Curators Needed

How Public Libraries are Bringing Community Members into their Web Archiving Practice

Karl-Rainer Blumenthal, Internet Archive
Emily Ward, East Baton Rouge Parish Library
Natalie Milbrodt, Queens Borough Public Library

#PDA18

#CommunityWebs

Community Webs:
Creating Community History Web Archives

Karl-Rainer Blumenthal
Web Archivist, Archive-It, Internet Archive

- > Web archiving - What is it?
- > Web archiving - Who does it?
- > What is the Community Webs project?
- > Community Webs Library Leaders

Web archiving is the process of **collecting, preserving,** and enabling **access** to web-published materials.

Average lifespan of a webpage

92 days

Web archiving

Goal: recreate the same experience you would have if visiting the live site on the day it was archived.

Process: capture as much information as possible from the original web resources and then these resources are used to playback the archived version of the site.

Brozzler
Heritrix
HTTrack
warcprox
wget

ARC
WARC

Wayback Machine
OpenWayback
pywb (Python Wayback)
oldweb.today

Brozzler
Heritrix
HTTrack
warcprox
wget

ARC
WARC

Archive-It
NetarchiveSuite (DK/FR)
PANDAS (AUS)
Web Curator (UK/NZ)
Webrecorder

Wayback Machine
OpenWayback
pywb (Python Wayback)
oldweb.today

Karl has dysentery.

Press ENTER to size up the situation

Date: July 28, 1848
Weather: warm
Health: poor
Food: 418 pounds
Next landmark: 35 miles
Miles traveled: 795 miles

Press SPACE BAR to continue

Internet Archive: Wayback Mach

+

← → ↺ 🏠

🔒 https://archive.org/web/

⋮ 🔔 🔍 Search

📄 📖 📽 🎧 📁 🖼

🏛

👤 Karl_Blum...

⬆

ABOUT CONTACT BLOG PROJECTS HELP DONATE JOBS VOLUNTEER PEOPLE

INTERNET ARCHIVE

WayBackMachine

Explore more than 324 billion web pages saved over time

BROWSE HISTORY

Is the Wayback Machine useful to you? Then keep it growing. Your gift matched today!

DONATE

buzzfeed.com
Oct 21, 2013 17:32:21

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

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 \(Dec 31, 2014\)](#)

It's like a whole other country

TravelTex.com

home

hot topics

events

free stuff

tours

postcards

reservations

QUICK SEARCH

GO!

TEXAS
TOUR
GUIDE

where do you want to go?

Anywhere

what are you interested in?

Historic

Howdy!

The page you requested was not found.
But don't worry, it's easy to find what you're
looking for by using Quick Search,
or click [here](#) to go to the home page.

A **web-based application** launched in 2006 that allows users to curate, manage, access and store collections of archived web content.

A **fully hosted** software - access and storage included, no hardware or installation needed.

Northeast Texas News and Organizations

Collected by: [Texas A&M University - Commerce](#)

Archived since: May, 2013

Description: A collection of newspaper and organizational websites from the Northeast Texas region.

Subject: [Spontaneous Events](#), [Society & Culture](#), [Electronic newspapers](#), [Internet-Periodicals](#), [Associations](#), [Institutions](#), etc.

Creator: [Texas A&M University - Commerce](#), [University Archives and Special Collections](#)

Type: [Interactive resource](#)

Rights: Texas A&M University-Commerce does not claim copyright to any of the materials within this web collection.

Collector: [Texas A&M University-Commerce](#)

Narrow Your Results

Creator

Sort By: Count | (A-Z)

The Sulphur Springs News-Telegram (2)
eParis Extra! (2)
Mount Vernon Optic-Herald (1)
North Texas e-News (1)
The Herald Banner (1)

Publisher

Sort By: Count | (A-Z)

eParis Extra! (2)
Mount Vernon Optic-Herald (1)
North Texas e-News (1)
The Herald Banner (1)
The Sulphur Springs News-Telegram (1)

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

Enter search terms here

Search

Sites

Search Page Text

Page 1 of 1 (9 Total Results)

Sort By: Title (A-Z) | Title (Z-A) | URL (A-Z) | URL (Z-A)

