

Web Archiving: How do we Preserve the Web?

Mark Phillips, Ph.D.

University of North Texas

UNT Libraries

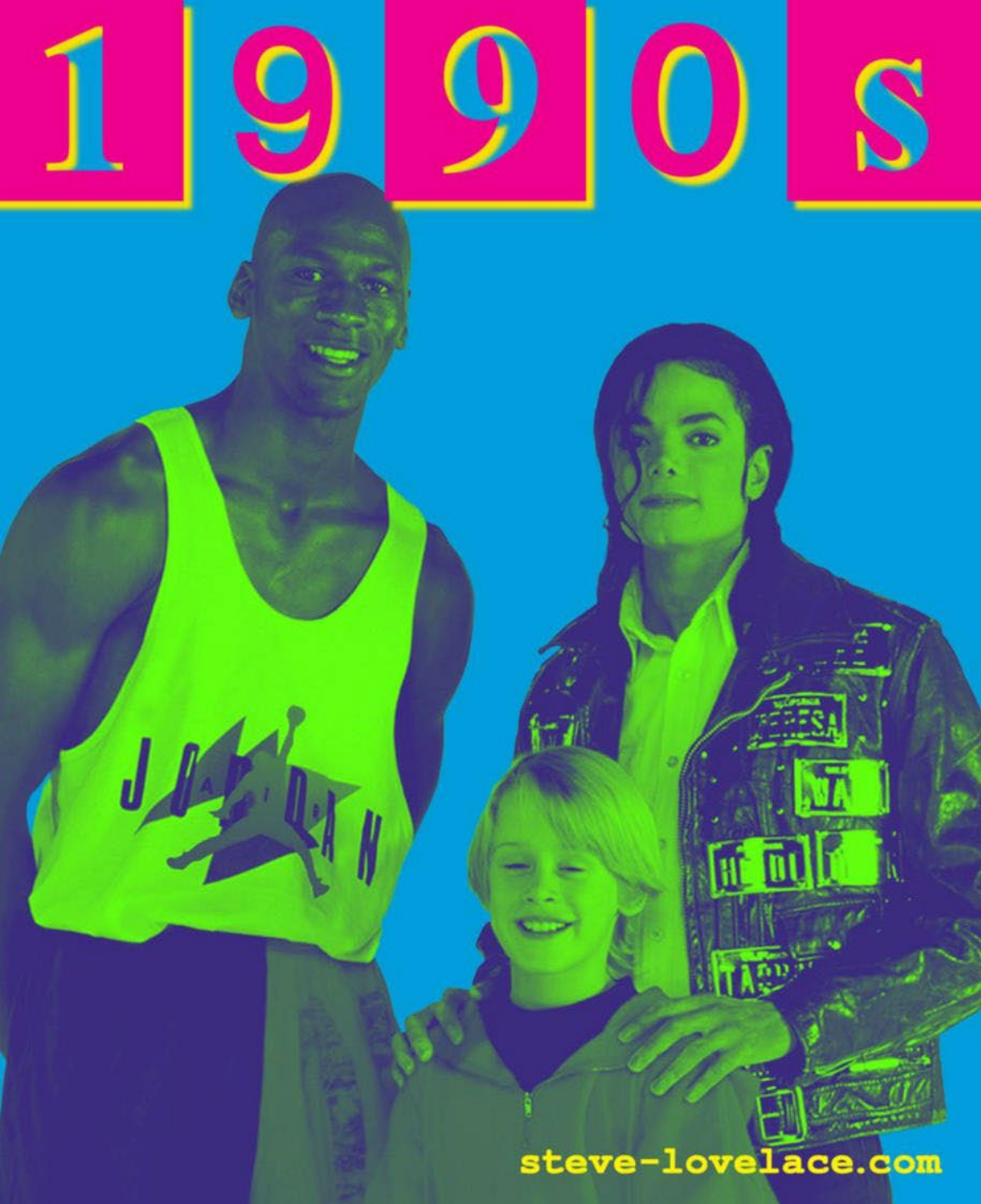
March 9, 2022

The Web is the platform for communication for the
21st century.

Each year more and more content
is created digitally.

And published to the world via the Web.

Ensuring the long-term preservation and access to these resources offer new challenges and opportunities to memory organizations such as libraries, museums, and archives.



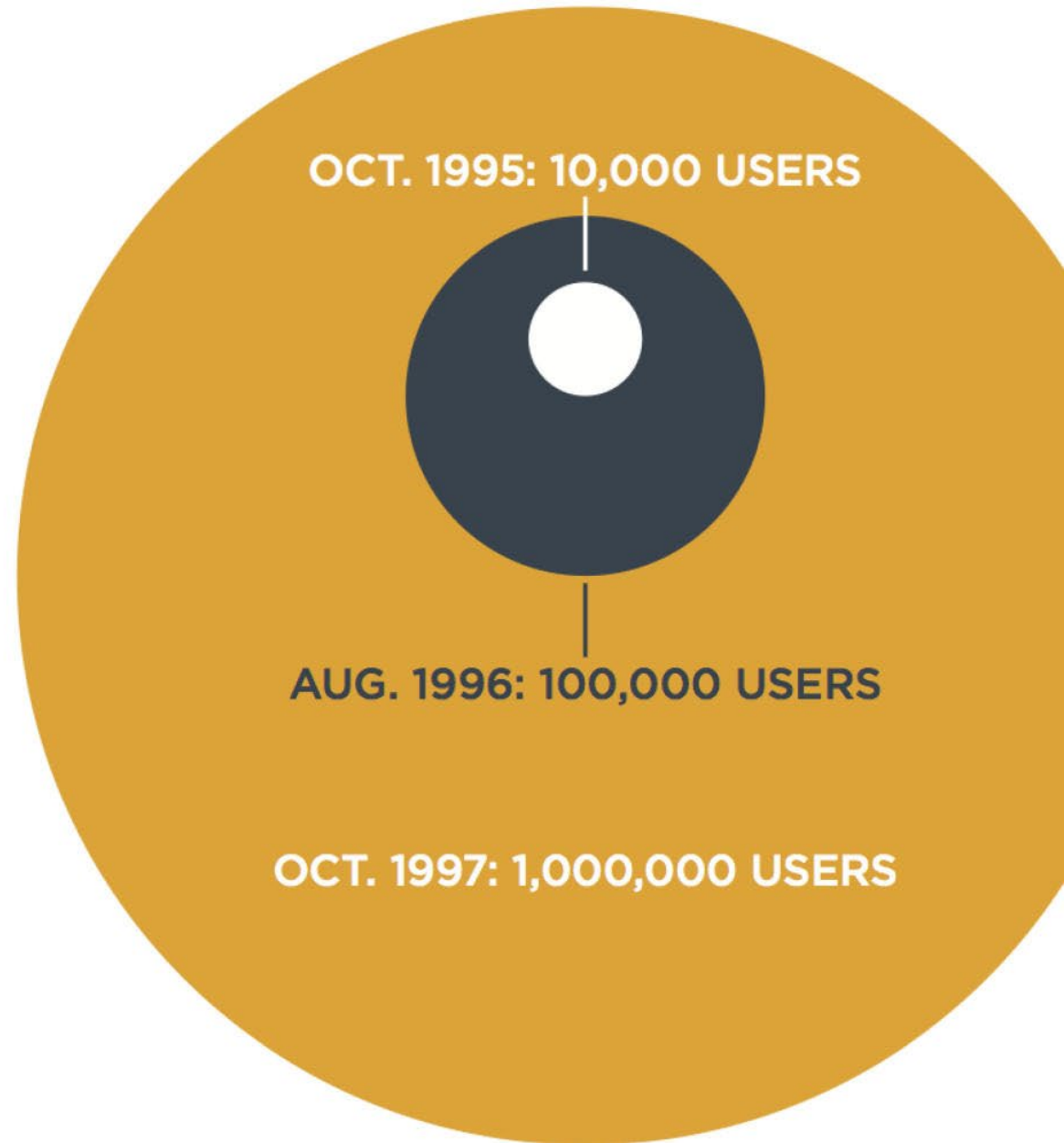
Could one
even study
the 1990s
and beyond
**without
web
archives?**

Ian Milligan
University of Waterloo

GEOCITIES USERS:

**What was
GeoCities?**

**Why does it
matter?**



A Minimalist Approach to Design of Walking Robots

Matt Moses, Robotic Vehicle Range
Sandia National Laboratories¹, P.O. Box 5800, Albuquerque, NM 87185-1125

RECEIVED
SEP 01 2000
OSTI

Abstract

A minimalist approach to the design of walking robots is presented, in which emphasis is placed on the simplification of control, the use of mechanics to solve problems, and the use of biologically inspired designs. Control is realized by artificial

References

1. Tilden, M. W., 1994, Adaptive Robotic Nervous Systems and Control Circuits Therefor. U.S. Patent # 5,325,031.
2. Haslacher, B., and Tilden, M. W., 1995, Living Machines. Robotics and Autonomous Systems: The Biology and Technology of Intelligent Autonomous Agents, L. Steels, Ed., Elsevier Publishers.
3. <http://www.geocities.com/SouthBeach/6897/beam2.html>
4. Moses, M., 2000, A Minimalist Approach to Design of Walking Robots, Sandia Report SAND2000-0095, Sandia National Laboratories.

YAHOO!

The requested URL was not found on this server.

Please check the URL for proper spelling and capitalization. If you're having trouble locating a destination on Yahoo, try visiting the [Yahoo home page](#) or look through a list of [Yahoo's online services](#). Also, you may find what you're looking for if you try searching below.

 · [advanced search](#)

Please try [Yahoo Help Central](#) if you need more assistance.

Copyright © 2019 Yahoo Inc. All rights reserved. [Privacy Policy](#) - [Terms of Service](#)

<https://digital.library.unt.edu/ark:/67531/metadc722979/>

Search the history of over 387 billion web pages on the Internet.

WayBackMachine



Internet Archive is a non-profit library of millions of free books, movies, software, music, websites, and more.

387B 22M 5.2M 7.6M 1.9M 474K 3.4M 209K 573K

GO

[Advanced Search](#)

Announcements

Unlocking the Potential for Every High School Library: 2019 Internet Archive Hero Award

How the Internet Archive is Digitizing LPs to Preserve Generations of Audio

Everyone Deserves to Learn

[SEE MORE](#)

Top Collections at the Archive

<p>Community Audio</p> <p>2,061,688 ITEMS</p>	<p>Community Video</p> <p>846,974 ITEMS</p>	<p>Community Texts</p> <p>877,057 ITEMS</p>	<p>American Libraries</p> <p>3,061,253 ITEMS</p>	<p>Community Data</p> <p>249,510 ITEMS</p>
<p>The LibriVox Free Audiobook</p>	<p>Electric Sheep</p>	<p>Canadian Libraries</p>	<p>Live Music Archive</p>	<p>University of Toronto - Roberts</p>

[Terms of Service](#) (last updated 12/31/2014)

INTERNET ARCHIVE Explore more than 387 billion web pages saved over time

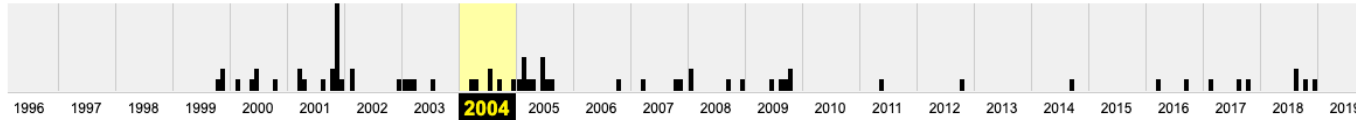
DONATE



http://www.geocities.com/SouthBeach/6897/beam2.html/

Calendar · Collections ^{beta} · Changes ^{beta} · Summary · Site Map

Saved 72 times between October 12, 1999 and December 30, 2018.



