

# K-12 Web Archiving Program Teacher Guide

## Program Description

There is a growing awareness among libraries, archives, and other cultural institutions of the importance and urgency of preserving the often transitory digital cultural artifacts distributed over the Web. But so far, most decisions about what Web sites will live into the future have been made by adults, and reflect adults' sensibilities about which records are historically significant. In order to broaden this perspective, the Internet Archive, the Library of Congress and California Digital Library launched the K-12 Web Archiving Program.

The K-12 Web Archiving Program has two primary objectives:

1. Archive the Web from the perspective of students to ensure that at risk digital content that is important to them is captured and preserved forever.
2. Stimulate students to think about history and actively participate in selecting the primary sources of today for historical research tomorrow.

The program explores archiving the Web from students' perspectives. Now in its second year, the program uses the Archive-It service from the Internet Archive (see Appendix 1) to create "time capsules" of born digital content available via the Web chosen by students to represent their world. Visit <http://archive-it.org/k12/> to see comments and collections from previous participants.

By allowing students to identify sites that will be preserved, the program gives K-12 students a chance to document their cultural history and the world that's important to them. Unlike time capsules of tangible objects, which usually remain hidden for decades or centuries, the resulting Web collections will be immediately visible and publicly accessible, with full text search for study and analysis. Stimulating young people to think about history in the context of their own lives will enrich their study of history, provide an opportunity to actively engage in selecting the matter of history in the future, and help students begin to grasp the tremendous challenges presented by a world in which information can be both generated and removed with a key stroke.

## Participants

- Students from schools across the country (5<sup>th</sup> grade to 12<sup>th</sup> grade)
- The group of students at each school may represent a class, student club, or other selected

group.

- Each group will be coordinated by a teacher and a student leader.

### **Selection Criteria**

- Teachers and students should be comfortable using the Web and Web-based tools.
- Teachers and students should be able to dedicate several hours per month to the project (see timeline below).
- Teachers and students must have an interest in history, culture, and technology.
- Teachers and students must have a strong understanding of primary sources and the role they play in documenting cultural heritage.

### **Timeline**

- August 14: deadline for submissions
- August 24: schools are selected and notified. Teachers gather participants and discuss parameters.
- September 28 - October 9: Introductions and training on the program, the web application and primary sources. (2 -3 hours total, time to be spread out)
- Schools define collections, gather Web sites, draft descriptive statements (7-9 hours total, time to be spread out)
- Thursday October 9: kick off meeting (via web conference) (1 hour)
- Weeks of October 12: Schools start crawling
- October through March: Schools continue to crawl web content and manage collections. (can be anywhere from 30 minutes to 1 hour per week)
- November and January: check-in conference calls (1 hour each)
- Early 2010: K-12 website (work in progress) is available online
- March 27, 2010: last day of crawling
- April 2010: evaluation process and completion with online survey (2 hours)
- Beginning of May 2010: final meeting via web or teleconference (1 hour)
- May 2010: K-12 website complete

### **Collection Parameters**

- Collections of Web sites that are primary sources reflecting the students' lives and interests. (See examples in Appendix 2)
- Selection criteria are purposely left broad, although several popular and very large websites need to be excluded (details to be provided)
- Each school will be responsible for selecting and curating at least one collection.
- Up to 300 Web sites (seeds) can be chosen for inclusion in each school's collection(s).

### **Student Responsibilities**

- Identify primary source sites for the "time capsule."
- Work with teacher to define a collection(s) and provide an overall description of the

“time capsule/collection.”

- For each site chosen, provide a brief description of the site and why the students chose the site to reflect their lives or interests.
- Answer a short online survey at the end of the project.

### **Teacher Responsibilities**

- Serve as facilitator for the student collection, assisting students in establishing a process for deciding on the sites to be included, and ensuring that the content selected does not include pornography, discriminatory materials, or gratuitous violence.
- Work with students to prepare an overall description of the “time capsule” or collection.
- Enter, or work with student leader to enter, URLs for selected sites into Archive-It.
- Answer a short online survey at the end of the project.

