



# How It Started / How It's Going: 25 Years of Web Archiving at the Library of Congress

February 19, 2026

**#DPClinic**  
**Digital Preservation Coalition**

Abbie Grotke, Section Head  
Melissa Wertheimer, Senior Digital Collections Specialist  
Tracee Haupt-Fugate, Digital Collections Specialist

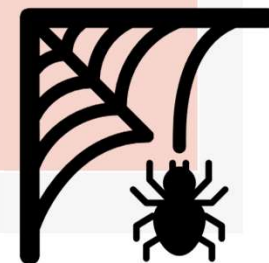
Web Archiving Section  
[webcapture@loc.gov](mailto:webcapture@loc.gov)

**LIBRARY  
OF CONGRESS**

Digital Collections Management and Services Division

# Web Archiving: What, Why, How

<b>What is it?</b>	<p>“Web archiving is the process of <b>collecting</b> portions of the World Wide Web, <b>preserving</b> the collections in an archival format, and then <b>servng</b> the archives for access and use.”</p> <p>-- <i>International Internet Preservation Consortium, “Web Archiving”</i> (<a href="https://netpreserve.org/web-archiving/">https://netpreserve.org/web-archiving/</a>)</p>
<b>Why do/should we do it?</b>	<p>Documents and cultural heritage content of interest to memory institutions are:</p> <ul style="list-style-type: none"><li>• On the web*</li><li>• Dynamic - change and get created rapidly</li><li>• Ephemeral - disappear at any time</li></ul> <p>* <i>Accessed via the web ≠ always web-archivable!</i></p>
<b>How is it done?</b>	<ol style="list-style-type: none"><li>1. <b>COLLECT:</b> Archival-quality software known as “crawlers”</li><li>2. <b>PRESERVE:</b> WARC files**</li><li>3. <b>SERVE:</b> Users interact with captures through browser-based replay software</li></ol> <p>** <a href="https://www.loc.gov/preservation/digital/formats/fdd/fdd000236.shtml">https://www.loc.gov/preservation/digital/formats/fdd/fdd000236.shtml</a></p>



# 25 Years of Web Archiving

**LIBRARY**  
LIBRARY OF CONGRESS

Note: External links, streaming audio/video, forms and search boxes may not function within this web archive  
[Back to Calendar View](#)

Library of Congress Web Archives  
<http://www.loc.gov/>  
Archived: 03/01/2000 at 06:43:40

◀ [First \(06/16/1997\)](#) | [Previous](#) #17 of 6830 | [Next](#) | [Last \(02/15/2023\)](#) ▶

[Ask a Librarian](#) ▲

*The* **LIBRARY of CONGRESS**

[SEARCH THE CATALOG](#) | [SEARCH OUR WEB SITE](#) | [WEB SITE MAP](#)

**USING the LIBRARY**  
Catalogs, Collections & Research Services



**AMERICAN MEMORY**  
America's Story in Words, Sounds & Pictures



**THOMAS**  
Congress At Work



**BICENTENNIAL**  
**1800-2000**  
Libraries • Creativity • Liberty



**EXHIBITIONS**  
An On-Line Gallery



**COPYRIGHT OFFICE**  
Forms & Information



**HELP & FAQs**  
General Information

**THE LIBRARY TODAY**  
News, Events & More



Above, the interior of the dome of the Main Reading Room of the Library of Congress

101 INDEPENDENCE AVE. S.E.  
WASHINGTON, D.C. 20540  
(202) 707-5000

Comments: [lcweb@loc.gov](mailto:lcweb@loc.gov)  
[NOTICE](#)

[USING the LIBRARY](#) | [THOMAS](#) | [COPYRIGHT OFFICE](#) | [AMERICAN MEMORY](#) | [EXHIBITIONS](#) | [The LIBRARY TODAY](#) | [BICENTENNIAL](#) | [HELP & FAQs](#) | [TOP of PAGE](#)

Source URL: <http://www.loc.gov/>

Capture Date: March 1, 2000

Web Archive: <https://www.loc.gov/item/lcwaN0003464/>

Digital Collections Management & Services Division

# 25 Years of Web Archiving

**LIBRARY**  
LIBRARY OF CONGRESS

Note: External links, streaming audio/video, forms and search boxes may not function within this web archive  
[Back to Calendar View](#)

**Library of Congress Web Archives**  
<http://lcweb2.loc.gov/cocoon/minerva/html/minerva-home.html>  
 Archived: 07/26/2006 at 20:08:14

[Ask a Librarian](#)

[« First \(07/26/2006\)](#)
[« Previous #1 of 108](#)
[Next»](#)
[Last \(11/11/2021\)»](#)

The Library of Congress >> More Online Collections

## MINERVA

Mapping the INTERNET Electronic Resources Virtual Archive

[home](#)

**Web Archives Available:**

- **107th Congress NEW!**  
December 12, 2002
- **Election 2002**  
Jul. 1, 2002 - Nov. 30, 2002
- **September 11, 2001**  
Sep. 11, 2001 - Dec. 1, 2001
- **Election 2000**  
Aug 1, 2000 - Jan 21, 2001

*Coming soon...*

- **Winter Olympic Games 2002**  
Feb. 8 - 24, 2002
- **9/11 Remembrance**  
Sep. 11 - 18, 2002
- **Iraq War**  
Mar. 2003 - Ongoing
- **108th Congress**  
Jan. 2004 - Dec. 2004
- **Election 2004 Web Archive**  
Oct. 2003 - Dec. 2004



[About MINERVA](#)

[Collection Policies](#)

[Documents & Articles](#)

[Metadata](#)

[Presentations](#)

[Awards](#)

[Contact Us](#)

The *MINERVA Web Archiving Project* was established to initiate a broad program to collect and preserve these primary source materials. A multi disciplinary team of Library staff representing cataloging, legal, public services, and technology services is studying methods to evaluate, select, collect, catalog, provide access to, and preserve these materials for future generations of researchers.

[home](#)

The Library of Congress >> More Online Collections  
November 22, 2005

[Contact Us](#)

**Capture Date:** June 7, 2006

**Web Archive:** <https://webarchive.loc.gov/all/20060726200814/http://lcweb2.loc.gov/cocoon/minerva/html/minerva-home.html>

**Digital Collections Management & Services Division**

# 25 Years of Web Archiving



**Library of Congress Web Archives**  
<http://lcweb2.loc.gov/cocoon/minerva/html/sept11/sept11-about.html>  
 Archived: 12/07/2006 at 03:20:22

[Ask a Librarian](#)

Note: External links, streaming audio/video, forms and search boxes may not function within this web archive  
[Back to Calendar View](#)

◀ First (12/07/2006) ◀ Previous #1 of 47 [Next](#) ▶ Last (11/21/2021) ▶

The Library of Congress >> More Online Collections

## MINERVA

Mapping the INternet Electronic Resources Virtual Archive

[about: collection overview](#)

### September 11 Web Archive

BROWSE
ABOUT

Overview Selection Criteria Metadata Technical Architecture Copyright FAQs Partners

**Overview**

The Library of Congress, in partnership with the [Internet Archive](#), [WebArchivist.org](#) and the [Pew Internet & American Life Project](#), has created a collection of digital materials known as the September 11 Web Archive.