Mon, 05 Jul 2004 08:45:44 GMT (why: alexa_dx, alexacrawls)

JAN			FEB							MAR						APR											
			1	2	3	1	2	3	4	5	6	7	1	2	3	4	5	6									
4	5	6	7	8	9	10	8	9	10	11	12	13	14	7	8	9	10	11	12	13	4	5	6	7	8	9	10
11	12	13	14	15	16	17	15	16	17	18	19	20	21	14	15	16	17	18	19	20	11	12	13	14	15	16	17
18	19	20	21	22	23	24	22	23	24	25	26	27	28	21	22	23	24	25	26	27	18	19	20	21	22	23	24
25	26	27	28	29	30	31	29	28	29	30	31	25	26	27	28	29	30										
MAY			JUN							JUL						AUG											
					1	1	2	3	4	5			1	2	3	1	2	3	4	5	6	7					
2	3	4	5	6	7	8	6	7	8	9	10	11	12	4	5	6	7	8	9	10	8	9	10	11	12	13	14
9	10	11	12	13	14	15	13	14	15	16	17	18	19	11	12	13	14	15	16	17	15	16	17	18	19	20	21
16	17	18	19	20	21	22	20	21	22	23	24	25	26	18	19	20	21	22	23	24	22	23	24	25	26	27	28
23	24	25	26	27	28	29	27	28	29	30	25	26	27	28	29	30	31	29	30	31							
30	31																										
SEP			OCT							NOV						DEC											
									1	2	1	2	3	4	5	6				1	2	3	4				
5	6	7	8	9	10	11	3	4	5	6	7	8	9	7	8	9	10	11	12	13	5	6	7	8	9	10	11
12	13	14	15	16	17	18	10	11	12	13	14	15	16	14	15	16	17	18	19	20	12	13	14	15	16	17	18
19	20	21	22	23	24	25	17	18	19	20	21	22	23	21	22	23	24	25	26	27	19	20	21	22	23	24	25
26	27	28	29	30	24	25	26	27	28	29	30	28	29	30	26	27	28	29	30	31							

Note

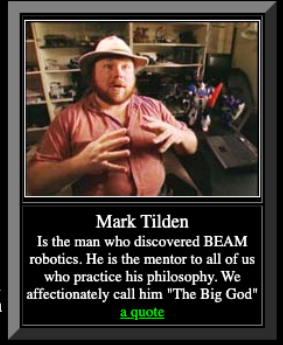
This calendar view maps the number of times <http://www.geocities.com/SouthBeach/6897/beam2.html/> was crawled by the Wayback Machine, *not* how many times the site was actually updated. More info in the [FAQ](#).

BEAM Robotics

BEAM is an acronym for:
Biology- Many aspects of the design and function of BEAM robots parallel nature.
Electronics- Simple control circuitry which allows for emergent behavior.
Aesthetics- Make it pretty and it will work better. My robots are very ugly :).
Mechanics- The mechanics of the BEAM robot's body are just as important as the electronics in it's brain.
These are the four words which BEAMers live by.

BEAM robotics is a new field of robotics. It uses minimalist electronics to create elegant mechanical creatures that parallel their natural counterparts in many ways. The simplicity of BEAM construction allows people from all walks of life and levels of education to create very capable autonomous robots without learning to program or a government research grant. If you have the mechanical skill to straighten out a paperclip until it can roll down an incline plane you can build a beambot. The potential for progressive learning makes BEAM an excellent hobby for young children who can learn basic construction techniques and mechanics, with parental aid of course, and move up to the electronics portion when they are ready.

BEAM devices come in infinite shapes and sizes. Some have wheels, some have legs, some even have wings and there are some which use completely different modes of transportation. The brains used to control BEAM "lifeforms" are very simple compared to the rats-nets of wiring and circuitry in conventional robots. These brains, also called nervous networks, contain no microprocessors with many BEAM critters having less transistors than a common radio. By wiring in basic sensors to influence the nervous network, we can control how the robot behaves much like putting blinkers on a race horse to make it run straight. These sensors include light detectors, touch feelers, heat sensors and just about anything you can think of. Your imagination is the limit.

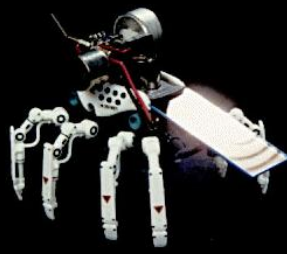


You can buy kits at [Solarbotics](#) the BEAM robotics resource company.
Join the [BEAM Robotics Club](#).
Got a question? Search through the [eGroups E-mail archive](#)
or
read the [BEAM Tek FAQ](#).
If you still can't find the answer [e-mail me](#).
Looking for robotics [BOOKS](#)?

BEAM Robotics News

Got a news tidbit you want posted? [E-mail](#) it to me.

News section is off-line for a while. Sorry. Tune in after the new year for my next contest.



[A Five Motor Advanced Walker](#)
(Still under experimentation)



[How to Build a Photovore](#)

- [August '99 News](#)
- [July '99 News](#)
- [June '99 News](#)
- [May '99 News](#)
- [April '99 News](#)
- [March '99 News](#)
- [February '99 News](#)
- [January '99 News](#)
- [December '98 News](#)
- [November '98 News](#)
- [October '98 News](#)
- [September '98 News](#)
- [August '98 News](#)
- [July '98 News](#)
- [June '98 News](#)



Background:

What is Web Archiving?

Web archiving is the process of collecting portions of the World Wide Web to ensure the information is preserved in an archive for future researchers, historians, and the public.

Web archiving, and web archives have been around since the mid-1990's with the start of the Internet Archive (archive.org) and the Wayback Machine

Since then, hundreds of institutions around the world have started to collect portions of the web for future researchers to use.

It involves concepts that are core to
libraries and archives.

Selection: Defining the scope of the Web for preservation.

National Domain
Federal Domain
State/Municipal Web
Subject Based
Event Based

Acquisition: Harvesting selected content by using Web crawlers such as Heritrix

**Access: Enabling interested users to accessing
Web resources from the past.**

Preservation: Ensuring the long-term access to these Web resources through systematic planning and curation.

These concepts enable memory organizations to align Web archiving with traditional collections, processes and missions

The web is *the* **big** dataset.

How might you use a web archive?

When you are following citations and reach a 404 on the web

- You can use a web archive

If what you are reading about a website doesn't match up with the redesign as it is today.

- You can use a web archive

If you need to interact with a website from ten years ago as it was

- You can use a web archive

If you need content from a long dead company and their website

- You can use a web archive

Let's Tour Some Web Archives

Internet Archive

<https://archive.org>

INTERNET ARCHIVE Explore more than 387 billion web pages saved over time



http://

BROWSE HISTORY

Find the Wayback Machine useful? DONATE



Tools

Wayback Machine Availability API
Build your own tools.

WordPress Broken Link Checker
Banish broken links from your blog.

404 Handler for Webmasters
Help users get where they were going.



Subscription Service

Archive-It enables you to capture, manage and search collections of digital content without any technical expertise or hosting facilities. Visit Archive-It to build and browse the collections.



Save Page Now

http:// SAVE PAGE

Capture a web page as it appears now for use as a trusted citation in the future.
Only available for sites that allow crawlers.

FAQ | Contact Us | Terms of Service (Dec 31, 2014)

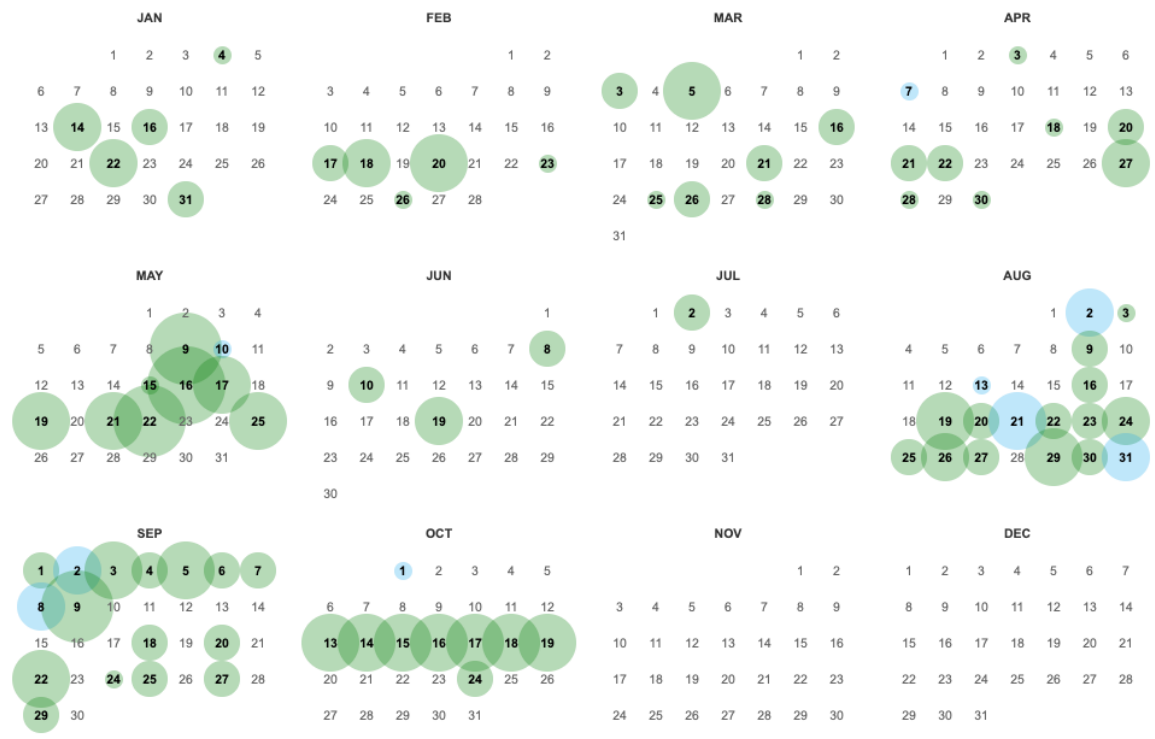
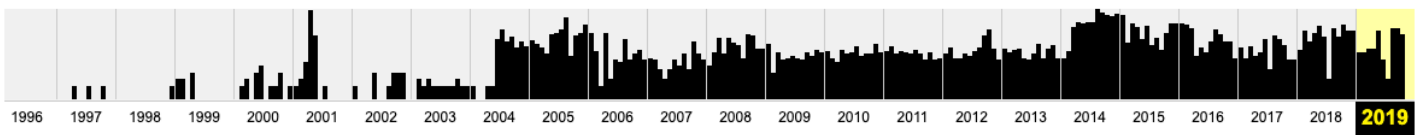
366 billion web pages

INTERNET ARCHIVE Explore more than 387 billion web pages saved over time

DONATE **Wayback Machine**

Calendar · Collections ^{beta} · Changes ^{beta} · Summary · Site Map

Saved **4,253** times between **April 10, 1997** and **October 24, 2019**.