### **Continuing (after May 2010)**

- Teachers are encouraged to have students revisit the sites that they have archived to see which have disappeared, and how the sites have changed over time.
- Teachers may be called upon for input into developing a broader range student-based Web capture program.
- Teachers and students may be called upon for testimonials of their experiences.

### **About the Partners**

It is at the very heart of the missions of the Internet Archive, CDL, and the Library of Congress to serve as stewards of content that has current and future value to the nation. In addition, all three institutions are committed to serving the K-12 community, particularly as it involves the use of primary source material in the understanding of history.

#### **Internet Archive**

Archive-It is a subscription service developed in 2005 by the Internet Archive to help partners capture manage and archive born digital content. The Internet Archive is a 501(c)(3) non-profit that was founded in 1996 to build an 'Internet library,' with the purpose of offering permanent access for researchers, historians, and scholars to historical collections that exist in digital format.

Internet Archive has the largest public web archive, located at (<http://www.archive.org>) comprising 150+ billion pages, and over 65 million websites in 37 languages for a total of 3 petabytes. In late 1999, the organization started to expand to include non born digital material. Now the Internet Archive includes over 200 terabytes of texts, audio, moving images, and software along with our web archive.

#### **Library of Congress**

For more than fifteen years, the Library of Congress has helped teachers integrate its unique primary sources into school curricula. The Library's Web site for teachers, [loc.gov/teachers](http://loc.gov/teachers), provides ready-to-use classroom materials and professional development opportunities, as well as highlighting some of the most interesting and useful primary sources from the Library's digital collections. Since 2001, the Library has also been involved in archiving Web sites. This program brings together elements of the Library's Teaching with Primary Sources and digital preservation programs.

### **California Digital Library (CDL)**

CDL is the University of California's system wide digital library that supports all 10 UC campuses in creating, managing, sharing and preserving their collections. In addition, through the Calisphere website (<http://www.calisphere.universityofcalifornia.edu/>), CDL provides free access to primary source materials designed around the needs of K-12 educators. Calisphere offers more than 170,000 primary sources from the libraries and museums of the UC campuses and cultural heritage organizations across California.

### **Contact Information**

- Application support: Archive-It
  - Molly Bragg [mbragg@archive.org](mailto:mbragg@archive.org) 415-561-6799 x6
  - Kristine Hanna [kristine@archive.org](mailto:kristine@archive.org) 415-561-6799 x5
- Program support: Library of Congress
  - Cheryl Lederle-Ensign [cled@loc.gov](mailto:cled@loc.gov) 202-707-3550

## Appendix 1

Internet Archive's subscription service, Archive-It, enables institutions to harvest, manage, search and preserve born digital web based content.

The Web based application features a user friendly interface allowing subscribers to easily set up and catalog collections, monitor crawling and make adjustments post crawl. Collections are hosted at the Internet Archive data center, accessible to the public with full text search.

Archive-It is designed to fit the needs of many types of organizations and individuals, including state archives, libraries, academic institutions, non profits, museums, historians, and independent researchers. The 100+ partners include: state archives, university libraries, federal institutions, state libraries, non-government non-profits, museums, historians, and independent researchers.

For more information regarding the Archive-It service, please visit [www.archive-it.org/onesheet.pdf](http://www.archive-it.org/onesheet.pdf)

## Appendix 2

Example of the types of site that may be appropriate for this project:

<http://inprogress.typepad.com/kidsforobama/> Blog encouraging young people to voice their thoughts in the 2008 elections

<http://www.ifbf.org/> Iowa Farm Bureau

<http://www.ifad.org/> Enabling poor rural people to overcome poverty

<http://www.insects.org/> Insects on the Web

<http://www.wikisky.org/> Astronomical Photographs and Information

<http://www.promspot.com/> Prom Planning Information

Examples of the types of sites that show an adult perspective on what kids want, or are compilations of sources and are NOT the target for this project:

<http://clerkkids.house.gov/laws/InspectLaw/index.html>

<http://www.ohiomemory.org/>

<http://encarta.msn.com/>