The September 11 Web Archive preserves the web expressions of individuals, groups, the press and institutions in the United States and from around the world in the aftermath of the attacks in the United States on September 11, 2001.


The Web Archive is important because it contributes to the historical record, capturing information that could otherwise be lost. With the growing role of the Web as an influential medium, records of historic events could be considered incomplete without materials that were "born digital" and never printed on paper.

The September 11 Web Archive consists of over 30,000 selected Web sites archived from September 11, 2001 through December 1, 2001.

Approximately 2,300 Web sites were identified for further processing and were cataloged using [MODS \(Metadata Object Description Schema\)](#), an XML schema for a bibliographic element set which enables the creation of original resource description records.

The collection uses the Wayback Machine interface, a display designed to display Web sites captured over time, which was pioneered by the [Internet Archive](#). Web sites in the collection can be discovered through browsable and searchable interfaces. Please review the [Technical Architecture](#) for more information on these interfaces.

>> [Search the September 11 Web Archive](#)




Library of Congress > Digital Collections > September 11, 2001 Web Archive > About this Collection [Share](#)


**COLLECTION**  
**September 11, 2001 Web Archive**

About this Collection Collection Items


**Featured Content**




WarTimeLiberty.com




You Will Never Be Forgotten  
-- A Tribute -- September 11, 2001



September 11th, Ongoing Updates -- George Mason University



The Avalon Project: September 11, 2001 Terrorist Attacks



American Red Cross

**About this Collection**

Rights & Access

**Expert Resources**

[Ask the Recommending Officer: The September 11, 2001 Web Archive](#)  
[The September 11, 2001 Web Archive: Twenty Years Later](#)  
[Main Reading Room](#)  
[Web Archiving Program Information](#)  
[Tips on Searching the Web Archive](#)

**Frequency of Collection:** Sites in the collection were targeted for capture weekly.

**Languages:** Collection material in English, with Spanish, German, Portuguese, French, Japanese, Flemish, Arabic, Dutch, Slovak, Thai, Italian, Chinese, Hebrew, Polish, Hungarian, Swedish, Romanian/Moldavian/Moldovan, Norwegian, Czech, Hindi, Danish, Malay, Armenian, Finnish, Malayalam, Icelandic, Indonesian, Russian, Albanian, and Greek.

**Capture Date:** December 7, 2006

<https://webarchive.loc.gov/all/20061207032022/http://lcweb2.loc.gov/cocoon/minerva/html/sept11/sept11-about.html>

\* Users accessed captures through the Internet Archive Wayback Machine \*

**Live Web:**

<https://www.loc.gov/collections/september-11-2001-web-archive/about-this-collection/>

• Users access captures through local replay \*

Digital Collections Management & Services Division

# 25 Years of Web Archiving

**LIBRARY**  
LIBRARY OF CONGRESS

Library of Congress Web Archives  
http://www.loc.gov/webcapture/  
Archived: 06/01/2006 at 03:01:17

[Ask a Librarian](#)

Note: External links, streaming audio/video, forms and search boxes may not function within this web archive  
[Back to Calendar View](#)

◀ First (06/01/2006) ◀ Previous #1 of 247 [Next](#) ▶ [Last \(10/14/2021\)](#) ▶

The LIBRARY of CONGRESS WEB CAPTURE

[More Search Options](#)

The Library of Congress > Web Capture

**Web Capture**

- Web Capture Home
- Technical Background
- Partners
- Projects
- FAQs
- Contact

**Related Resources at the Library**

- Digital Preservation
- Digital Collections & Programs
- MINERVA
- U.S. Copyright Office

The Library of Congress preserves the nation's cultural artifacts and provides enduring access to them. The Library's traditional functions of acquiring, cataloging, preserving and serving collection materials of historical importance to the Congress and the American people to foster education and scholarship extend to digital materials, including Web sites.

In 2000, the Library of Congress established a pilot project to collect and preserve these primary source materials. A multidisciplinary team of Library staff representing cataloging, legal, public services, and technology services studied methods to evaluate, select, collect, catalog, provide access to, and preserve these materials for future generations of researchers. The Library developed thematic Web archives on such topics as the United States National Elections of 2000, 2002, and 2004, the Iraq War, and the events of [September 11](#). More about these collections can be found at the [MINERVA Web Preservation project](#) Web site.

In July 2003, the Library and the national libraries of Australia, Canada, Denmark, Finland, France, Iceland, Italy, Norway, Sweden, the British Library (UK), and the Internet Archive (USA) acknowledged the importance of international collaboration for preserving Internet content for future generations and formed the [International Internet Preservation Consortium](#). The goals of the Consortium include collecting a rich body of Internet content from around the world and fostering the development and use of common tools, techniques and standards that enable the creation of international archives.

In 2004, the Library's Office of Strategic Initiatives created a Web Capture team to support the goal of managing and sustaining at-risk digital content. The team is charged with building a Library-wide understanding and technical infrastructure for capturing Web content. The team, in collaboration with a variety of Library staff, and national and international partners, is identifying policy issues, establishing best practices and building tools to collect and preserve Web content.

[Site Map](#) | [Contact](#) | [Accessibility](#) | [Legal](#)

**LIBRARY**  
LIBRARY OF CONGRESS

Everything

Library of Congress > Programs > Web Archiving > Web Archiving overview

PROGRAM  
**Web Archiving**

[Web Archiving overview](#) [For Researchers](#) [For Site Owners](#) [Web Archives](#)

**Web Archiving overview**

[Web Archiving Collaborations](#)

[Frequently Asked Questions](#)

[Glossary](#)

[Contact Us](#)

**Access the Library of Congress Web Archive Collections**  
Browse and search the Library's event and thematic web archives  
[Learn more](#)

**Contact Us**  
Comments, questions, and suggestions related to Web Archiving and this website can be sent to us online.

**Location**  
Web Archiving Program  
Library of Congress

The Library of Congress Web Archive manages, preserves, and provides access to archived web content selected by subject experts from across the Library.