Note
This calendar view maps the number of times **www.unt.edu** was crawled by the Wayback Machine, *not* how many times the

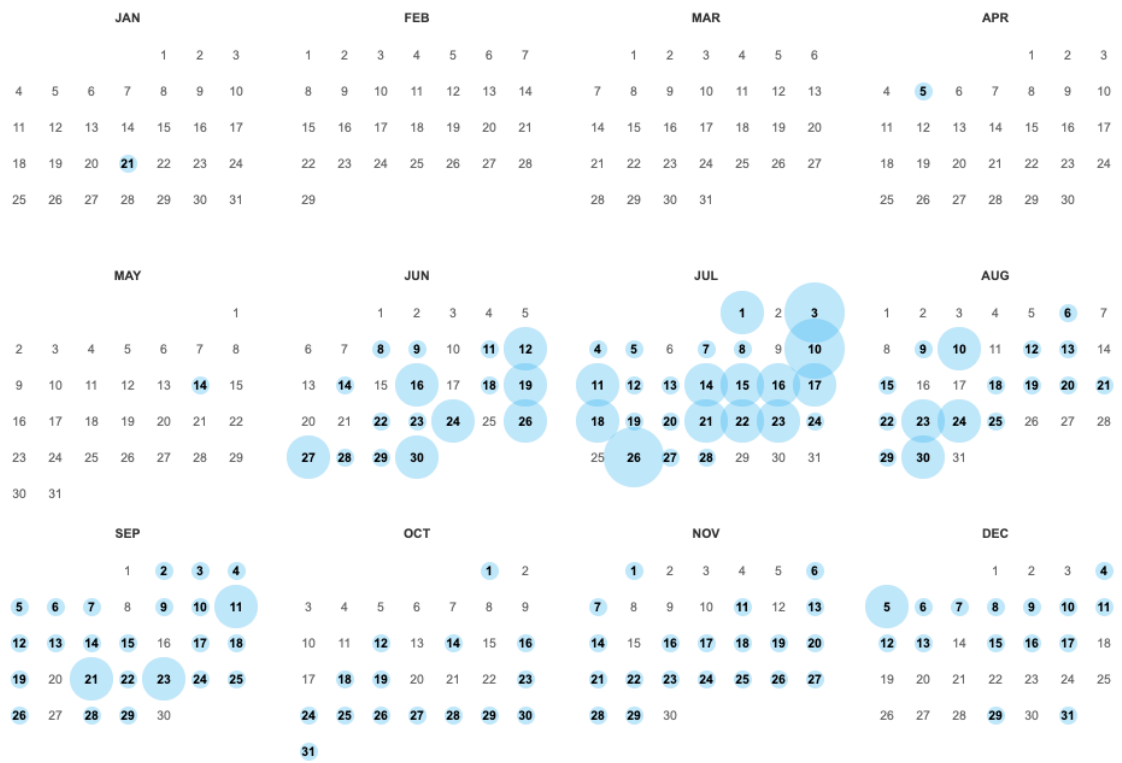
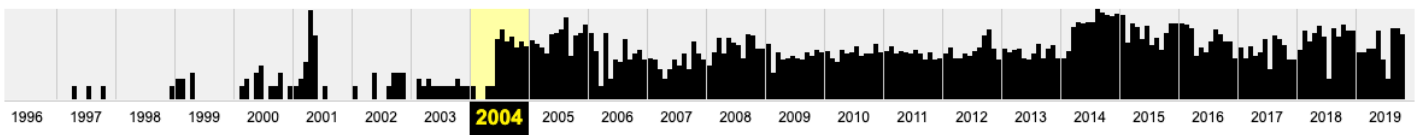
INTERNET ARCHIVE Explore more than 387 billion web pages saved over time

DONATE **Wayback Machine**

www.unt.edu

Calendar · Collections ^{beta} · Changes ^{beta} · Summary · Site Map

Saved 4,253 times between April 10, 1997 and October 24, 2019.



Note
This calendar view maps the number of times **www.unt.edu** was crawled by the Wayback Machine, *not* how many times the

INTERNET ARCHIVE Explore more than 387 billion web pages saved over time

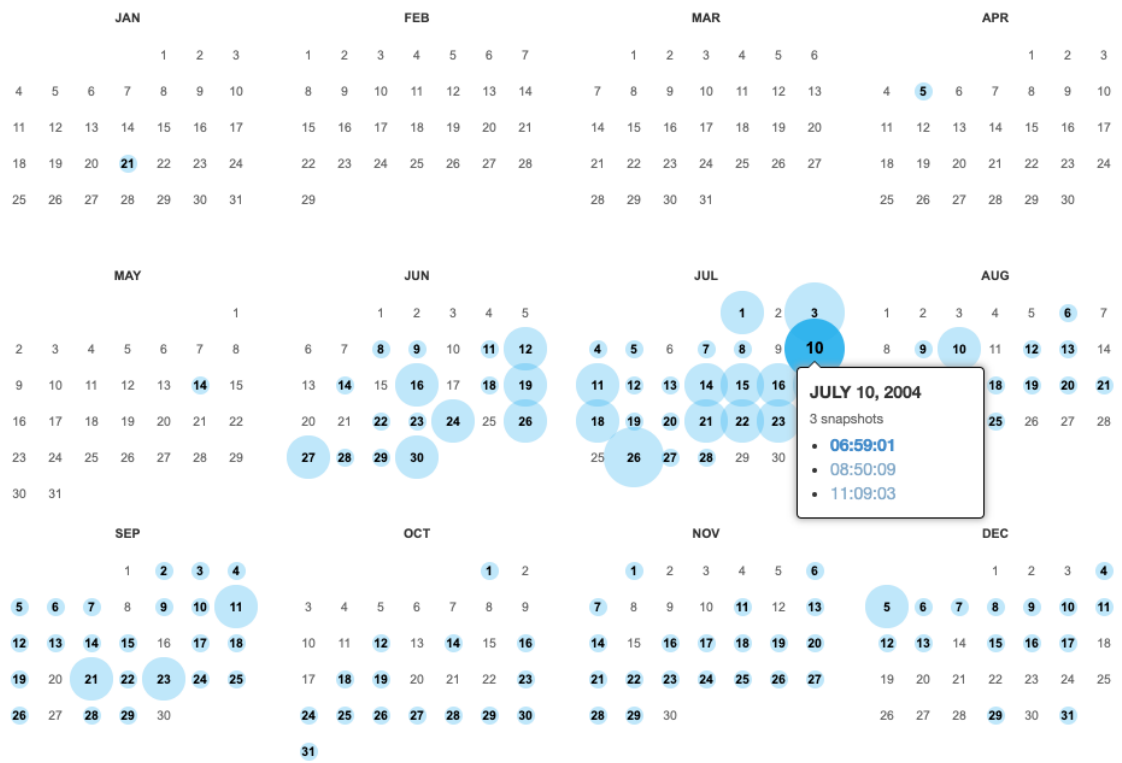
DONATE **Wayback Machine**

Calendar · Collections ^{beta} · Changes ^{beta} · Summary · Site Map

Saved **4,253** times between **April 10, 1997** and **October 24, 2019**.



Sat, 10 Jul 2004 06:59:01 GMT (why: alexa_dx, alexacrawls)



Note
This calendar view maps the number of times **www.unt.edu** was crawled by the Wayback Machine, *not* how many times the

UNIVERSITY of NORTH TEXAS

UNT is the leading university of the Dallas-Fort Worth region

links especially for: Go search the UNT web: search

Thinking about UNT?

- Admissions
- Registering for and attending classes
- Departments and contact information
- Campus life
- Working at UNT
- UNT Research
- Athletics
- Tours and maps
- Computing resources and support
- Distance education and WebCT
- Libraries
- Resources for UNT faculty and staff
- UNT president and administration
- Alumni and friends
- Community services and resources




Get new schedule for summer commencement Saturday, August 7

Beware: construction on Eagle Drive Begins Monday June 21

UNT News

Chemist wins semiconductor award
For the third consecutive year, Dr. Jeffrey Kelber received the Semiconductor Research Corporation inventor recognition award for his computer chip research.

Trailblazers gather on campus for reunion during 50 years of progress and opportunity celebrations
Members of the University of North Texas Trailblazers, an alumni organization including the university's earliest African American students, are gathering on campus July 15-17 for their sixth annual reunion.

Not your usual summer jobs
The Texas Academy of Mathematics and Science Summer Research Program pays high school students to conduct research at a UNT laboratory or at another research laboratory in Texas.

Greater Texas Foundation gives \$25,000 endowment to UNT
The Greater Texas Foundation of Bryan, Texas, has donated \$25,000 for a scholarship endowment at the University of North Texas Dallas Campus and the future UNT at Dallas.

[More news at UNT >>](#)

UNT Web Services

Register online

- [Register for Fall 2004 classes](#)

Pay your tuition online

- [Pay your tuition online for Summer classes](#)
- [Pay your tuition online for Fall 2004 classes](#)

Financial Aid online services

- [Financial Aid services online for current recipients](#)
- [Financial Aid services online for recipients applying for Fall 2004](#)

Check your grades and view your records

- [Access Student Services Information System by Technology \(ASSIST\) online](#)

UNT Events

Selected upcoming events

		
UNT Trailblazers Reunion	Rec sports summer adventure trips	Schedule a campus tour online
Honoring all former African American students who attended North Texas July 15 - 17	Sign up for some adventure this summer	A free campus tour is the best way to see a lot of the university in a short time.

Today at UNT

[Center for Play Therapy Summer Institute](#)
9 a.m. until 4:30 p.m.
Location: Silver Eagle Suite A
Check event for participation information

[Jay Novacek Football Camp](#)
Check event for participation information

Featured web sites

- [UNT System](#)
- [UNT Dallas Campus](#)
- [Undergraduate majors and](#)

INTERNET ARCHIVE
Wayback Machine

http://www.unt.edu/ JUN JUL AUG
2003 2004 2005

4,253 captures
10 Apr 1997 - 24 Oct 2019

COLLECTED BY

Organization: [Alexa Crawls](#)
Starting in 1996, [Alexa Internet](#) has been donating their crawl data to the Internet Archive. Flowing in every day, these data are added to the [Wayback Machine](#) after an embargo period.

Collection: [Alexa Crawl DX](#)
Crawl DX from Alexa Internet. This data is currently not publicly accessible.

TIMESTAMPS

<http://www.unt.edu/imageassets/constructfeature.gif> +1 year 3 months
http://www.unt.edu/imageassets/finalindex_r1_c4.gif -5 months 25 days
<http://www.unt.edu/imageassets/spacer.gif> -5 months 24 days
<http://www.unt.edu/imageassets/50yearsprogressfeature.gif> +5 months 16 days
<http://www.unt.edu/imageassets/gradfeature.jpg> +4 months 14 days
<http://www.unt.edu/imageassets/webservicestabletop.gif> +2 months 10 days
<http://www.unt.edu/imageassets/untlogo.gif> +2 months 10 days
<http://www.unt.edu/imageassets/transparent.gif> +2 months 10 days
<http://www.unt.edu/imageassets/thinkingfeature3.gif> +2 months 10 days
<http://www.unt.edu/imageassets/researchhomeb.gif> +2 months 10 days
<http://www.unt.edu/imageassets/outdfeature.gif> +2 months 10 days
http://www.unt.edu/imageassets/newsbit_r1_c1.gif +2 months 10 days
<http://www.unt.edu/imageassets/newhomerandten.jpg> +2 months 10 days
<http://www.unt.edu/imageassets/libraryhomeb.gif> +2 months 10 days
http://www.unt.edu/imageassets/finalindex_r8_c1.gif +1 month 13 hours
http://www.unt.edu/imageassets/finalindex_r7_c1.gif +1 month 13 hours
http://www.unt.edu/imageassets/finalindex_r4_c1.gif +1 month 1 hour
http://www.unt.edu/imageassets/finalindex_r3_c1.gif +1 month 1 hour

law schedule for summer commencement
 Saturday, August 7
 Beware: construction on Eagle Drive Begins Monday June 21
 UNT Web Services
 register online
 Register for Fall 2004 Classes

