VOLUNTEER, and PEOPLE. The main content area features the Newark Public Library logo and a list of navigation links such as 'Browse Collections A to Z', 'Lesson Plans & Subject Guides', and 'Library Exhibits'. A search bar is visible on the right. Below the navigation, there's a section for 'COLLECTION' with 'ABOUT' and '53,752 Results'. A search filter section allows users to search metadata or text contents, and a 'Sort by' dropdown is set to 'Weekly views'. The bottom part of the screenshot shows a grid of collection thumbnails, including 'Newark Yearbooks' and 'Al Henderson Photos'.

Search the history of over 737 billion
web pages on the Internet.

WayBack Machine

enter URL or keywords

Mobile Apps

Wayback Machine (iOS)

Wayback Machine (Android)

Browser Extensions

Chrome

Firefox

Safari

Edge

Archive-It Subscription

Explore the Collections

Learn More

Build Collections

Save Page Now

Capture a web page as it appears now
for use as a trusted citation in the future.

https://

SAVE



Internet Archive Services and Tools

- Digitization
 - Interlibrary Loan
 - Controlled Digital Lending
 - Digital Preservation
 - Web Archiving
 - Custom Harvesting
 - Scholarly Preservation
 - Access Services
- **Archive-It** - subscription web archiving service
 - **Vault** - digital preservation repository and storage system
 - **IA Scholar** - scholarly material access interface
 - **ARCH** - platform for digital collection analysis



Community Programs at the Internet Archive

Community Programs helps memory and research organizations around the world provide universal access to all knowledge. We achieve this mission through programs backed by Internet Archive's internationally distributed nonprofit infrastructure and services. We work with partners to develop sustainable efforts that respond directly to community needs.



Priorities

- Increase preservation of and access to underrepresented and/or under resourced knowledge
- Strengthen memory worker and researcher communities (e.g., professional development, and stakeholder convening and action)
- Create and contribute to open and interoperable knowledge infrastructures (e.g., persistent IDs, standards, preservation and access systems)
- Support responsible computational engagement with the cultural record (e.g., text and data mining, machine learning, AI)

Community Programs at the Internet Archive



- **Program Officers**
 - Communications Management
 - Collection building
 - Project Management of technical work
- **Engineering Staff**
 - Build Access Portal
 - Generate datasets
 - Metadata integrations
- **Technical Services**
 - Archive-It Managed Account
 - Hosting Research and Access Portal



Advance the capacity for public libraries and other community-based memory organizations to build digital collections and archives documenting local and community histories, with an emphasis on communities that have been underrepresented in the historic record.



Community Webs includes...

- 180+ Member Organizations in:
 - 7 Countries
 - 41 US States and Territories
 - 7 Canadian Provinces

Expanding a national cohort of librarians, community organizers, who are building a vast, inclusive digital archive that prioritizes underrepresented voices through skills development and digital library services



archive-it.org

HOME EXPLORE LEARN MORE CONTACT US

The leading web archiving service for collecting and accessing cultural heritage on the web
Built at the Internet Archive

Welcome to Archive-It! Contact Us to learn more about our products and services.

Explore Collections Find a Collection by Name Search Show All Collections

Clinical Translational Science Award
By National Institutes of Health
The Clinical Translational Science Award Collection document and preserve the progress of the Clinical Translational Science Award websites.

Culture and Sensitivity
By NYC Public School 56 Q - The Harry Eichler School
This collection, from fifth graders at New York City PS 56 in Queens is their student's guide to the wonderful arts and varieties of cultures.

Jasmine Revolution - Tunisia 2011
By Internet Archive Global Events
A collection of websites, news coverage, and commentary surrounding the 2011 Jasmine Revolution in Tunisia. Our partners at Library of Congress and Bibliothèque Nationale de...

Explore Collecting Organizations Schomburg Center for Research in Black Culture Search Show All Organizations

Home

Home Collections Crawls Archives ARS Help Center Welcome, system

Community Webs Training

2.1 GB archived since Apr 21, 2021

2.1 GB ARCHIVED

Data Budget Usage

Active Collection List (in Active Collections)

Type to Filter Active Collections

Create a Collection Download Collection List

Collection Name	Data (this period)	Docs (this period)	Active Seeds	Last Crawl
Washington County Organiz...	619.3 MB	27,983	4	Aug 21, 2022
Community Webs	471.2 MB	5,690	4	Sep 1, 2022
Calgary Farmers' Markets	0 bytes	0	5	Dec 1, 2022
Source Collection	0 bytes	0	3	
Target Collection	0 bytes	0	0	