**Capture Date:** June 1, 2006

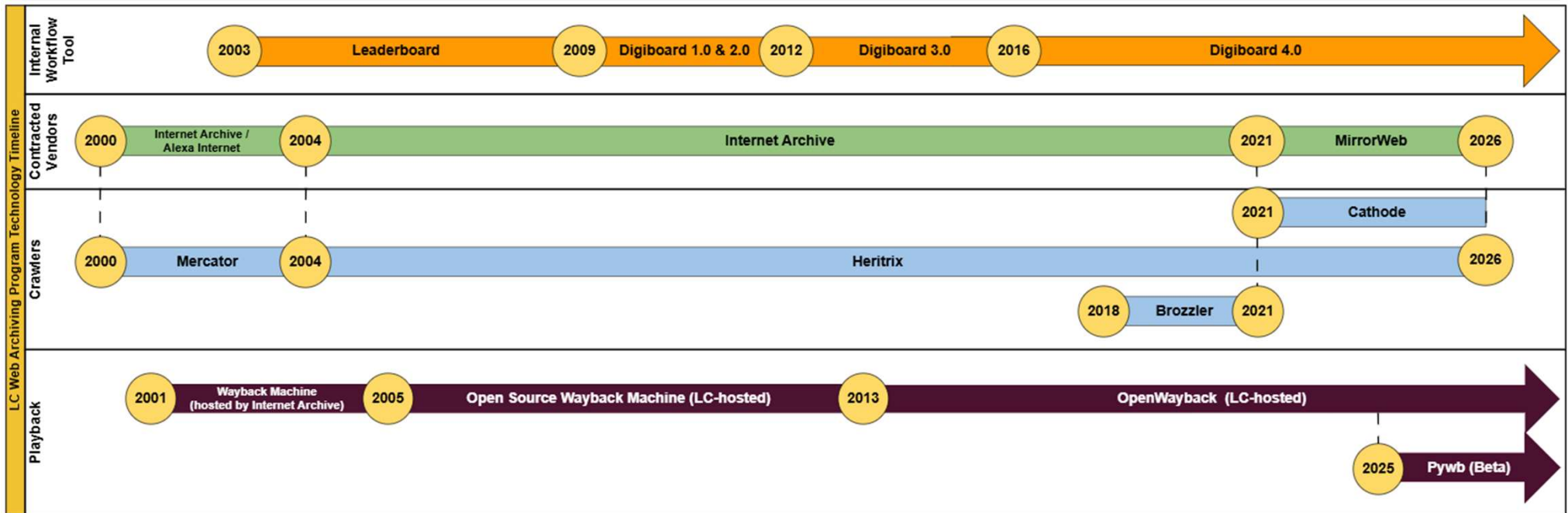
<https://webarchive.loc.gov/all/20060601030117/http://www.loc.gov/webcapture/>

**Live Web:** February 11, 2026

<https://www.loc.gov/programs/web-archiving/about-this-program/>

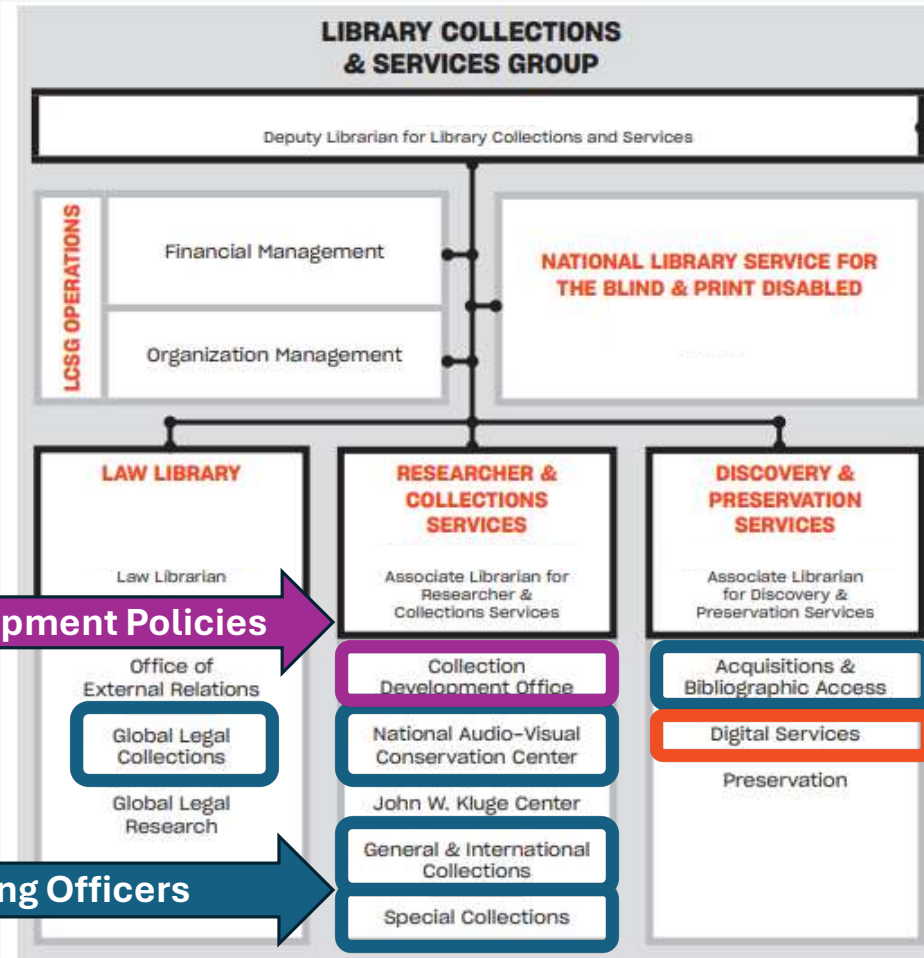
Digital Collections Management & Services Division