Websites that aren't online anymore

<http://oceancommission.gov>

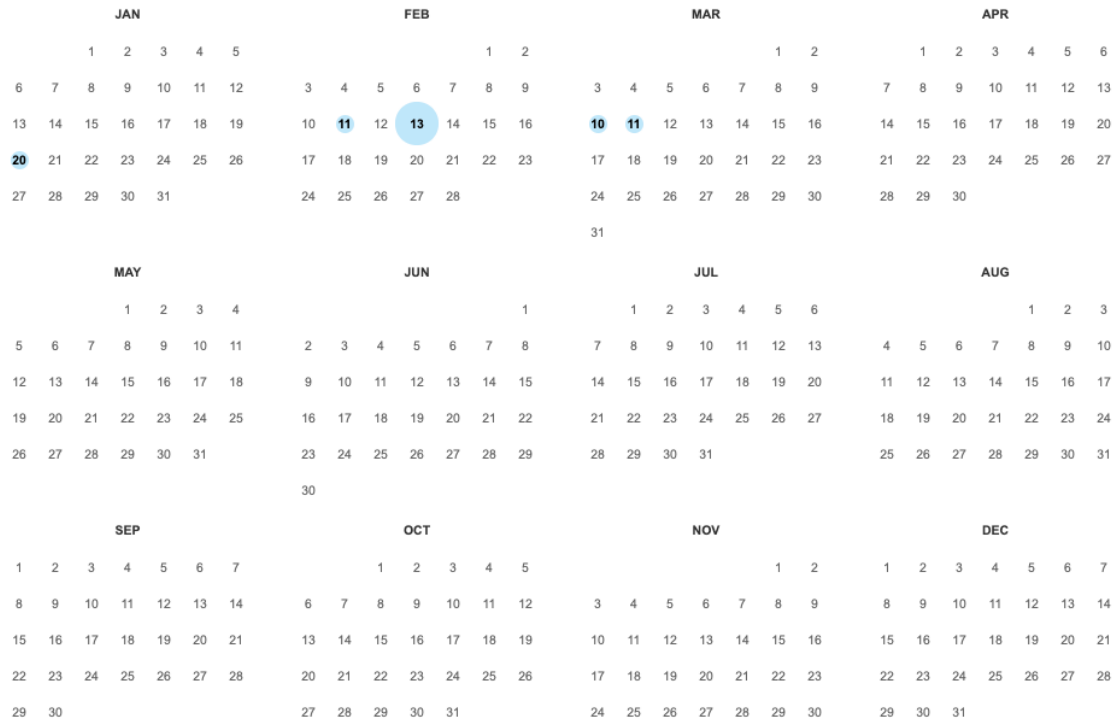
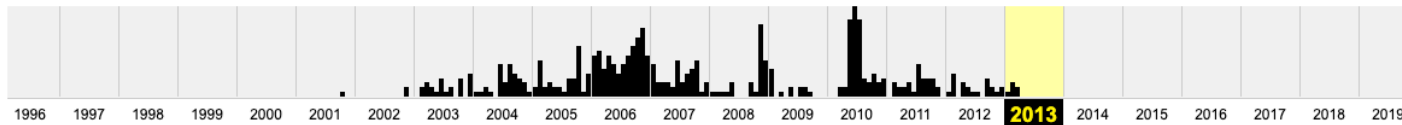


Explore more than 387 billion web pages saved over time

http://www.oceancommission.gov/

Calendar · Collections · Changes · Summary · Site Map

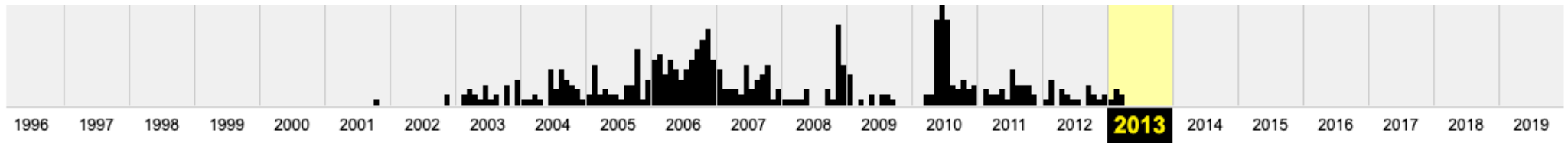
Saved 455 times between October 11, 2001 and March 11, 2013.



Note

This calendar view maps the number of times <http://www.oceancommission.gov/> was crawled by the Wayback Machine,

Saved **455 times** between **October 11, 2001** and **March 11, 2013**.





The U.S. Commission on Ocean Policy

Mandated by the [Oceans Act of 2000](#) (Public Law 106-256), authorized by Congress and appointed by the President, the U.S. Commission on Ocean Policy is required to establish findings and make recommendations to the President and Congress for a coordinated and comprehensive national ocean policy. The new policy will address a [broad range of issues](#) (pdf, 96kb), from the stewardship of marine resources and pollution prevention to enhancing and supporting marine science, commerce and transportation.

The [16-member Commission](#), which began its work in September 2001, launched a series of [public meetings](#) to gather information and hear about the most pressing issues that the Nation faces regarding the use and stewardship of the oceans. After hearing from 440 presenters in 10 cities over 11 months, the Commission completed its information-gathering phase in October 2002. The Commission entered its [deliberative phase](#) in November 2002.

The Commission's schedule to complete its work is determined by the breadth of its mandate, the complexity of the issues it must address, and the need to analyze the voluminous comments received on various draft recommendations that have been put forward.

Early 2004 - Draft Report Circulated

A draft will be made available for the Governors and other stakeholders to review sometime in early 2004. A notice of availability of the draft report will be published in the Federal Register and a copy will be available on the Commission's website at that time.

Final Report Delivered

The Commission will provide the Governors and other stakeholders a reasonable time frame to review and comment on the draft. Sufficient time will be allowed for any needed changes to the report as a result of gubernatorial or other stakeholder comments. When such changes are made and the final report is approved by the Commission, it will be transmitted to the President and Congress.

Recent milestones from the Commission's work include the Draft Policy Options documents which were discussed at the public meetings in [November 2002](#), [January 2003](#) and [April 2003](#). In May 2003, the Commission posted the initial framework for the report in its first [Table of Contents](#) (pdf, 180kb) document. That document will be updated regularly as the report evolves. Two more Commission documents were completed in June 2003, [A Synthesis of Testimony Organized by Policy Topic](#) and [A Summary of Testimony Indexed by Presenters](#).

Search go button

- [About the Commission](#)
- [Commission Meetings](#)
- [Commission Documents](#)
- [News and Notices](#)
- [Public Comment](#)
- [Calendar](#)

New This Month

[Statement on the Release of the Report](#)



Library of Congress

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

PROGRAM
Web Archiving

- About This Program
- For Researchers
- For Site Owners
- Archived Websites

About This Program

Web Archiving Collaborations

Frequently Asked Questions

Glossary

Contact Us

Related Resources

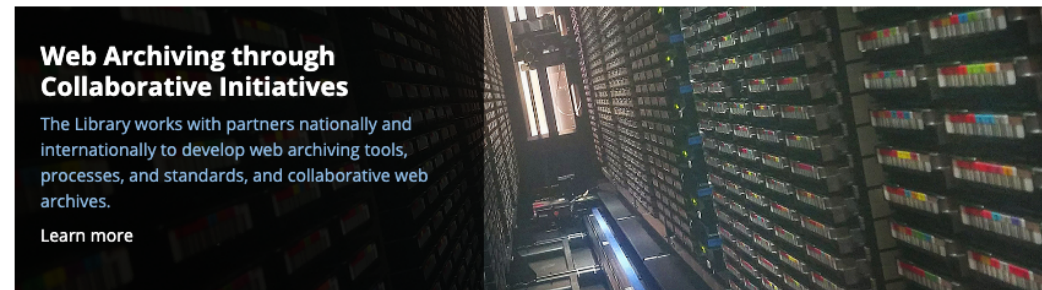
The Signal

Collection Policy Statements and Supplementary Guidelines

Recommended Formats Statement for Websites

Digital Collections Management

About This Program



Web Archiving through Collaborative Initiatives

The Library works with partners nationally and internationally to develop web archiving tools, processes, and standards, and collaborative web archives.

[Learn more](#)

The Library of Congress Web Archive selects, preserves, and provides access to archived web content selected by subject experts from across the Library, so that it will be available for researchers today and in the future. Websites are ephemeral and often considered at-risk born-digital content. New websites form constantly, URLs change, content changes, and websites sometimes disappear entirely. Websites document current events, organizations, public reactions, government information, and cultural and scholarly information on a wide variety of topics. Materials that used to appear in print are increasingly published online.

This site provides general information about program activities, information for researchers who are interested in using the web archives, and information for site owners who might be included in the archives.

For Researchers

The Library of Congress web archives are organized in thematic and event-based collections, and contain websites documenting a variety of U.S. and international organizations representing a broad range of subjects and topic areas. Examples include select U.S. government sites from the Legislative, Judicial, and Executive branch agencies; select foreign government sites; campaign websites and political parties documenting U.S. and select foreign elections; non-profit organizations; journalism and news; creative sites such as those documenting comics, music, authors, and art; legal sites; and international organizations. While most web archives are collected as a part of one or more event or thematic archives, the Library also

CONTACT US

Comments, questions, and suggestions related to Web Archiving and this website can be sent to us [online](#).

ADDRESS:

Web Archiving Team
Library of Congress
101 Independence Avenue, SE
Washington DC 20540-1310



COLLECTION

Foreign Law Web Archive

Collection Period: June 20, 2016 to ongoing. The Foreign Law Web Archive is a collection of foreign legal materials, including gazettes and judicial sites. Many foreign legal materials are now posted online, with some jurisdictions dispensing with a print publication entirely. Certain jurisdictions' legal materials are challenging to acquire or considered at-risk of disappearing from the web. The Law Library of Congress is now...

Contributor: Library of Congress - Library of Congress. Law Library

Date: 2016

Collection Items: [View 18 Items](#)



COLLECTION

General News on the Internet Web Archive

Collection Period: June 2014 to ongoing. The General News on the Internet Web Archive contains general news sites with a focus on stories of national interest for all audience levels. The collection contains born-digital publications and publications that were once only available in print and are now only available online. Given the dynamic nature of the 24-hour news cycle of today, these archives are...

Contributor: Library of Congress - Library of Congress. Serial & Government Publications Division

Date: 2014

Collection Items: [View 54 Items](#)



COLLECTION

Indian General Election 2009 Web Archive

Collection Period: April 9, 2009 to June 29, 2009. This collection of websites documents the 2009 Indian General Elections. As the world's most populated democratic nation with the largest number of eligible voters, India's political landscape encompasses an electorate comprised of 44 major linguistic groups and over 61 officially recognized political parties at the national and state level. To constitute the 15th Lok Sabha...

Contributor: Library of Congress

Date: 2012

Collection Items: [View 59 Items](#)



COLLECTION

Indonesian General Election 2009 Web Archive

Collection Period: March 10, 2009 to October 21, 2009. This collection of websites documents the Indonesian general elections held in 2009. The collection covers three separate rounds of elections during that year - the parliamentary election held in April, the first round of the presidential election held in April followed by the second/final round of the presidential election in September. As the third most...

Contributor: Library of Congress

Date: 2009

Collection Items: [View 80 Items](#)



COLLECTION

International Tribunals Web Archive

Collection Period: June 13, 2014 to ongoing. The "International Tribunals Archive"(ITA) is an archive organized and managed by the Library of Congress with the purpose of digitally storing relevant websites hosting information about the most important

COLLECTION General News on the Internet Web Archive

About this Collection [Collection Items](#)

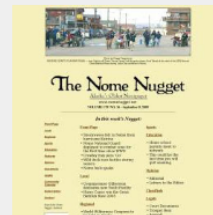
Featured Content



Huffington Post



The Texas Tribune



The Nome Nugget



Indian Country Media Network



Slate Magazine

About this Collection

Rights & Access

Expert Resources

[Web Archiving Program Information](#)
[Tips on Searching the Web Archive](#)

Listen to this page

About this Collection

Collection Period:
June 2014 to ongoing.

The General News on the Internet Web Archive contains general news sites with a focus on stories of national interest for all audience levels. The collection contains born-digital publications and publications that were once only available in print and are now only available online.

Given the dynamic nature of the 24-hour news cycle of today, these archives are meant to capture as much of the news distribution as possible given current limitations in technology and resources. For most news-based sites, we use a hybrid approach of weekly captures of the websites, augmented with twice-daily capture of known RSS feeds (Real Simple Syndication). This produces a more complete news archive.

The General News on the Internet Web Archive is an ongoing collecting effort.

Follow Us | [Take our survey](#)

FORMAT
Archived Web Site

Search Archived Web Sites Collections with Archived Web Sites

Results: 1-25 of 17,446 | Refined by:

Refine your results

- Available Online 17,446
- All Items 17,446

- Original Format**
- Archived Web Site
 - Web Page 3
 - Legislation 1

- Online Format**
- Web Page 17,355
 - PDF 47
 - Image 2
 - Online Text 2
 - Audio 1

- Date**
- 2000 to 2099 17,441
 - 1900 to 1999 681

- Location**
- United States 12,112

Search Archived Web Sites

View Sort By



ARCHIVED WEB SITE
Online Encyclopedia of Silver Marks, Hallmarks & Makers' Marks

A comprehensive internet resource for research of Silver Marks, Hallmarks & Maker's Marks found on antique and vintage silver and silverplate. Web Site. electronic | Electronic (Form).