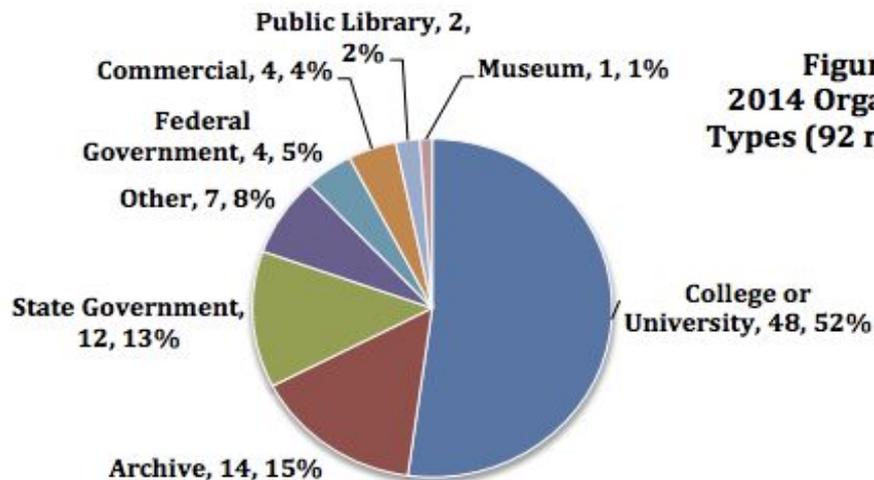
You are viewing an archived web page collected at the request of [Community Archives of Belleville and Hastings County](#) using [Archive-It](#). This page was captured on 06:00:52 Mar 15, 2024, and is part of the [Community Groups](#) collection. The information on this web page may be out of date. See [All versions](#) of this archived page. Found 0 archived media items out of 0 total on this page. [Metadata](#) [Enable QA](#)

Islamic Society of Belleville

Serving the City of Belleville, Bay of Quinte, Prince Edward County, and Trenton

Over 180 public libraries and community archives participate in the program from 41 states and 4 U.S. territories. Collectively, Community Webs members have built nearly 1,300 web archive collections of over 30,000 websites totaling hundreds of millions of primary source records and over 90 terabytes of archived data.

Why Community Webs?



**Figure 2:
2014 Organization
Types (92 responses)**

Web Archiving in the United States: A 2013 Survey: An NDSA Report, 2014:
https://ndsa.org/documents/NDSA_USWebArchivingSurvey_2013.pdf

Why Community Webs?

As of 2014...

2%

As of 2023...

18%

of Web Archiving Organizations were
Community Memory Organizations

Community Webs members receive...

- Archive-It Subscription
- Vault Repository + Storage (500GB)
- Training in Web Archiving, Digital Preservation and Community Archiving
- Events & Webinars
- Networking, Peer Support, Collaboration
- Collaboration and Pathways for Community Partnership
- Professional Development Stipends
- Access to Collections
- Digitization opportunities

The screenshot shows the website for the Community Archives of Belleville and Hastings County. At the top, there is a navigation menu with links for Home, About, Collections, Participants, Curriculum, News, and Apply. The main header features the organization's logo and a description: "The Community Archives collects and shares information about the history of Belleville and Hastings County region of Ontario through the records of its people, organizations, and local governments." Below the header is a search bar with a "Search" button and a "Clear" button. Underneath the search bar, there are tabs for "Collections", "Sites", and "Page Text". The search results are displayed under the heading "Narrow Your Results". On the left side, there is a filter for "Organization" with a dropdown menu showing "Community Archives of Belleville and Hastings County (8)". The main content area shows two search results. The first result is for the "2021 Federal election", with the organization listed as "Community Archives of Belleville and Hastings County" and a description: "Websites of candidates in the Hastings County region for the 2021 Federal Election." A small thumbnail image of a news broadcast is visible next to this result. The second result is for the "2022 Municipal election". At the top of the results area, there are "Previous Page" and "Next Page" buttons, and a page indicator showing "Page 1 of 1 (8 Total Results)".

Collaborative Access to Diverse Public Library Local History Collections

- The National Historical Publications and Records Commission (NHPRC) of the National Archives support projects that promote access to America's historical records to encourage understanding of our democracy, history, and culture.
- The collaborative effort will digitize and provide access to a diverse range of content from across the country that represents the experience of immigrant, indigenous, and African American communities throughout the United States.



Community Webs' members have...