# 25 Years of Web Archiving



Digital Collections Management & Services Division

# Where We Sit Organizationally



Collection Development Policies

Recommending Officers

OCIO/IT  
Office of General Counsel

Web Archiving Section

Contracted Harvesting Vendor

Digital Collections Management & Services Division

# Web Archiving Section Duties

**Set requirements** for transfer, replay, storage, and access; co-manage online releases

**Inform** policy from a technical standpoint

**Manage** web crawling contract

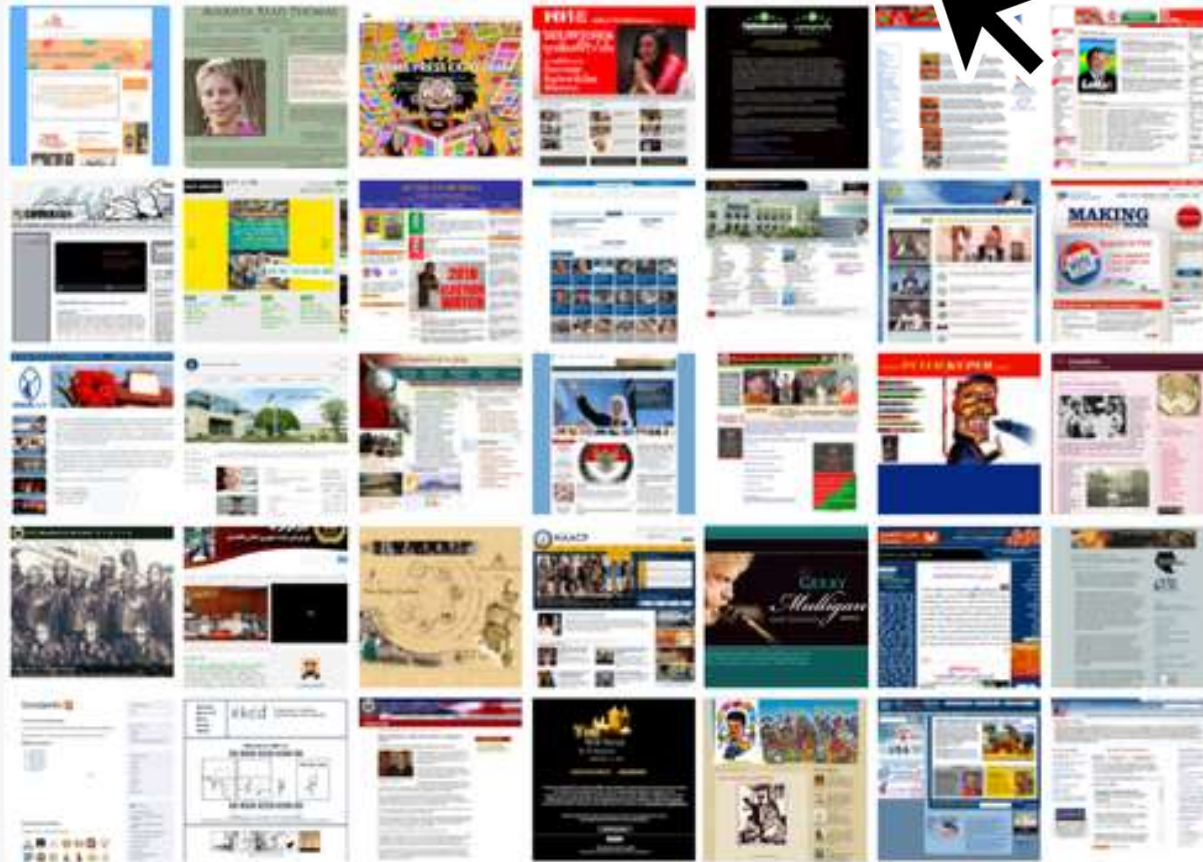
**Perform** quality control to ensure content is archived to the best of our technical capabilities

**Support** curatorial/selecting staff through documentation, trainings, and [webcapture@loc.gov](mailto:webcapture@loc.gov)

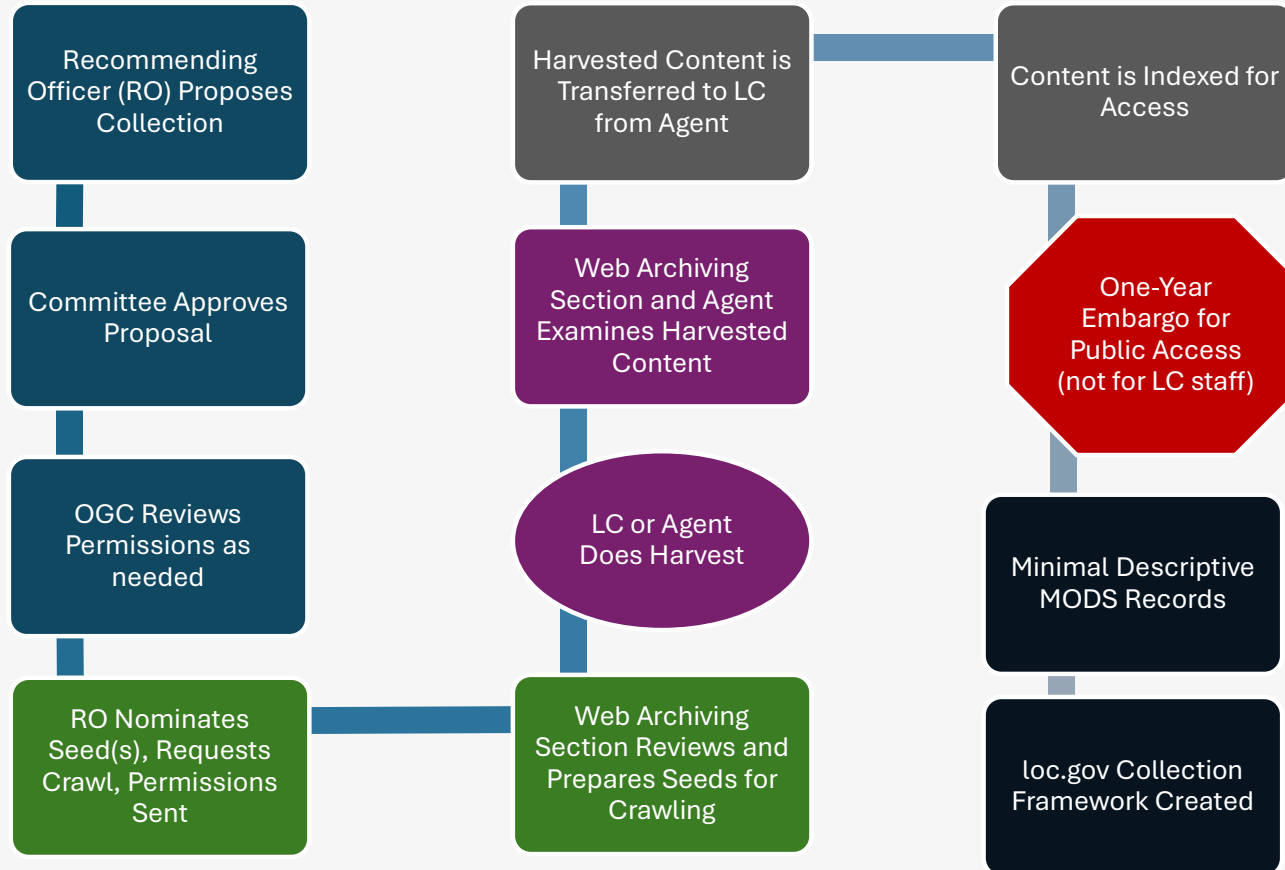
**Promote** the web archiving program & engage with international web archiving community