Date: 2010-08-03



ARCHIVED WEB SITE
NOYNOY AQUINO FOR PRESIDENT

The Official Website of Benigno 'Noynoy' Aquino III, his Platform, History, Current Issues, and Statements on Magindanao, and Martial Law. A home for his Volunteers and Supporters to help fight against corruption. Web Site. electronic | Electronic (Form).

Contributor: Aquino, Benigno Simeon Cojuangco

Date: 2010-04-14

Arquivo.pt (Portuguese Web Archive)

<https://arquivo.pt>

Q Words or URL

To look for

1991

6 Ago

2022

9 Mar

Pages

Images

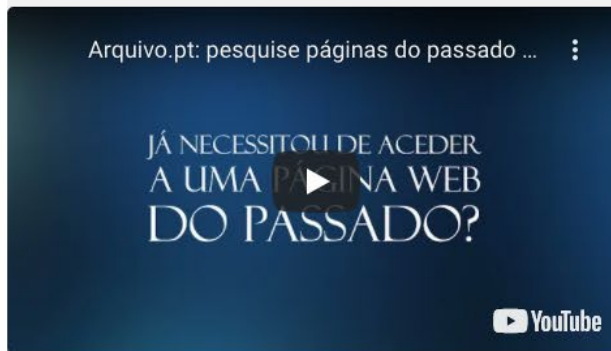
Narrative

Advanced Search

Arquivo.pt Award 2022



What is Arquivo.pt?



User Testimonials



Examples of preserved pages



thematic exhibitions



UK Web Archive

<https://www.webarchive.org.uk>

Collection guides

UK Web Archive



Web Archiving collects, makes accessible and preserves web resources of scholarly and cultural importance from the UK domain.

About the collection

The Library has three main collections of archived websites:

1. **Open UK Web Archive** has been collecting websites by permission since 2004 and is the smallest of the three collections. Its selective nature means that the sites are of high quality, and the archival copies have been manually quality-checked and carefully annotated. This curated collection covers aspects of UK life and culture, as well as specific events such as general elections, the 2012 Olympic Games, and the commemoration of the centenary of the First World War. Selected websites are continuously added and anyone can [nominate a UK website](#) for inclusion.
2. **JISC UK Web Domain Dataset** (1996-2013) contains all of the resources from the [Internet Archive](#) that were hosted on domains ending in '.uk', or that are required in order to render those UK pages.
3. **Non-Print Legal Deposit UK Web Archive** includes millions of websites obtained through our annual archiving of the entire UK domain. This was enabled by the non-print legal deposit regulations introduced by the government in April 2013.

Read the [UK Web Archive blog](#) to keep up-to-date with research and developments.

Related guides

Digital mapping

Geospatial data, cartographic applications, digital aerial photography and scanned-in historic map materials

Datasets for image analysis

Image collections suitable for large-scale image-analysis-based research

Datasets for content mining

Content suitable for use in text and data mining research

Web news

Our Legal Deposit UK Web Archive includes millions of websites from the UK domain

Television and radio news

The Broadcast News service provides onsite access to recent news programmes

The Fay Godwin collection

The collection comprises the photographer's complete studio collection

Russia in the UK Web Archive Collection

Russia in the UK Special Collection, UK Web Archive

Datasets about our collections

Bibliographic datasets relating to our published and archival holdings

Digitised printed books (18th-19th century)

Over 60,000 out-of-copyright digitised books from the 18th and 19th century available for viewing, search, download and digital research.

Asian Britain: music and audiovisual

Search the UK Web Archive

Enter a specific website URL (e.g. www.bl.uk) or any word or phrase... ⓘ

What we do

The UK Web Archive (UKWA) collects millions of websites each year, preserving them for future generations. Use this site to discover old or obsolete versions of UK websites, search the text of the websites and browse websites curated on different topics and themes.

The UKWA is a partnership of the six [UK Legal Deposit Libraries](#).

Topics and Themes

Topics and Themes are groups of websites brought together on a particular theme by librarians, curators and other specialists, often working in collaboration with key organisations in the field.



[British Stand-up Comedy Archive](#)

March 9, 2022



[French in London](#)

IS - Brown Bag



[News Sites](#)

Archive-It

<https://archive-it.org>



Welcome to Archive-It! Attend a live informational webinar and demo to learn more about the service

Contact Us to sign up for an upcoming session: Oct 31 2019, 11:00 AM PDT Nov 14 2019, 11:00 AM PST

Explore Collections Find a Collection by Name Search Show All Collections



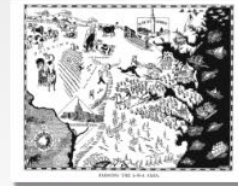
Avery Library By Columbia University Libraries

A collection of websites chosen by subject specialists from the Avery Architectural and Fine Arts Library at Columbia University.



Earthquake in Haiti By Internet Archive Global Events

This collection is currently documenting the events of the January 2010 earthquake in Haiti and the aftermath, including the rescue efforts from around the world and the stories and circumstances of...



Maryland State Document Collection By University of Maryland

This collection contains material created by the State of Maryland related to state planning.

Explore Collecting Organizations Find an Organization by Name Search Show All Organizations



Cornell University Library

Based on the number of volumes in its collections, Cornell University Library (CUL) is one of the ten largest academic research libraries in the United States. Within its 20 unit libraries, holdings...



Bodleian Libraries

The Bodleian Libraries is the integrated library service of the University of Oxford. Established in 2000, it comprises nearly 40 libraries. Among these are major research libraries - including the...



New Brunswick Free Public Library

New Brunswick Free Public Library





Explore >> Columbia University Libraries >> Human Rights



Human Rights

Collected by: [Columbia University Libraries](#)

Archived since: May, 2008

Description: An initiative of CUL's Center for Human Rights Documentation and Research, the Human Rights Web Archive is an effort to preserve and ensure access to freely available human rights resources created mainly by non-governmental organizations, national human rights institutions, and individuals.

Subject: [Society & Culture](#), [Human rights](#), [Non-governmental organizations](#), [Human rights workers](#), [National human rights institutions](#), [Web archives](#)

Creator: [Columbia University, Center for Human Rights Documentation and Research](#)

Collector: [Columbia University Libraries, Center for Human Rights Documentation and Research](#)

Narrow Your Results

Group

Sort By: [Count](#) | [\(A-Z\)](#)

- [Amnesty International sections \(66\)](#)
- [Blogs by individuals \(16\)](#)
- [National human rights institutions \(142\)](#)
- [Non-governmental organizations \(660\)](#)
- [Truth commissions, tribunals, and courts \(27\)](#)

Subject

Sort By: [Count](#) | [\(A-Z\)](#)

- [Human rights \(700\)](#)
- [Human rights advocacy \(209\)](#)
- [National human rights institutions \(132\)](#)
- [Democracy \(64\)](#)
- [Civil rights \(62\)](#)

[More ▾](#)

Creator

Sort By: [Count](#) | [\(A-Z\)](#)

- [Asociación Paz con Dignidad \(4\)](#)
- [Comité de Familiares de Detenidos-Desaparecidos en Honduras \(4\)](#)
- [National Human Rights Commission \(Mauritius\) \(4\)](#)
- [Sánchez, Yoani, 1975- \(4\)](#)
- [Amnesty International Hong Kong \(3\)](#)

[More ▾](#)

Language

Sort By: [Count](#) | [\(A-Z\)](#)

- [English \(655\)](#)
- [Spanish \(150\)](#)
- [Arabic \(122\)](#)

Sites for this collection are listed below. Narrow your results at left, or enter a search query below to find a site, specific URL or to search the text of archived webpages.

Search

Clear

[Sites](#) | [Search Page Text](#)

Page 1 of 10 (922 Total Results)

[Next Page ▶](#)

Sort By: [Title \(A-Z\)](#) | [Title \(Z-A\)](#) | [URL \(A-Z\)](#) | [URL \(Z-A\)](#)

Title: [5-18 Kinyŏm Chaedan](#)

URL: <http://518.org/>

Captured 72 times between [Jun 18, 2010](#) and [Jul 23, 2019](#)

Videos: [9 Videos Captured](#)

Subject: [Human rights](#), [Political violence](#), [Collective memory](#)

Group: [Non-governmental organizations](#)

Creator: [5-18 Kinyŏm Chaedan \(Korea\)](#)

Language: [Korean](#), [English](#)

Coverage: [Korea \(South\)](#)

Collector: [Columbia University Libraries, Center for Human Rights Documentation and Research](#)

Title: [Assistance Association for Political Prisoners \(Burma\)](#)

URL: <http://aappb.org/>

Description: Burma-focused human rights organization based in Thailand that assists current and former political prisoners and their loved ones and tries to publicize the conditions of political prisoners in Burma. Contains reports and photo galleries.

Captured 64 times between [May 14, 2009](#) and [Jul 24, 2019](#)

Videos: [58 Videos Captured](#)

Subject: [Human rights](#), [Political prisoners](#), [Prisoners—Civil rights](#)

Group: [Non-governmental organizations](#)

Creator: [Assistance Association for Political Prisoners \(Burma\)](#)

Language: [English](#)

Coverage: [Burma](#)

0

UNT Libraries' CyberCemetery

<https://cybercemetery.unt.edu>



About this Collection

Overview

At a Glance

Latest Additions

Cite This Collection

Explore Holdings

Contact Us

Items

Titles

Dates

Locations

Statistics

API

Share



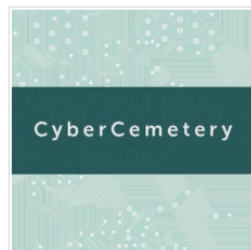
Search this Collection

Search across 121 items



Feed

CyberCemetery



About This Collection

The CyberCemetery is an archive of government web sites that have ceased operation (usually web sites of defunct government agencies and commissions that have issued a final report). This collection features a variety of topics indicative of the broad nature of government information. In particular, this collection features web sites that cover topics supporting the university's curriculum and particular program strengths.

Background

Foreseeing the potential preservation problems created by federal agencies' ventures into electronic publishing, UNT became the second depository library in the nation to join the Federal Depository Library Program Content Partnerships Program. This program attempts to ensure permanent public access to electronic federal information. As a participant, the [UNT Libraries](#), part of the Federal Depository Library Program (FDLP), were designated the host of the permanent online collection of the defunct Advisory Commission on Intergovernmental Relations (ACIR).

We've since expanded our Content Partnership with the federal government to include dozens of other defunct federal agency web sites. This electronic repository is popularly known as the CyberCemetery. In recognition of our work in this area, the UNT Libraries were designated an [Affiliated Archives](#) of the National Archives and Records Administration (NARA) in 2006. Under this agreement, the UNT Libraries will continue to preserve and provide access to the records of defunct government web sites, while the [NARA](#) will legally accession the records as part of the Archives of the United States and will join the UNT Libraries and the [Government Printing Office \(GPO\)](#) in ensuring the preservation of these valuable records. We are currently the only non-military educational institution to be designated an Affiliated Archives.

Source: Kelly, Melody Specht. [Our History: Celebrating 60 Years of Public Service](#). Accessed March 31, 2010.

How to Get Your Website in the CyberCemetery

If your agency or commission is closing, you can make sure that your website's information is preserved. For more information, see our [CyberCemetery FAQs](#).

<https://cybercemetery.unt.edu/>

Archive of Federal commissions, committees and initiatives that are no longer active.

Started in 1997 with the Advisory Commission on Intergovernmental Relations (ACIR)



Advisory Commission on Intergovernmental Relations

Welcome to The U.S. Advisory Commission on Intergovernmental Relations' Home Page

The Advisory Commission on Intergovernmental Relations (ACIR) is a permanent, independent, bipartisan intergovernmental agency established by Public Law 86-380 in 1959. As it was established, ACIR's mission is:

To strengthen the American federal system and improve the ability of federal, state, and local governments to work together cooperatively, efficiently, and effectively.