Title: The Herald Banner, Greenville, TX

URL: <http://heraldbanner.com/>

Captured 87 times between May 3, 2013 and Jul 1, 2016

Videos: 4 Videos Captured

Creator: The Herald Banner

Publisher: The Herald Banner

Rights: © Community Newspaper Holdings, Inc

Title: North East Texas, Chapter 416, Air Force Association

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

Captured 9 times between Apr 10, 2014 and Apr 10, 2016

Videos: 1 Videos Captured

Rights: © North East Texas, Chapter 416

URL: <http://www.commercejournal.com/>

Captured 126 times between Dec 9, 2016 and Mar 30, 2018

Videos: 241 Videos Captured

You are viewing an archived web page, collected at the request of [Texas A&M University - Commerce](#) using [Archive-It](#). This page was captured on 14:13:20 May 03, 2013, and is part of the [Northeast Texas News and Organizations](#) collection. The information on this web page may be out of date.

See [All versions](#) of this archived page. [Metadata](#)

eParisTexas.com

EXTRA!

Search...

HOME

COLUMNS

PARIS FLASH

BUSINESS

SCHOOLS

SPORTS

REAL ESTATE

OBITUARIES

9-1-1

CONTACT

Horses fill the streets, officially kicking off rodeo weekend – Click for gallery!

The Paris Rodeo & Horse Club has officially kicked off its annual rodeo weekend with the 2013 Rodeo Parade on Thursday before the rodeo...click to read more and view gallery.

READ MORE

Welcome to the eParis Extra!

Your source for local news and media coverage. You can find breaking news, upcoming local

eParisExtra.com

18,728 likes

Archive-It partners

Organizations with web archiving programs by type
NDSA, Web Archiving in the United States: A 2016 Survey

Perceived roadblocks

- > Lack of funding or institutional buy-in
- > Lack of training resources
- > Staff time requirements
- > Technical requirements

Community Webs

A two-year IMLS and Kahle-Austin grant funded program to empower public libraries to build collections of historically-valuable, web published materials documenting their local communities.

Cohort members:

- > Participate in online and in-person meetings
- > Participate in and provide feedback on an online course on web archiving and community engagement.
- > Collaborate with their local communities to build web archive collections using Archive-It (~ 2 hours per week)
- > Contribute to blogs and other outreach to the public library community

110 applications
from public libraries
across the country

A cohort of 28
small, medium, and
large public libraries

15 IMLS
Participants

13 Kahle Austin
Participants

- ★ Athens Regional Library System
- ★ Birmingham Public Library
- ★ Brooklyn Collection, Brooklyn Public Library
- ★ Buffalo & Erie County Public Library
- ★ Cleveland Public Library
- ★ Columbus Metropolitan Library
- ★ DC Public Library
- ★ East Baton Rouge Parish Library
- ★ Forbes Library
- ★ Grand Rapids Public Library
- ★ Henderson District Public Libraries
- ★ Kansas City Public Library
- ★ Lawrence Public Library
- ★ County of Los Angeles Public Library
- ★ Marshall Lyon County Library

- ★ Denver Public Library
- ★ Metropolitan Library System (Oklahoma)
- ★ New Brunswick Free Public Library
- ★ Patagonia Library
- ★ Pine River Library
- ★ Pollard Memorial Library
- ★ Queens Public Library
- ★ San Diego Public Library
- ★ San Francisco Public Library
- ★ Schomburg Center for Research in Black Culture
- ★ The Urbana Free Library
- ★ West Hartford Public Library
- ★ Westborough Public Library

Program outcomes:

- > A corpus of publicly available web archiving documentation and training resources for public libraries
- > A cohort of leaders in the field of local history web archiving
- > 40+ Terabytes of archived local history web content

Archiving the Web @ East Baton Rouge Parish Library

Emily Ward, Digital Archivist
East Baton Rouge Parish Library
eward@ebrpl.com

Overview

- Introduction to East Baton Rouge Parish Library (EBRPL)
- EBRPL's first web archive collections
- Collection development
- Outreach Models for Public Libraries
 - Methods
 - Challenges
- Questions

East Baton Rouge Parish Library

- Main Library and 13 branch or regional libraries throughout East Baton Rouge Parish
- Plans to add a South Branch are in the works.
- The library system is staffed by 540 City-Parish employees.
- The Library is open a cumulative 958 hours a week.
- In 2017, EBRPL branches were visited 2,212,004 times by registered library patrons.
- A total of 885,784 reference queries were answered.