# Web Archive Collections

 <https://www.loc.gov/web-archives/collections/>



# Overview of Our Web Crawling Process



# Web Archive Collections

## Comprehensive Collecting

**COLLECTION**  
United States Elections Web Archive

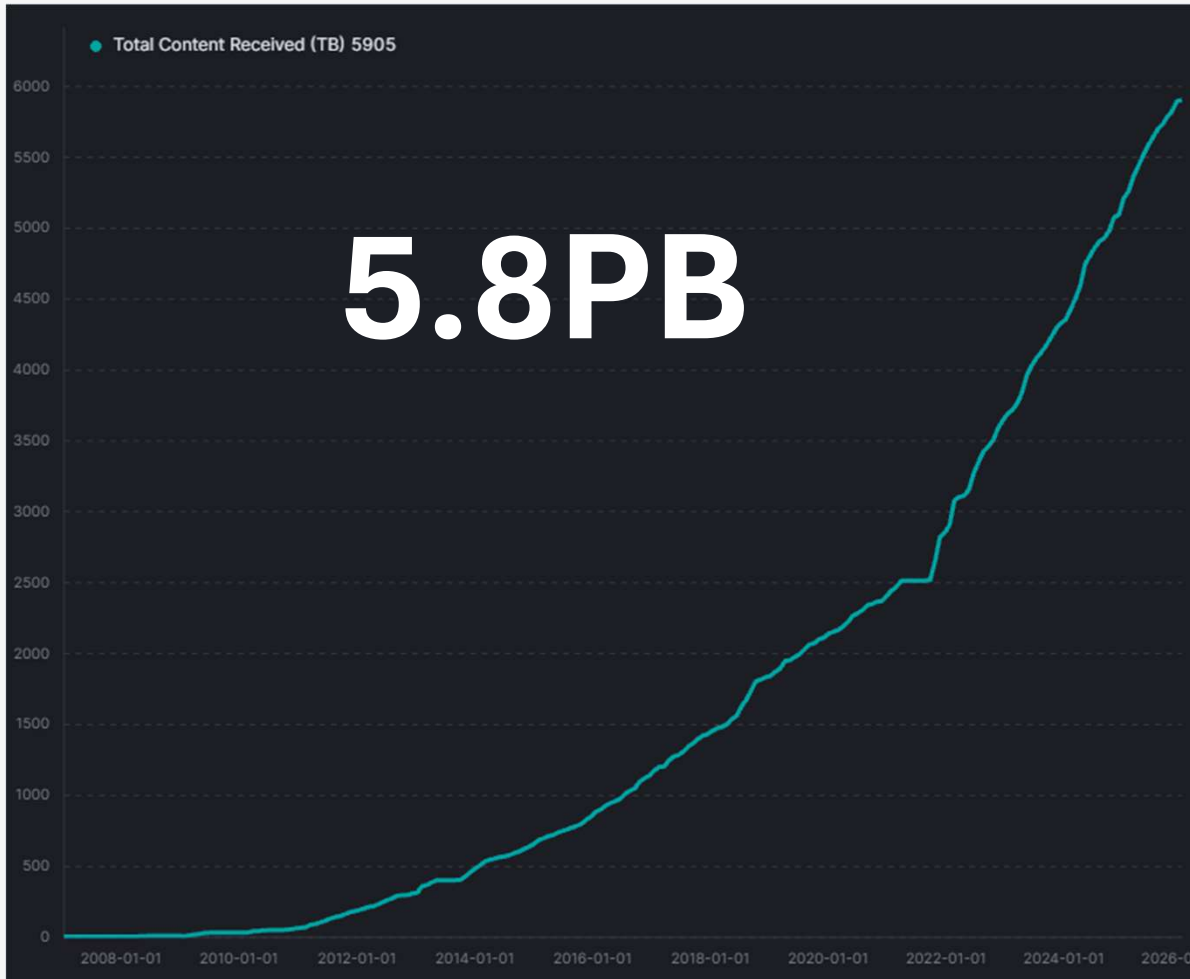
**COLLECTION**  
United States Congressional Web Archive

**COLLECTION**  
Legislative Branch Web Archive

## Representative Collecting

**COLLECTION**  
World War I Centennial Web Archive

**COLLECTION**  
Brazil Cordel Literature Web Archive



## How big is the web archive?

204 Total Collections

83 Active Collections

16,847 Records in Crawl

217 Countries

138 Languages

# Web Archive Collections as Data

**LIBRARY OF CONGRESS**

Library of Congress > Programs > Web Archiving > For Researchers > Web Archive Datasets

**PROGRAM**  
Web Archiving

Web Archiving overview | For Researchers | For Site Owners | Web Archives

For Researchers

## Web Archive Datasets

The Library of Congress makes web archive datasets available for researchers to explore, visualize, and re-use the web archives. Datasets are made available either as experimental data packages or as permanent collection datasets. See more info below.

### Data Packages with LC Labs

Data packages are datasets packaged together with robust documentation and how-to information. All Library of Congress data packages are available at <https://data.labs.loc.gov/packages/> as LC Labs experiments.

United States Elections, Web Archives Data Package

### Permanent Collection Datasets

Datasets that have been added to the Library of Congress permanent collections can be found listed in the Selected Datasets collection, such as the Dinosaur comics, Meme generator, and Giphy datasets that contain CSV metadata for images from the Web Cultures Web Archive. The Dinosaur comics dataset also includes the original images.

See all >

**COLLECTION**  
**Dinosaur comics**

This dataset was generated from content harvested from the Library of Congress's web archive of [qwantz.com](https://www.loc.gov/item/lcwaN0009953/) (Dinosaur Comics!). It includes minimal metadata about 3,325 image objects from the Dinosaur Comics! web archive as well as the files themselves. This dataset was created as apart of exploratory work done by the Library of Congress's Web Archiving Team. README file This dataset includes: lcwa\_dinosaurcomics\_image\_data.zip compressed...

**Contributor:** Library of Congress Web Archiving Program - North, Ryan  
**Date:** 2019

**COLLECTION**  
**Meme Generator : collected datasets**

Meme Generator dataset

Meme Generator allows users to create and share image macros (featuring a picture, or artwork, superimposed with text) in the style of popular internet memes. The site also serves as a searchable collection of user-created images. Meme images/image macros are widely used images that include text written to a range of templates used in online communication.

**Contributor:** Library of Congress - Library of Congress Web Archiving Program - American Folklife Center  
**Date:** 2010

**SOFTWARE, E-RESOURCE**  
**1000 .gov image dataset**

Dot gov image dataset | One thousand dot gov image dataset

"This dataset of 1,000 images was generated from indexes of the Web archives, which were used to derive a random list of 1,000 items identified as image files and hosted on .gov domains. The set includes 1,000 unique image files (primarily with GIF, JPG, PNG, and TIFF extensions) and minimal associated metadata, including links to their locations within the Library's web archive."-- Web archive...