Presented their work at 20+ conferences, and written dozens of blog posts and publications

Outreach to local groups, organizations, and individuals for URL nominations

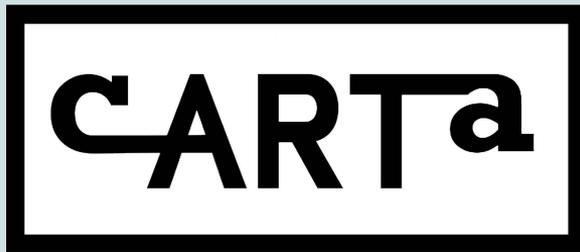
Developed coursework and workshops for local MLIS students

Collaborated on successful grant proposals for digitization and digital preservation projects

Hosted web archiving panel discussions

Served as Chair/Vice Chair of the SAA Web Archiving Section





CARTA is a collaborative entity of art libraries building collections of archived web-based content related to art history and contemporary art practice.

Through this collaborative approach, the project leverages shared infrastructure, expertise and collecting activities amongst participating organizations, scaling the extent of web-published, born-digital materials preserved and accessible for art scholarship and research.



NATIONAL
ENDOWMENT
FOR THE
HUMANITIES

CARTA Program Origins

The Collaborative ART Archive (CARTA) is a collaborative program developed from a partnership between the Internet Archive and the New York Art Resources Consortium (NYARC) aimed at capturing and preserving at-risk web-based art materials.

Project Origins:

- 2018 - IMLS National Leadership Grant for Libraries
 - Environmental scans, catalyze conversations, mobilize community
 - Events, outreach, training, meetings to propel this work
- 2021 - Grant funding from NEH and IMLS
 - Formalized a collaborative group for coordinated collecting and access
 - Built access portal, scaled collaborative collecting, datasets and other methods for promoting use



Goals and Activities

- **Community formation and growth**
 - Member recruitment, collaborative meetings and discussions
- **Collection building**
 - Creating collections, preserving content, adding metadata
- **Facilitating Research and Access Methods**
 - Building access portal
 - Metadata integrations
 - Datathons and promotion of use with end users/researchers
- **Sustainability, Promotion and Evaluation**
 - Publicizing the access portal, research use cases
 - Sustainability planning
 - Reviewing metrics for collections/use and fill gaps as appropriate

CARTA Members

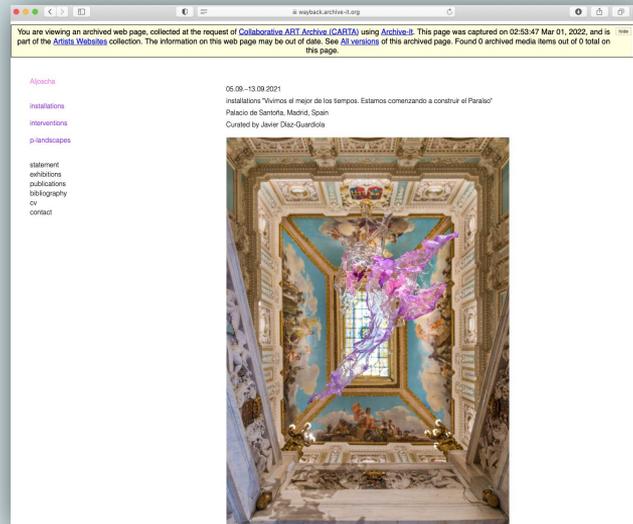
41 participating organizations so far:

American Craft Council
American Folk Art Museum
Amon Carter Museum of American Art
ART | library deco
Art Gallery of Ontario
Art Institute of Chicago
Artexite
Bibliothèque et Archives nationales du Québec (BANQ)
Birmingham Museum of Art
Cleveland Museum of Art
Detroit Institute of Arts
Emily Carr University of Art + Design
Fashion Institute of Technology - SUNY
Getty Research Institute (Getty Library)
Harvard University - Fine Arts Library
Harvard University - Graduate School of Design
Indianapolis Museum of Art at Newfields
Leonardo / ISAST
Los Angeles County Museum of Art (LACMA)
Maryland Institute College of Art (MICA)

Museum of Contemporary Art Chicago
Museum of Contemporary Art of Georgia (MOCA GA)
National Gallery of Art Library
National Gallery of Canada
New York Art Resources Consortium (NYARC)
Oklahoma State University
Philadelphia Museum of Art
Rhode Island School of Design
Rungt Cultural Society
San Francisco Museum of Modern Art (SF MOMA)
Smithsonian Libraries and Archives
Smithsonian Institution - Archives of American Art
Sterling and Francine Clark Art Institute Library
The Clyfford Still Museum
The Corning Museum of Glass
The Menil Collection
The Metropolitan Museum of Art
The Nelson-Atkins Museum of Art, Spencer Reference Library
University of Hawaii at Manoa, Hamilton Library
University of North Carolina, Greensboro
Walker Art Center