Special Collections

Located at the Main Library

Current staff:

- 2 full-time archivists
- 1 part-time student worker
- 1 full-time genealogist
- 4 full-time techs
- 1 part-time librarian
- 1 moonlighter

Web archivists: → Just me :)

- About 15% of each working week is devoted to tasks related to web archiving.

Why did we decide to archive the web?

Initial Takeaway

- Can't archive everything
- Stay vigilant
- Time, patience, and commitment!

Collection Development Policy

Mission of the Bentley Historical Library

The Bentley Historical Library was established in 1935 by the University of Michigan Regents to carry out two functions: to serve as the official archives of the university and to document the history of the state of Michigan and the activities of its people, organizations and voluntary associations. The library is currently comprised of two departments: Collections and Curation.

The Bylaws of the University of Michigan Board of Regents (Sec. 12.04) establish the mission of the Bentley Historical Library:

The Michigan Historical Collections will be maintained for the purpose of collecting, preserving, and making available to students manuscripts and other materials pertaining to the state, its institutions, and its social, economic, and intellectual development. The University Archives and Records Program shall be maintained to collect, preserve, and make available the records generated by the university in the conduct of its business.

Mission of the East Baton Rouge Parish Library Special Collections Department

The mission of the Baton Rouge Room Collection is collect, manage, preserve, and provide access to items that represent significant historical actions of local governments, businesses, residents and institutions of the City of Baton Rouge and East Baton Rouge Parish. These items include but are not limited to photographs, manuscripts, documents, periodical publications, audiotapes and memorabilia.

C.P.D. Points of Interest- Intellectual Property Rights

Intellectual Property Rights excerpt:

- All robot.txt exclusions with exception to those used for social media will be respected. If EBRPL Web Archive wishes to pursue a crawl that utilizes robot.txt it will contact the content creator to negotiate access to the site.
- EBRPL will send out notices to content creators when their sites will be crawled to give ample notice and to allow content creators to set access restrictions to certain parts or all of their content.

C.P.D. Points of Interest- Scope and Content

Collections will build off of physical collections present in the Baton Rouge Room Archives to fill holes in collection content and to add additional context. Furthermore, the web archivist will actively seek out web content from people and organizations that have not been already included in the physical holdings of the archive. Content will cover the following areas:

- Residential life
- Geography
- Economic Activity
- Government Activities
- Business Activities
- Cultural Events
- Spontaneous Events
- Educational Institutions
- Local Art
- Local Community Groups/Organizations
- Social Movements/Grassroots Movements
- Race and Ethnicity
- Minorities such as LGBTQIA persons or groups
- Religious Groups
- Transportation
- Leisure
- Labor

C.P.D. Points of Interest- Scope and Content Continued...

- The institutional record
- Closing the gap

Outreach for Public Libraries

You are viewing an archived web page, collected at the request of [East Baton Rouge Parish Library](#) using [Archive-It](#). This page was captured on 18:52:00 Mar 31, 2018, and is part of the [Arts and Culture](#) collection. The information on this web page may be out of date. See [All versions](#) of this archived page. [Metadata](#) [Enable QA](#)

phoebe

[submissions](#)

[midrange a blog](#)

[print issues](#)

[online issues](#)

[newsletter sign-up](#)

[about](#)

[f](#)

[t](#)