To learn more about ACIR, please click on one of the following subjects,

- [WHAT IS ACIR?](#)
- [ACIR NOW](#)
- [SERVICES, PROGRAMS, AND FUNDING](#)

scroll down this page, or
contact us directly at:
800 K St., NW
Suite 450 South Building
Washington, DC 20575

WHAT IS ACIR?

The Advisory Commission on Intergovernmental Relations (ACIR) is a permanent, independent, bipartisan agency that was established under Public Law 86-380 in 1959 to study and consider the federal government's [intergovernmental relationships](#) and the nation's intergovernmental machination.

[The Commission](#) is composed of 26 individuals who represent the interests of the federal system's intergovernmental partners in matters of intergovernmental concern, and is the only established, freestanding part of the federal system in which the views of the federal government's intergovernmental partners are openly aired, and in which difficulties and inefficiencies in the federal system's Intergovernmental Relationship's are examined.

Intergovernmental Relations:

The studies of Intergovernmental Management and Intergovernmental Relations are fields in which researchers study how different levels of government interact with one another, and attempt to define how they should interact with one another in the context of The Constitution. Intergovernmental Relationships have profound effects on self-government in America, and determining the appropriate role for federal government within the governmental system is as high on the national agenda today as it was when The Constitution was written in 1787.

As the nation's foremost repository of experience and information on intergovernmental structure, finance, process, and practice, the challenge of studying the federal government's relationships with State, Local, and Tribal governments falls to ACIR.

ACIR works with State, Local, and Tribal governments, as well as other interested parties to:

- Identify emerging intergovernmental issues, trends, and turning points;
- Stimulate thought about American federalism and intergovernmental relationships;
- Educate leaders and the public about the impacts of intergovernmental reforms; and
- Promote stronger intergovernmental communication, cooperation, and coordination as the critical basis for an effective federal system.

Currently, ACIR is the hub of the intergovernmental community and maintains relationships with many Intergovernmental organizations. Several related documents and other sources of information are:

Western Hemisphere Drug Policy Commission

[HOME](#)

[DOWNLOAD REPORT](#)

[READ REPORT ONLINE](#)

[VIEW HEARING](#)

[LEARN ABOUT](#) ▼

CHARTING A NEW PATH FORWARD

REPORT OF THE WESTERN HEMISPHERE
DRUG POLICY COMMISSION

[f](#) Share

[t](#) Tweet

[w](#) Share

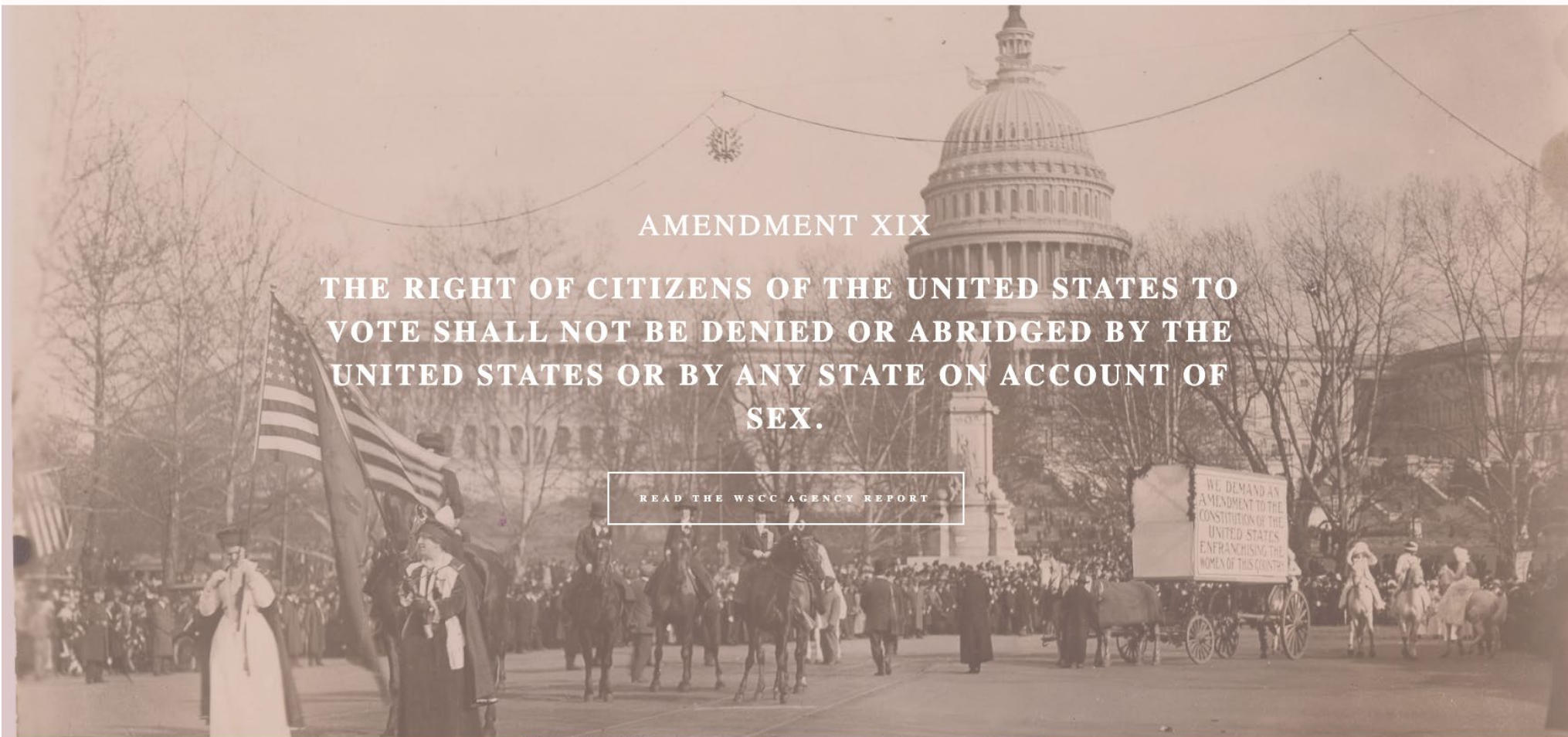
[e](#) Email

[p](#) Print

Executive Summary



ABOUT **PROGRAMS** **LEARN**
partners media federal commemorations





UNITED STATES OF AMERICA
CYBERSPACE
SOLARIUM
COMMISSION



Introduction

The Cyberspace Solarium Commission (CSC) was [established](#) in the John S. McCain National Defense Authorization Act for Fiscal Year 2019 to "develop a consensus on a strategic approach to defending the United States in cyberspace against cyber attacks of significant consequences." The finished report was presented to the public on March 11, 2020.

Our Report

The Cyberspace Solarium Commission's proposes a strategy of layered cyber deterrence. Our report consists of over 80 recommendations to implement the strategy. These recommendations are organized into 6 pillars:

1. *Reform the U.S. Government's Structure and Organization for Cyberspace.*
2. *Strengthen Norms and Non-Military Tools.*
3. *Promote National Resilience.*
4. *Reshape the Cyber Ecosystem.*
5. *Operationalize Cybersecurity Collaboration with the Private Sector.*
6. *Preserve and Employ the Military Instrument of National Power.*

[DOWNLOAD OFFICIAL REPORT](#)

[DOWNLOAD EXECUTIVE SUMMARY](#)

[DOWNLOAD LEGISLATIVE PROPOSALS](#)



COMMISSION ON EVIDENCE-BASED POLICYMAKING

[REPORT](#) [ABOUT](#) [COMMISSIONERS](#) [MEETINGS](#) [HEARINGS](#) [LIBRARY](#) [NEWS](#) [CONTACT US](#)

The Commission on Evidence-Based Policymaking (CEP) was established to develop a strategy for increasing the availability and use of data in order to build evidence about government programs, while protecting privacy and confidentiality.

[Read More](#)

The Commission on Evidence-Based Policymaking sunsets on September 30, 2017. While this website is no longer being updated, it will be available until July 1, 2018.

For updates on related evidence-based policymaking efforts, see the Evidence-Based Policymaking Initiative of the Bipartisan Policy Center at bipartisanpolicy.org/evidence.

■ [CEP Final Report: The Promise of Evidence-Based Policymaking](#)



Latest

- [News](#)
- [Meetings](#)
- [Hearings](#)



Commission on Evidence-Based Policymaking Releases Final Report

September 06, 2017

Members of the Commission on Evidence-Based Policymaking announce the Commission's final report and recommendations.



Commission Meeting on March 13

March 01, 2017

The Commission on Evidence-Based Policymaking (CEP) will hold a public meeting on federal statistical and partnership infrastructure for evidence building.



Commission Meeting on February 24

February 22, 2017

The Commission will hold a public meeting on the role of legal standards and technology in maximizing data security and privacy.

[View All News](#)

Today contains over 120 websites

UNT Libraries is an Affiliated Archives with the
U.S. National Archives and Records
Administration (NARA) because of the
CyberCemetery

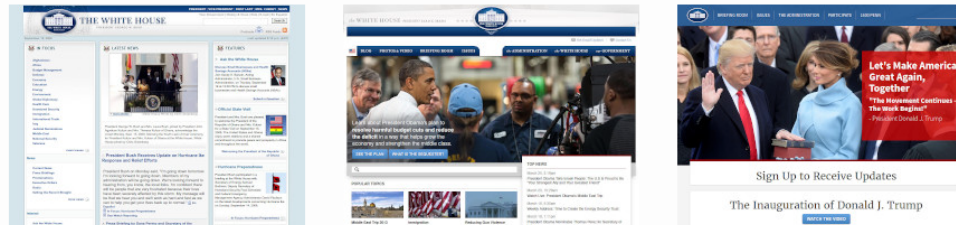
End of Term Web Archive

<https://eotarchive.org>

End of Term Presidential Web Archives 2008, 2012, 2016, 2020

Purpose

The End of Term Web Archive captures and saves U.S. Government websites at the end of presidential administrations. The EOT has thus far preserved websites from administration changes in 2008, 2012, 2016, and 2020.



Whitehouse.gov captures from: September 15, 2008; March 21, 2013; and February 3, 2017.

Archive Scope

The End of Term Web Archive contains federal government websites (.gov, .mil, etc) in the Legislative, Executive, or Judicial branches of the government. Websites that were at risk of changing (i.e., whitehouse.gov) or disappearing altogether during government transitions were captured. Local government websites, or any other site not part of the federal government domain were out of scope.

U.S. Federal Government Domain End of Term 2020 Web Archive

For the End of Term 2020, The Library of Congress, University of North Texas Libraries, Internet Archive, Stanford University Libraries, and the U.S. Government Publishing Office (GPO) joined efforts again, this time with new partners Environmental Data & Governance Initiative (EDGI) and the National Archives and Records Administration (NARA), to preserve public United States Government websites at the conclusion of the presidential administration ending January 20, 2021. This web harvest – like its predecessors in 2008, 2012, and 2016 – was intended to document the federal government's presence on the World Wide Web during the transition of presidential administrations and to enhance the existing collections of the partner institutions. This broad comprehensive crawl of the .gov domain includes as many federal .gov sites as we could find, plus federal content in other domains (such as .mil, .com, and social media content) and FTP'd datasets.

Nominations made by individual URL for inclusion in the End of Term Presidential Harvest 2020 are available to view in the [Nomination Tool](#). URLs submitted for consideration in bulk form via files were added to a separate [bulk Nomination Tool instance](#). The files containing the bulk list URLs have also been added to a [GitHub repository](#). The entirety of the archived content is currently [being held by the Internet Archive](#).

Browse the End of Term Web Archive

- [2008-2009](#)
- [2012-2013](#)
- [2016-2017](#)
- [Browse All](#)
- [Search Full Text](#)

Collaboration between institutions throughout the United State to capture the Federal (.gov) Web at the end of each presidential term.