Collection Topics

<u>Collection Name</u>	<u>Description</u>
	Web-based content and blogs providing critical discussion and evaluation of works of art.
Art Criticism	
Art Fairs and Events	International fairs and events specializing in the exhibition and sale of art.
Art Galleries	Art galleries, gallery exhibitions, and art dealers.
	Web resources and documentation related to the study of art and scholarship in art history.
Art History/Scholarship	
Artists Websites	Websites created by individual artists and/or estates to showcase their artworks and biographies.
Arts Education	Web-based educational materials and programs for artists.
	Organizations, museums, and artist-run centers that focus on community arts and empowerment.
Arts Organizations	
Auction Houses	Auction houses and catalogs that specialize in arts.

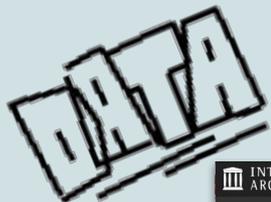


Web Archive Datathon/Workshops

<https://archive-it.org/blog/data-analysis-workshop-series/>



CARTA Whitepaper & Research Datasets



The screenshot shows the Internet Archive search results for 'CARTA'. The search bar at the top indicates 15 results. The main content area displays the title 'CARTA: Promoting Computational Research with Art on the Web' by Thomas Padilla, Sarah Beth Seymore, Sumitra Duncan, and Karl Blumenthal, dated 9/23/2023. Below the title is a paragraph of text describing the project. To the right, a grid of 15 search results is visible, each featuring the Internet Archive logo and a brief description of a dataset, such as 'Collaborative ART Archive (CARTA) word processing file information dataset' and 'Collaborative ART Archive (CARTA) domain graph data'. The interface includes navigation links like 'ABOUT', 'BLOG', 'PROJECTS', and 'HELP', as well as a search bar and a 'SIGN UP | LOG IN' button.



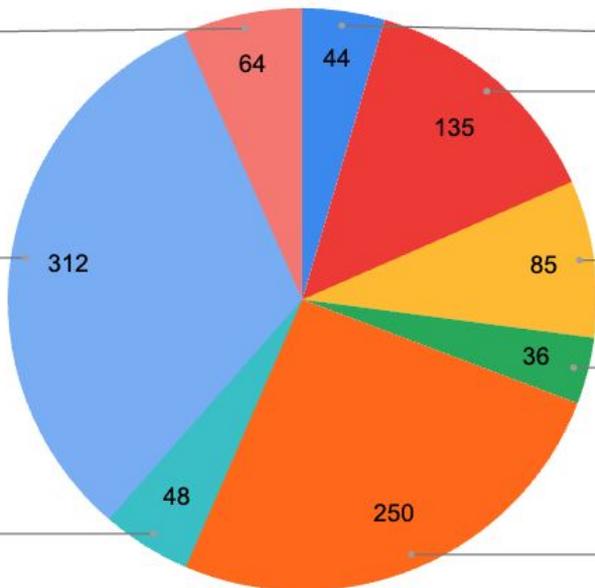
Collections: 900+ websites (16 TB data)

of Seeds*

Auction Houses
6.6%

Arts Organizations
32.0%

Arts Education
4.9%



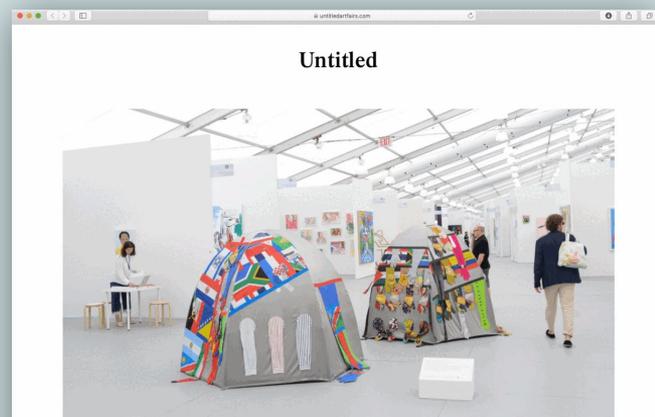
Art Criticism
4.5%

Art Fairs and Events
13.9%

Art Galleries
8.7%

Art History/Scholarship
3.7%

Artists Websites
25.7%



Untitled Art is an innovative and inclusive platform for discovering contemporary art. It balances intellectual integrity with cutting-edge experimentation, refreshing the standard fair model by embracing a unique curatorial approach.