[q](#)

cesarean vegetarian

© May 4, 2017 [46.2 Poetry](#), 46.2 Spring 2017, Poetry

Dylan Krieger

and then the oinkles wriggle out of me in all their mammary glamor, in their interminable gallows laugh of nothing left but bereft receptacles and pentacles that lift, on top of every barren birth you pinned to me, there was also the matter of your hips' wanting width, of sisters destined to linger down the canal of your bone trap, unknissed, in that vice-grip is where they will forever simmer and twitch, shimmer and glitch until the whole joke turns into just a tiptoe up the monolith of mother-bitch, the hard dissociation switch that leaves us itching to stitch pity into the space between our lips, so now our childhood chicken dinners become beloved pets, the ancient jainist wisdom of our ancestors rolling spliffs on house arrest, and somewhere between the changing station and the sacrificial altar, we start to wonder whether all is just another word for nothing, whether the parent of compassion is just a jackal with a beehive in its ribs

You are viewing an archived web page, collected at the request of [East Baton Rouge Parish Library](#) using [Archive-It](#). This page was captured on 21:29:29 Apr 08, 2018, and is part of the [Arts and Culture](#) collection. The information on this web page may be out of date. See [All versions](#) of this archived page. [Metadata](#) [Enable QA](#)

[ABOUT US](#) [ADVERTISE](#) [CAREERS](#) [EVENTS](#) [CONTACT US](#)

[Q](#)

225
COMMUNITY

[FOOD & DRINK](#)

[STYLE & HOME](#)

[THINGS TO DO](#)

[OUR CITY](#)

[BEST OF 225](#)

[PHOTOS & VIDEO](#)

[NEWSLETTERS](#)

Resources

- [Baton Rouge Room: East Baton Rouge Parish Library Web Archiving Initiative](#)
- [EBRPL Web Archive on Archive-it](#)
- [East Baton Rouge Parish Library Web Archives Collection Development Policy](#)
- [Bentley Historical Library Web Archives: Collection Development Policy](#)

Thank you!

Emily Ward

eward@ebrpl.com

225-231-3752

Queens Library

Natalie Milbrodt

Coordinator, Metadata Services Division

Director, Queens Memory Program

Queens Borough Public Library, New York City

Metadata Services at Queens Library

Cataloging

Digitization

Queens Memory

Queens Memory

Broad Channel Community Library, April 2013

Interviewer Training Workshop, September 2013

Local History + Outreach

Volunteers in the Vaults, July 2017

Wikipedia Edit-a-Thon, December, 2013

Web Archiving Stakeholders at Queens Library

Communications

**Government
Affairs**

Archives

Sites We Archive

- Queens Library websites
- Queens Library press clippings and social media
- Social media (Flickr, Instagram) for elected officials
- Local newspapers
- Local history organizations
- Local history websites of note made by private individuals

Creating a **Collection Development Policy**

- Using examples
 - Columbia
 - Stanford
 - Library of Congress
 - East Baton Rouge Public Library!
- Stating our scope and selection criteria
- How we will handle robots.txt exclusions

Public Programming Plans in Development

Developing workshops for **3 audiences**:

1. MLIS candidates in Library Studies program
2. Middle or high school Social Studies students
3. Cultural heritage professionals in Queens

Workshops for MLIS Students

Session One: 90 minutes

- Intro to Web Archiving and the Internet Archive
- Queens Library web archiving as a case study: How we decided what to archive and how we manage our web archiving
- Queens Library's participation as a Library Lead in the Community Webs project, funded by IMLS
- A peek at the Archive It backend with Queens Library's existing crawls
- Lesson plan ideas for teaching a public program to adults or youth in your library or off-site at a school or historical society
- Setting up a new crawl that we will explore in the next session

Session Two: 90 minutes

- Exploring the new crawl we set up in the last session (going back and forth between the Archive It backend and the Wayback Machine)
- How did the settings impact the resulting archived page?
- How much space is it using?
- What information did we capture that we might not want?
- How frequently should we do this same crawl?
- Changing the settings to modify frequency and depth of the crawl

Session #1: March 5, 2018

Session #2: March 12, 2018

Different content for different audiences...

Middle and High School Students:

- What are the qualities of a trustworthy information resource?
- How quickly does website content disappear and why?
- Help the library select our seeds

Cultural Heritage Professionals:

- What is web archiving?
- How can you search and use archived websites?
- What website content is your organization or community stakeholder creating that will disappear?
- Help the library select our seeds

Thanks!

Learn more about Community Webs and more Archive-It partners:
archive-it.org

#PDA18

#CommunityWebs