Good Evening



*Welcome to the
White House*

**[Social Security Administration Announces
New Online Retirement Planner](#)**



[The President & Vice President:](#)

Their accomplishments, their families, and how to send them electronic mail --



[Commonly Requested Federal Services:](#)

Direct access to Federal Services



[Interactive Citizens' Handbook:](#) Your guide to information about the Federal government



[What's New:](#)
What's happening at the White House
[-Expanding Opportunities for America's Senior Citizens](#)
[-A National Call to Action to Close the Digital Divide](#)



[White House History and Tours:](#)
Past Presidents and First Families, Art in the President's House and Tours -- [Tour](#)



[Site News:](#)
Featured sections of our site
[-White House Easter Egg Roll](#)
[-China Trade Relations Working](#)



The White House

PRESIDENT GEORGE W. BUSH

President * News * Vice President * History & Tours * First Lady * Mrs. Cheney

YOUR GOVERNMENT KIDS ESPAÑOL CONTACT PRIVACY POLICY SITE MAP SEARCH



EMAIL UPDATES

SEARCH

Issues

- Economy
- Iraq
- Education
- National Security
- Homeland Security
- More Issues
- En Español

News

- Current News
- Press Briefings
- Proclamations
- Executive Orders
- Radio

Appointments

- Nominations
- Application

Offices

- Global Comm
- Freedom Corps
- Faith-Based & Community
- OMB
- More Offices

Major Speeches

- Iraq Transition
- State of the Union
- Saddam Capture
- UN Address
- National Address
- Iraqi Freedom
- National Address



PHOTOS

- Photo Essays
- Search photos by date

July 20, 2004 | Last Updated 11:00 p.m. (EST)

Secretary Snow Discusses Economy on "Ask the White House"



Treasury Secretary John Snow on Tuesday discussed the economy on "Ask the White House".

Read the Full Transcript [full story](#)

Steady Job Growth Continues in States Across the Country

June's employment data indicate that the President's economic policies are creating jobs and driving steady economic growth, but there is still more work to be done. [full story](#)

President Bush Welcomes President Lagos of Chile



President Bush and Chilean President Lagos discussed trade issues on Monday at the White House. [full story](#)

Press Briefing by Scott McClellan



Watch Monday's Press Briefing with White House Press Secretary Scott McClellan. [full story](#)

President Bush Welcomes Malaysian Prime Minister to White House



More Photos

White House photo by Paul Morse

President George W. Bush meets with Freedom Corps greeter Dean Gesme Sr. after arriving in Cedar Rapids, Iowa on July 20, 2004.

WHITE HOUSE RADIO

More ▶

Rumsfeld on the Radio

Secretary Rumsfeld talks with Martha Zoller, WDUN, Gainesville, GA

[Click here](#) for more ♦♦♦♦



ASK THE WHITE HOUSE

More ▶

Apollo 11 35th Anniversary

Discuss the 35th Anniversary of Apollo 11 tomorrow with Neil Armstrong, Michael Collins and Buzz Aldrin at 11am. Submit your question now.



WHITE HOUSE INTERACTIVE

More ▶

Email the White House...

Q: Marcus, Oaktown.



March 9, 2022

IS - Brown Bag



THE WHITE HOUSE

PRESIDENT GEORGE W. BUSH

Search

Podcasts RSS Feeds

November 26, 2008

Last updated 2:35 p.m. (EDT)

IN FOCUS

- Afghanistan
- Africa
- Budget Management
- Defense
- Economy
- Education
- Energy
- Environment
- Global Diplomacy
- Health Care
- Homeland Security
- Immigration
- International Trade
- Iraq
- Judicial Nominations
- Middle East
- National Security
- Veterans

[more issues](#)

News

- Current News
- Press Briefings
- Proclamations
- Executive Orders
- Radio
- Setting the Record Straight

[more news](#)

LATEST NEWS



White House Photo by Eric Draper

President George W. Bush waves to the troops at Fort Campbell, Ky., Tuesday, Nov. 25, 2008, after being introduced on stage during a Thanksgiving visit. Said the President, "Over the past seven years, folks from this base have done exactly what they were trained to do." "You have taken the battle of the terrorists overseas so we do not have to face them here in the United States. You have helped counter the hateful ideology of tyranny and terror with a more hopeful vision of justice and liberty. You're part of the great ideological struggle of our time. With the soldiers of Fort Campbell out front, the forces of freedom and liberty will prevail." White House photo by Eric Draper

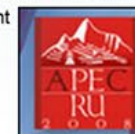
President Bush Pardons "Pumpkin and Pecan" in National Thanksgiving Turkey Ceremony

President Bush on Wednesday said, "This is my final Thanksgiving as the President. Over the past eight years, I have been given many reasons to be thankful. I'm thankful to our men and women in uniform and I am incredibly proud to have been their Commander-in-Chief. I am thankful for the armies of compassion volunteers who feed the hungry and shelter the poor. I am

FEATURES

APEC 2008

On November 21-23, 2008, President Bush attended the 16th Asia-Pacific Economic Cooperation (APEC) Leaders' Meeting in Lima, Peru.



[APEC 2008](#)

Celebrate the 61st Annual Thanksgiving Turkey Pardon

This year marks the 61st anniversary of the National Thanksgiving Turkey presentation. Though live Thanksgiving turkeys have been presented intermittently to presidents since the Lincoln administration, the current ceremony dates to 1947, when the first National Thanksgiving Turkey was presented to President Harry Truman.



[Thanksgiving at the White House](#)

Presidential Transition

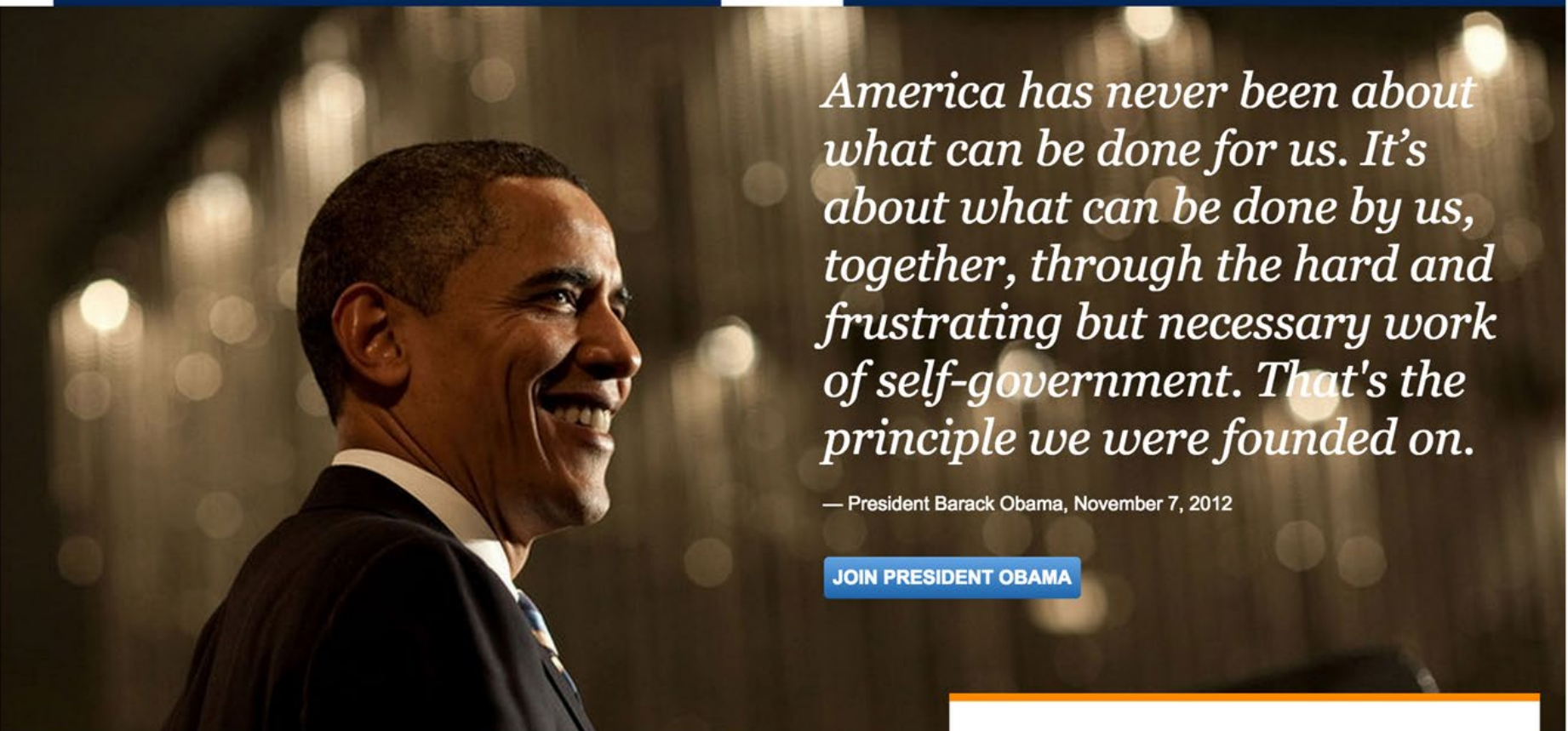
The peaceful transfer of power from one Presidential Administration to the next is a hallmark of American democracy. With our Nation at war, our homeland targeted by terrorist adversaries, and our economy facing serious challenges, the Administration is committed to establishing and executing a transition plan that minimizes disruption, maintains continuity, and addresses the major changes in government since the 2000 transition, including the Intelligence





Get Email Updates | Contact Us

[BLOG](#) [PHOTOS & VIDEO](#) [BRIEFING ROOM](#) [ISSUES](#) [the ADMINISTRATION](#) [the WHITE HOUSE](#) [our GOVERNMENT](#)



America has never been about what can be done for us. It's about what can be done by us, together, through the hard and frustrating but necessary work of self-government. That's the principle we were founded on.

— President Barack Obama, November 7, 2012

[JOIN PRESIDENT OBAMA](#)



IN FOCUS

TOP NEWS

November 11, 11:55am
Thank an American Hero

November 10, 5:45am
Weekly Address: Extending Middle Class Tax Cuts to Grow the Economy



BRIEFING ROOM

ISSUES

THE ADMINISTRATION

PARTICIPATE

1600 PENN

Search



Stay in the Loop

From photos to the President's schedule to policy updates, the Daily Snapshot has everything you need to know about a day at the White House.

START NOW



POPULAR TOPICS



Net Neutrality

President Obama is asking the...



2015 State of The Union

Did you miss this year's speech?



Immigration Action

The President is taking...

TOP NEWS

February 12, 4:20pm
The Clay Hunt Act: What the President Just Signed

February 12, 5:32pm
President Obama, a Selfie Stick, and HealthCare.gov

February 11, 8:28pm
The Authorization of Military Force Against ISIL Terrorists: What You Need to Know



BRIEFING ROOM

ISSUES

THE ADMINISTRATION

PARTICIPATE

1600 PENN



“Believe in yourselves. Believe in your future.
And believe, once more, in America.”

- President Donald J. Trump

Sign Up to Receive Updates

FEATURED CONTENT



President Trump Praises Law Enforcement While Delivering Remarks on MS-13

President Trump gave remarks to local, State, and Federal law enforcement about the challenges of, and the Administration's success regarding, MS-13.