**Contributor:** Library of Congress Web Archiving Program  
**Date:** 2018

<https://www.loc.gov/programs/web-archiving/for-researchers/web-archive-datasets/>

Digital Collections Management & Services Division

# The Web Archiving Community



Digital Preservation Coalition

iPRES

NDSA



DLF

MARAC  
Mid-Atlantic Regional Archives Conference

*asis&t*

# Learn More About Web Archiving

Organization	Resource
Digital Preservation Coalition (DPC)	<p>“Web-archiving” in the Digital Preservation Handbook  <a href="https://www.dpconline.org/handbook/content-specific-preservation/web-archiving">https://www.dpconline.org/handbook/content-specific-preservation/web-archiving</a></p>
	<p>“Preserving the Web” (Digital Preservation Topical Note 10)  <a href="https://www.dpconline.org/docs/dpc-technology-watch-publications/topical-notes-series/1861-dp-note-10-preserving-the-web/file">https://www.dpconline.org/docs/dpc-technology-watch-publications/topical-notes-series/1861-dp-note-10-preserving-the-web/file</a></p>
	DPC Web Archiving and Preservation Working Group (WAPWG)
International Internet Preservation Consortium (IIPC)	<p>IIPC Training Materials (Created with DPC)  <a href="https://netpreserve.org/web-archiving/training-materials/">https://netpreserve.org/web-archiving/training-materials/</a></p>
	IIPC Blog: <a href="https://netpreserveblog.wordpress.com/">https://netpreserveblog.wordpress.com/</a>
	<p>Past IIPC Conference Presentations:  <a href="https://digital.library.unt.edu/explore/collections/IIPCM/">https://digital.library.unt.edu/explore/collections/IIPCM/</a></p>
Library of Congress (LC)	<p>The Signal Blog (“web archiving” category)  <a href="https://blogs.loc.gov/thesignal/category/web-archiving-2/">https://blogs.loc.gov/thesignal/category/web-archiving-2/</a></p>
	<p>Web Archiving Program  <a href="https://www.loc.gov/programs/web-archiving/about-this-program/">https://www.loc.gov/programs/web-archiving/about-this-program/</a></p>
Society of American Archivists (SAA)	<p>SAA Web Archiving Section Blog  <a href="https://webarchivingrt.wordpress.com/">https://webarchivingrt.wordpress.com/</a></p>
National Digital Stewardship Alliance (NDSA)	<p>NDSA Web Archiving Surveys  <a href="https://osf.io/4ytb2/overview">https://osf.io/4ytb2/overview</a></p>














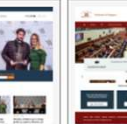
























# Learn More About Tools & Formats

Organization	Resource
Library of Congress (LC)	“WARC, Web ARChive file format” <a href="https://www.loc.gov/preservation/digital/formats/fdd/fdd000236.shtml">https://www.loc.gov/preservation/digital/formats/fdd/fdd000236.shtml</a>
Community-Owned Digital Preservation Tool Registry (COPTR)	“Web” <a href="https://coptr.digipres.org/Web">https://coptr.digipres.org/Web</a>
	“WARC” <a href="https://coptr.digipres.org/WARC">https://coptr.digipres.org/WARC</a>
International Internet Preservation Consortium (IIPC)	“Tools & Software” <a href="https://netpreserve.org/web-archiving/tools-and-software/">https://netpreserve.org/web-archiving/tools-and-software/</a>