CARTA Portal Launched: April 2023

CARTA

About

Collections

Participants

News

Join CARTA



Photo by Michael Bodycomb of The Frick Collection

About the Collaborative ART Archive (CARTA)

The Internet Archive and the New York Art Resources Consortium (NYARC) have spearheaded this collaborative project aimed at capturing and preserving at-risk web-based art materials. CARTA is a collaborative entity of art libraries building collections of archived web-based content related to art history and contemporary art practice. Through this collaborative approach, the project leverages shared infrastructure, expertise and collecting activities amongst participating organizations, scaling the extent of web-published, born-digital materials preserved and accessible for art scholarship and research. The goals are to promote streamlined access to art reference and research resources, enable new types of scholarly use for art-related materials, and ensure that the art historical record of the 21st century is readily accessible far into the future.

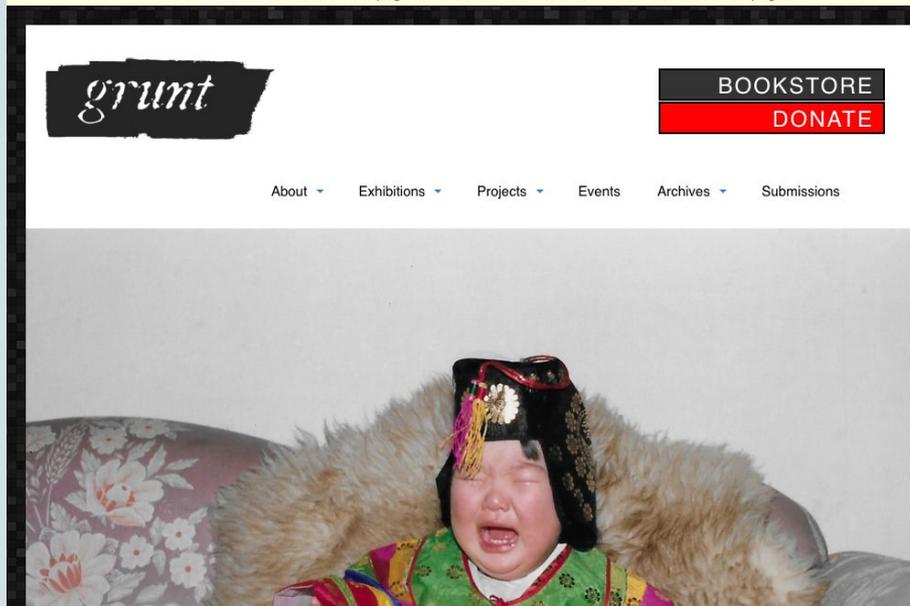
<https://carta.archive-it.org/>



Emily Carr University of Art + Design

You are viewing an archived web page collected at the request of [Collaborative ART Archive \(CARTA\)](#) using [Archive-It](#). This page was captured on 23:26:38 May 25, 2023, and is part of the [Arts Organizations](#) collection. The information on this web page may be out of date. See [All versions](#) of this archived page. Found 0 archived media items out of 0 total on this page. [Enable QA](#) hide

You are viewing an archived web page collected at the request of [Collaborative ART Archive \(CARTA\)](#) using [Archive-It](#). This page was captured on 23:23:13 Nov 25, 2023, and is part of the [Arts Organizations](#) collection. The information on this web page may be out of date. See [All versions](#) of this archived page. Found 0 archived media items out of 0 total on this page.



RE A

ional Centre
r Asian Art

ntorship

ter Anti-
igital

ents

 Search

CARTA: What's next?

- Continued growth of collections and membership after the grant period
- Developed/finalized [membership model](#), aimed at covering program costs in a flexible and sustainable manner, and encouraging participation from a diverse range of interested organizations and individuals
- Formation of Advisory Board in late 2023 and transition to self-governance

How do I join?

- Submit an application form:
 - communitywebs.archive-it.org/apply
 - <https://carta.archive-it.org/join/>
- We need:
 - Information about **you**
 - Information about **your organization**
 - Information about **your community**
 - Information about **your collecting interests**
- Rolling deadline!
- Approval take 1-2 weeks to process



Thank you!

sarahbeth@archive.org
communitywebs@archive.org
carta@archive.org