[READ MORE](#)



HELP IS HERE

WE PASSED THE AMERICAN RESCUE PLAN



- See How The American Rescue Plan Helps People Like You
- Submit Your Story About The American Rescue Plan
- Learn How Working Americans Will Receive Their Direct Payments



**Selection, harvesting, archiving distributed to meet
local collection needs.**

Centralized access

UNT Libraries has acquired archive for 2008 and 2012 resulting in a 16TB and 30TB dataset

2016 - 159TB (HTTP) 150TB (FTP)
2020 - 300TB

Working to move a copy of
EOT archives to Open Data on AWS
<https://aws.amazon.com/opendata/>

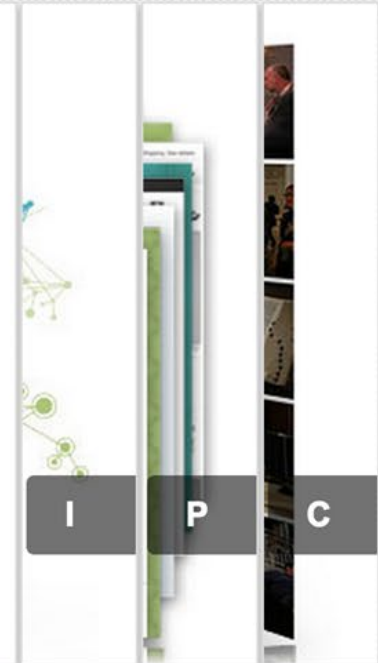
Based on the work by Common Crawl
Designed for computational access to datasets.
<https://commoncrawl.org>

International Internet Preservation Consortium

<https://netpreserve.org>

A GLOBAL NETWORK OF EXPERTS ARCHIVING THE WEB FOR FUTURE GENERATIONS.

[Learn More About IIPC](#)



March 9, 2022

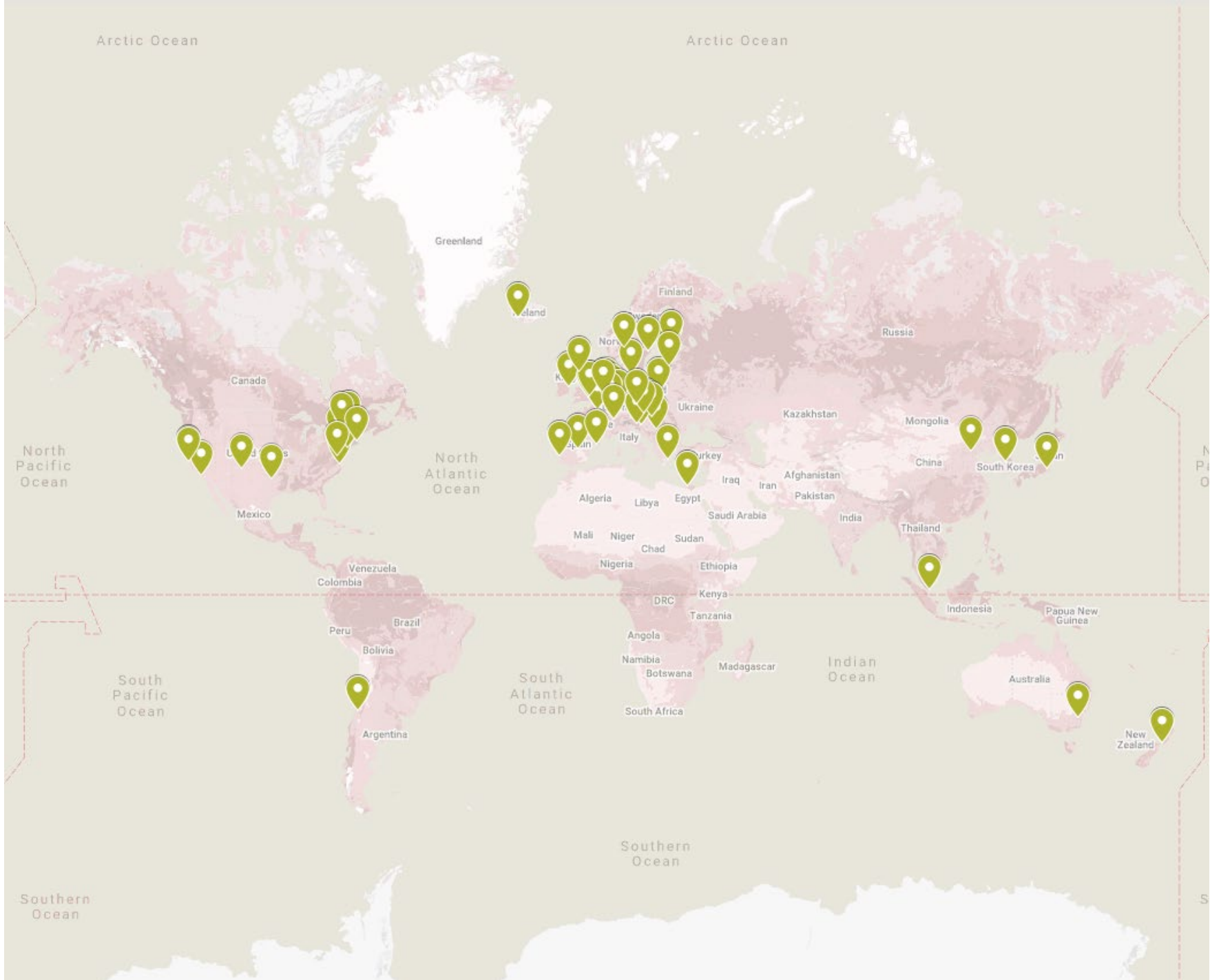
IS - Brown Bag

WHY ARCHIVE THE WEB?

Information published on the World Wide Web today will be the primary resources for future



55 member institutions



Members from 25 countries: North America,
Europe, Asia, Africa

National Libraries
National Archives
Research Libraries
Government Agencies
Private Companies

Common goal to preserve the Web

Collaboratively develop tools for the community such as the Heritrix Web crawler and the Open Wayback system for replay of Web archives.

Lead work on the formalization of the WARC File Format (ISO 28500:2009)

General Assembly & Web Archiving Conference 2022



The General Assembly (GA) and the Web Archiving Conference (WAC) 2022 will be hosted online by the Library of Congress and the IIPC. We will be organizing the GA **May 17-19** as a series of meetings for IIPC members and will accommodate all IIPC time zones. WAC will begin with a day of workshops on **May 23**, and continue with the conference from **May 24-25**. We will try to accommodate as many time zones as possible.



IIPC WEB ARCHIVING CONFERENCE
ONLINE 23-25 May 2022
[#WhyWebArchiving](#)

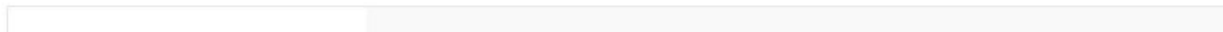


Call for
Proposals

[Click here](#) for information about the 2021 virtual GA/WAC hosted by the National Library of Luxembourg.

[Click here](#) to view recorded sessions from the 2021 GA/WAC.

Preliminary Schedule



Searching Across Many Web Archives

<http://timetravel.mementoweb.org/>



time travel

http://			
2008-11-21	13:16:29	Find	Reconstruct

Find Mementos in Internet Archive, Archive-It, British Library, archive.today, GitHub and [many more!](#)



http://www.unt.edu/

2019-10-24

01:01:23

Find

Reconstruct



24 Oct 2019 00:43:06 GMT



Mementos closest to the requested date 24 Oct 2019 01:01:23 GMT

[Internet Archive Memento, 1 hour before](#) Embed

<https://web.archive.org/web/20191024004306/http://www.unt.edu/>
24 Oct 2019 00:43:06 GMT [<1 hour from requested date]

[Previous Memento](#)

24 Oct 2019 00:43:05 GMT
[<1 hour]

[First Memento](#)

10 Apr 1997 02:35:43 GMT
[-22 years 201 days]

[All captures from Internet Archive between 1997 and 2019](#)

[Archive-It Memento, 1 day before](#) Embed

<http://wayback.archive-it.org/all/20191022052025/https://www.unt.edu/>
22 Oct 2019 05:20:25 GMT [-1 day from requested date]

[Previous Memento](#)

22 Oct 2019 05:20:13 GMT
[-1 day]

[First Memento](#)

17 Feb 2006 01:23:48 GMT
[-13 years 251 days]

[All captures from Archive-It between 2006 and 2019](#)

[Arquivo.pt Memento, 2 years 112 days before](#) Embed

<https://arquivo.pt/wayback/20170703070339/https://www.unt.edu/>
03 Jul 2017 07:03:39 GMT [-2 years 112 days from requested date]

[Previous Memento](#)

07 Jun 2017 04:25:45 GMT
[-2 years 138 days]

[First Memento](#)

07 Nov 2013 22:57:54 GMT
[-5 years 351 days]

[All captures from Arquivo.pt between 2013 and 2017](#)

[archive.is Memento, 4 years 141 days before](#) Embed

<http://archive.md/20150605134419/http://www.unt.edu/>
05 Jun 2015 13:44:19 GMT [-4 years 141 days from requested date]

[Previous Memento](#)

04 May 2013 19:50:57 GMT
[-6 years 173 days]

[First Memento](#)

25 May 2012 07:07:36 GMT
[-7 years 152 days]

[All captures from archive.is between 2012 and 2015](#)

[Web Citation Memento, 7 years 195 days before](#) Embed

<http://www.webcitation.org/66re3vYCz>
12 Apr 2012 07:57:18 GMT [-7 years 195 days from requested date]

experience web time travel



install Memento for Chrome

enable web time travel



MEDIAWIKI

install Memento for MediaWiki

say **no** to "404 Not Found"



Robust Links

use **links that refuse to die**



Using Web Archives

There are many challenges

(possible research areas)

- The web keeps evolving technology wise
- A move from a document platform to applications in the browser
- Replay of captured websites often breaks
- You can accidentally “leak” onto the live web from a web archive
- Hard to know if the view of the website you see actually existed
 - cnn.com with the wrong images from a different story in time.
- Web archives are far from complete
- Web archives need better tools for understanding what they hold.

Research at UNT

National Digital Information and Infrastructure Preservation Program: Web-at-Risk (2005-2008)

- Web-at-Risk: Preserving Our Nation's Cultural Heritage - UNT Digital Library
- Seneca, T. (2009). The Web-at-Risk at Three: Overview of an NDIIPP Web Archiving Project. *Library Trends*, 57(3), 427-441. <http://hdl.handle.net/2142/13606>

Expanding Collection Development Practices to Web Archives (EOTCD) (2009-2013)

- Hartman, C. N., Murray, K. R., & Phillips, M. E., (2013). Classification Of The End-Of-Term Archive: Extending Collection Development Practices To Web Archives. <https://digital.library.unt.edu/ark:/67531/metadc152437/>
- Murray, K. R., & Hartman, C. N. (2012). Classifying the End-of-Term Archive. IS & T--the Society for Imaging Science and Technology Archiving Conference, 2012, Copenhagen, Denmark. <https://digital.library.unt.edu/ark:/67531/metadc93305/>
- Phillips, M. E., & Murray, K. R. (2013). Improving Access to Web Archives through Innovative Analysis of PDF Content. IS & T--the Society for Imaging Science and Technology Archiving Conference, 2013, Washington, D.C., United States. <https://digital.library.unt.edu/ark:/67531/metadc155622/>

UNT Research (cont.)

Programmatic Extraction of 'Documents' from Web Archives (2017-2020)

- Phillips, M. E. & Caragea, C. (2017) Programmatic Extraction of 'Documents' from Web Archives <https://www.ims.gov/grants/awarded/lg-71-17-0202-17>
- Fox, N. T., Phillips, M. E., & Tarver, H. (2020). Programmatic Extraction of 'Documents' from Web Archives: Identifying Document Characteristics from Content Selector Interviews. <https://digital.library.unt.edu/ark:/67531/metadc1757659/>
- Patel, K., Caragea, C. Phillips, M. E., & Fox., N. (2020). Identifying Documents In-Scope of a Collection from Web Archives. Proceedings of the ACM/IEEE Joint Conference on Digital Libraries, 167-176 2020. <https://doi.org/10.1145/3383583.3398540>

Web Archive Quality Assurance (ongoing)

- Reyes, B. (2018). A Grounded theory of information quality in web archives. <https://digital.library.unt.edu/ark:/67531/metadc1248497/>

Spring 2022

INFO 5960.001 Section – Web Archiving

12 Students, primary from the Archival
Studies and Imaging Technology
concentration

Web archiving is an important component in preserving a national heritage for future generations.

Thank you.

Mark Phillips

mark.phillips@unt.edu

@vphill