# Use Collections

## Library of Congress Web Archive Collections

 <p><b>COLLECTION</b> Afghanistan, Iran, Pakistan, and Tajikistan Elections Web Archive The Afghanistan, Iran, Pakistan and Tajikistan Elections Web Archive includes language sites archived during these countries' elections since 2014. This archive collection includes websites for government leaders, political parties, presidential candidates...</p> <p>Collection Items: View 183 Items</p>	 <p><b>COLLECTION</b> Afghanistan, Iran, Pakistan and Tajikistan Government Web Archive This collection consists of official government websites for Afghanistan, Iran, Pakistan and Tajikistan and provides an overview of the political, economic, administrative and social situation in these four countries. Several of them...</p> <p>Collection Items: View 437 Items</p>	 <p><b>COLLECTION</b> Afghanistan Web Archive The Afghanistan Web Archive is composed primarily of websites from key Afghan government agencies, specifically the executive branch. Also included are a selection of sites from economic reporting agencies, banking institutions, universities...</p> <p>Collection Items: View 22 Items</p>	 <p><b>COLLECTION</b> African Government Web Archive The African Government Web Archive provides links to information from key African government agencies for the 51 countries in Africa south of the Sahara. This web archive ensures that the...</p> <p>Collection Items: View 343 Items</p>	 <p><b>COLLECTION</b> Industry Associations Web Archive This is a collection of select industry association websites. Industry associations produce original content such as reports, statistics, and information on industry trends and advocacy work by associations, which often reflect broader...</p> <p>Collection Items: View 194 Items</p>	 <p><b>COLLECTION</b> Intergovernmental Organizations Web Archive The Intergovernmental Organizations Web Archive contains a selection of websites of international, regional, and multilateral organizations to which member states send official representatives. The collection is not comprehensive but aims to complement...</p> <p>Collection Items: View 43 Items</p>	 <p><b>COLLECTION</b> International Political Events Web Archive The International Political Events Web Archive is an ongoing collaborative collection that preserves websites about developing and spontaneous international events, such as elections, diplomatic meetings, bombings, rebellions, revolutions, political protests and other...</p> <p>Collection Items: View 420 Items</p>	 <p><b>COLLECTION</b> International Tribunals Web Archive The International Tribunals Web Archive contains selected websites on the most important international courts and tribunals created since World War II to adjudicate legal issues of a transnational nature. The collection primarily...</p> <p>Collection Items: View 19 Items</p>	 <p><b>COLLECTION</b> Science Blogs Web Archive The Science Blogs Web Archive captures an element of science communication, predominantly originating in the United States, in the digital age. From microbiology to climate activism, and food science to space exploration...</p> <p>Collection Items: View 94 Items</p>	 <p><b>COLLECTION</b> September 11, 2001 Web Archive The September 11, 2001 Web Archive preserves the web expression of individuals, groups, the press and institutions in the United States and from around the world in the aftermath of the attacks...</p> <p>Collection Items: View 2,312 Items</p>	 <p><b>COLLECTION</b> Serbian Historical and Social Issues Web Archive The Serbian Political and Social Issues Web Archive is composed of sites that address political and social issues and inform public policy. Content includes websites of Serbian politicians, political parties, non-governmental organizations...</p> <p>Collection Items: View 79 Items</p>	 <p><b>COLLECTION</b> Singaporean Elections Web Archive The Singaporean Elections Web Archive brings together a collection of websites that document Singaporean general elections, including government, non-profit, political party, candidate, and news sites. The web archive covers general elections...</p> <p>Collection Items: View 48 Items</p>
 <p><b>COLLECTION</b> American Civil War Sequenential Web Archive The American Civil War Sequenential Web Archive is a collection of websites presenting various graphics and text in commemoration of the 150th anniversary of the American Civil War. This anniversary is historically...</p> <p>Collection Items: View 199 Items</p>	 <p><b>COLLECTION</b> American Folklife Center Web Archive The American Folklife Center Archive, established in the Library of Congress Music Division in 1928 as the first office of American folk song is now one of the largest archives of ethnographic materials...</p> <p>Collection Items: View 23 Items</p>	 <p><b>COLLECTION</b> American Music Creators Web Archive The American Music Creators Web Archive is a collection of websites of musicians, performers, and composers contributing to the American music scene. The collection aims to be representative of all genres of music. This includes mainstream and independent records...</p> <p>Collection Items: View 69 Items</p>	 <p><b>COLLECTION</b> American Music Industry Web Archive The American Music Industry Web Archive captures the activities of individuals and companies that represent music creation and are involved broadly in the business of music. This includes mainstream and independent records...</p> <p>Collection Items: View 48 Items</p>	 <p><b>COLLECTION</b> Iraq War 2003 Web Archive On March 28, 2003, the United States began military action against Iraq for the stated purpose of removing Iraqi leader Saddam Hussein and preventing the use of suspected nuclear weapons (weapons of...</p> <p>Collection Items: View 312 Items</p>	 <p><b>COLLECTION</b> January 6th Attack on the U.S. Capitol Web Archive On January 6, 2021, the U.S. Capitol was attacked by a crowd attempting to stop a joint session of Congress from certifying the electoral votes of the 2020 presidential election. The January...</p> <p>Collection Items: View 137 Items</p>	 <p><b>COLLECTION</b> Kashmir Legal Status 2019 Web Archive In August 2019, the special legal status of Kashmir defined in section 370 of the Indian Penal Code was rescinded. This web archive collection covers a range of local and regional responses...</p> <p>Collection Items: View 39 Items</p>	 <p><b>COLLECTION</b> Kosovo Political and Social Issues Web Archive The Kosovo Political and Social Issues Web Archive is composed of sites that address political and social issues and inform public policy. Content includes websites of Kosovo political parties, non-governmental organizations, voluntary...</p> <p>Collection Items: View 30 Items</p>	 <p><b>COLLECTION</b> Small Press Expo Comic and Comic Art Web Archive Collection Period: August 9, 2011 to ongoing. The Small Press Expo (SPE) is a registered 501(c)(3) nonprofit created in 1984 to promote artists and publishers who produce independent comics. SPE hosts an...</p> <p>Collection Items: View 35 Items</p>	 <p><b>COLLECTION</b> South Pacific Islands Web Archive The South Pacific Islands Web Archive covers the following twenty island/nation/territories: American Samoa, Cook Islands, Fiji, French Polynesia, Guam, Kiribati, Marshall Islands, Federated States of Micronesia, Nauru, New Caledonia, Niue, Northern Mariana...</p> <p>Collection Items: View 172 Items</p>	 <p><b>COLLECTION</b> Southeast Asia Banking and Finance Web Archive The Southeast Asia Banking and Finance Web Archive focuses on the role of the government in the banking and financial sectors in Southeast Asia in much of the region. Governments play a...</p> <p>Collection Items: View 85 Items</p>	 <p><b>COLLECTION</b> Southeast Asia Defense and Military Web Archive This collection brings together the primarily state-operated and national military websites from across the Southeast Asian region. At the top level, there are websites from the various national Department of Military Affairs...</p> <p>Collection Items: View 75 Items</p>
 <p><b>COLLECTION</b> Art and Design Web Archive The Art and Design Web Archive focuses on art and design-related websites that promote, complement and contextualize the Library's rich holdings of materials related to the field of the and decorative arts...</p> <p>Collection Items: View 22 Items</p>	 <p><b>COLLECTION</b> Author Websites Web Archive The Author Websites Web Archive is a collection of official author websites that promote, complement and contextualize the Library's rich holdings of materials related to the field of the and decorative arts...</p> <p>Collection Items: View 83 Items</p>	 <p><b>COLLECTION</b> Azerbaijan, Kazakhstan, Kyrgyzstan, Turkmenistan, and Uzbekistan Government Web Archive The countries of Central Asia represent a key geographic area of strategic interest for the United States and other world powers, including China, the European Union, India, Iran, Russia and Turkey...</p> <p>Collection Items: View 376 Items</p>	 <p><b>COLLECTION</b> Bangladesh Elections Web Archive This collection of websites documents the Bangladesh General Elections beginning with the 2018 2019 election cycle, up to the present. The websites in this collection cover the political parties taking part in the...</p> <p>Collection Items: View 28 Items</p>	 <p><b>COLLECTION</b> Lebanon Elections Web Archive This Lebanon Elections Web Archive covers elections for the country of Lebanon, and currently includes archived websites on the parliamentary and presidential elections of 2011...</p> <p>Collection Items: View 17 Items</p>	 <p><b>COLLECTION</b> LC Commissioned Composers Web Archive The Music Division has commissioned over 400 composers since 1925, and continues to do so annually. World premiere performances of these commissions usually take place on the Library of Congress Concert Series...</p> <p>Collection Items: View 288 Items</p>	 <p><b>COLLECTION</b> Legal Blogs Web Archive This is a selective collection of authoritative sites (associated with American Bar Association approved law schools, research institutes, think tanks, subject matter experts, and expertise based organizations) that contain useful information on specific topics or...</p> <p>Collection Items: View 281 Items</p>	 <p><b>COLLECTION</b> Legislative Branch Web Archive The Legislative Branch Web Archive includes a variety of websites created by Congressional groups and agencies that fall within the legislative branch. This includes websites created by Congress for specific topics or...</p> <p>Collection Items: View 102 Items</p>	 <p><b>COLLECTION</b> Sri Lankan Elections Web Archive The Sri Lankan Elections Web Archive covers the presidential and general elections in 2015, 2015, and 2019/2020...</p> <p>Collection Items: View 81 Items</p>	 <p><b>COLLECTION</b> State Government Websites of the United States Web Archive The State Government Websites Web Archive contains websites of state government agencies and entities from all 50 states, U.S. territories, and DC. The collection seeks to represent U.S. state-level...</p> <p>Collection Items: View 3,158 Items</p>	 <p><b>COLLECTION</b> The Elections Web Archive The The Elections Web Archive covers the general elections of 2011, 2014, and 2018...</p> <p>Collection Items: View 181 Items</p>	 <p><b>COLLECTION</b> The Different Faces of Nationalism in Pakistan Web Archive The Different Faces of Nationalism in Pakistan Web Archive collects websites and blogs of Pakistani organizations, individuals, and groups who engage with the political scene in Pakistan, and the language of nationalism...</p> <p>Collection Items: View 39 Items</p>

<https://www.loc.gov/web-archives/collections/>

Digital Collections Management & Services Division

# Use Collections

## IIPC Collaborative Collections



### International Internet Preservation Consortium

Archive-It Partner Since: Dec, 2013

Organization Type: [National Institutions](#)

Organization URL: <http://netpreserve.org/>

Description: The International Internet Preservation Consortium (IIPC) is a membership organization dedicated to improving the tools, standards and best practices of web archiving while promoting international collaboration and the broad access and use of web archives for research and cultural heritage. The Content Development Group leads the IIPC's [collaborative collection projects](#).

#### Street Art

**Archived since:** Jan, 2023

**Description:** A collection created by the Content Development Group of the International Internet Preservation Consortium in collaboration with Archive-It to preserve web content related to Street Art. The types of street art covered by this collection include but are not limited to mural art, graffiti, stencil art, fly-posting (gluing posters), stickering, yarn-bombing, mosaic. The collection will also include a number of different types of websites such as image collections on databases, blogs, artists' websites, forums, non-profit organizations sites, social media, governmental, institutional, and commercial.

**Subject:** Arts & Humanities, Society & Culture, Street art, Graffiti

**Creator:** International Internet Preservation Consortium

#### War in Ukraine

**Archived since:** Jul, 2022

**Description:** On 23 February 2022, the armed forces of the Russian Federation invaded the territory of Ukraine, annexing certain regions and cities and carrying out a series of military strikes throughout the country, thus triggering a war in Ukraine. Since then, the clashes between the Russian military and the Ukrainian army and population have had an unprecedented impact on the situation in the region and on international relations. This collection aims to collect web content related to this event in order to map the impact of this conflict on digital history and culture. Identification of seed websites and initial web crawling began in July 2022, and the collection will continue to add new content as the conflict evolves. High priority subtopics include: general information on military confrontations; consequences of the war on the civilian population; refugee crisis and international relief efforts; political consequences; international relations; diaspora communities—Ukrainians around the world; human rights organisations; foreign embassies and diplomatic relations; sanctions imposed on Russia by foreign powers; consequences on energy and agri-food trade; and public opinion (blogs/protest sites/activists). Websites from all over the world and in all languages are in scope.

**Subject:** Spontaneous Events, Russian Invasion of Ukraine, 2022, Wars

**Creator:** International Internet Preservation Consortium

**Coverage:** Ukraine

**Collector:** International Internet Preservation Consortium

#### Intergovernmental Organizations

**Archived since:** Apr, 2015

**Description:** A collection of the official websites of intergovernmental organizations, created by the International Internet Preservation Consortium's Content Development Group beginning in 2015. The initial core of the collection consisted in a comprehensive list of active websites from the .int top-level domain (available only to organizations created by treaties). In 2020 the collections was expanded to include the websites of a wider selection of intergovernmental organizations from other top-level domains, prioritizing organizations appearing on the UN's list of IGO's with observer status, Wikipedia's list of Intergovernmental Organizations, and the subset of IGOs from UIA Yearbook of International Organizations designated as FIGOs by Thomas J. Volgy et al in their 2008 article "Identifying Formal Intergovernmental Organizations."

**Subject:** Government, Science & Health, Society & Culture, International agencies, International cooperation—Societies, etc., Intergovernmental cooperation—Societies, etc.

**Creator:** International Internet Preservation Consortium

**Collector:** International Internet Preservation Consortium

#### National Olympic and Paralympic Committees

**Archived since:** Mar, 2016

**Description:** A collection created by the International Internet Preservation Consortium including the websites of national Olympics and Paralympics committees and associations, as identified from the official listings of these groups found on the official sites [www.olympic.org](http://www.olympic.org) and [www.paralympic.org](http://www.paralympic.org).

**Subject:** Society & Culture, Olympics

**Creator:** International Internet Preservation Consortium

**Collector:** International Internet Preservation Consortium

#### World War I Commemoration

**Archived since:** Oct, 2015

**Description:** A collection of websites related to the First World War, including sites and events marking the centenary of the war and those that study WWI in general. The collection aims to include a range of different types of website, from official commemorations to amateur history websites, and the reporting of the centenary in the media. Websites from several different countries and many languages have been selected by the member libraries of the International Internet Preservation Consortium, to provide a transnational perspective on the war's centennial commemoration.

**Subject:** Society & Culture, World War, 1914-1918, World War, 1914-1918—Centennials, Wars

**Creator:** International Internet Preservation Consortium

**Collector:** International Internet Preservation Consortium

#### World War Two 80th Anniversary Commemoration

**Archived since:** May, 2025

**Description:** The year 2025 is the 80th anniversary of the end of World War II, which concluded in 1945. This year includes specific dates of note that commemorate events that led to the war's end. The 80th anniversary of Victory in Europe Day (VE Day) is 8 May 2025. The 80th anniversary of the atomic bombings of Hiroshima and Nagasaki are 6 and 9 August 2025. The 80th anniversary of Victory over Japan Day/Victory in the Pacific Day (VJ Day/VP Day) is 15 August 2025 in the United Kingdom and 2 September 2025 in the United States. This event-based collaborative collection between IIPC members and the public will include websites and individual web pages that document anniversary events, including physical and online sites of memory, such as memorial ceremonies, veterans' activities, and military cemeteries; physical exhibits and related events, such as those hosted by museums and historical societies; information about commemorative works of literature, visual art, and performing arts; memorial events related to specific battles in the European, Pacific, and North African theaters; and memorial events that commemorate specific events including VE Day, VJ Day, the Holocaust and concentration camp liberations, and the atomic bomb attacks on Hiroshima and Nagasaki. This collection relates thematically to IIPC's 2015 World War I Commemoration Web Archive (<https://archive-it.org/collections/6415>).

**Subject:** Arts & Humanities, Society & Culture, World War, 1939-1945, Holocaust memorials, Military history, Japanese Americans—Forced removal and internment, 1942-1945

**Creator:** International Internet Preservation Consortium

<https://archive-it.org/home/IIPC>

Digital Collections Management & Services Division

# Reflections at 25 Years

**LIBRARY**  
LIBRARY OF CONGRESS

NOTE: EXTERNAL LINKS, STREAMING AUDIO/VIDEO, FORMS AND SEARCH BOXES MAY NOT FUNCTION WITHIN THIS WEB

[BACK TO CALENDAR VIEW](#)

Library of Congress Web Archives  
HTTP://XKCD.COM/1348/  
ARCHIVED: 07/15/2014 AT 22:52:29

◀ FIRST (07/15/2014) ◀ PREVIOUS #1 OF 80 NEXT ▶ LAST (11/17/2022) ▶

[ASK A LIBRARIAN](#)

**ARCHIVE**

**WHAT IF?**

**BLAG**

**STORE**

**ABOUT**

**A WEBCOMIC OF ROMANCE, SARCASM, MATH, AND LANGUAGE.**

NEWS: I'M PUBLISHING A *WHAT IF* BOOK!

**BEFORE THE INTERNET**

< < PREV RANDOM NEXT > >

< < PREV RANDOM NEXT > >

PERMANENT LINK TO THIS COMIC: [HTTP://XKCD.COM/1348/](http://xkcd.com/1348/)  
 IMAGE URL (FOR HOTLINKING/EMBEDDING): [HTTP://IMGS.XKCD.COM/COMICS/BEFORE\\_THE\\_INTERNET.PNG](http://imgs.xkcd.com/comics/before_the_internet.png)

“Before the Internet” | Capture Date: July 15, 2014

Web Archive:

<https://webarchive.loc.gov/all/20140715225229/http://xkcd.com/1348/>

Collection: Webcomics Web Archive

<https://www.loc.gov/collections/webcomics-web-archive/about-this-collection/>

Digital Collections Management & Services Division

**THANK YOU!**



**webcapture@loc.gov**



<https://www.loc.gov/programs/web-archiving/about-this-program/>

**Digital Collections Management & Services Division**