

Wisconsin Digital Archives

Guidelines for State Agencies

Background

Wisconsin statutes, as amended by 1991 Wisconsin Act 285, require Wisconsin state agencies to send copies of their publications to the Wisconsin Reference and Loan Library for distribution to libraries through the Wisconsin Document Depository Program. Many of these publications are now born-digital, that is, published only on the Web. As more documents are provided exclusively in digital formats, particularly on the Internet, the requirement to comply with this law has become more challenging.

The Wisconsin Digital Archives provides state agencies a means to continue to fulfill their statutory obligation to participate in the Wisconsin Document Depository Program with electronic formats. **The Wisconsin Digital Archives is not an indication that state agencies should stop printing, drastically reduce printing or cease participation with print documents in the Wisconsin Document Depository Program.** Print is still the most stable format for long-term preservation.

The Wisconsin Digital Archives is a collaborative effort involving the Wisconsin Reference and Loan Library, the Wisconsin Historical Society, the Wisconsin Legislative Reference Bureau, the Wisconsin Department of Transportation, the University of Wisconsin-Madison, and the Wisconsin State Law Library.

Primary Goals

- To provide authoritative permanent electronic access to select web content originally accessed from Wisconsin state agency websites.
- To provide the opportunity for state agencies to continue to fulfill their statutory obligation to participate in the Wisconsin Document Depository Program with electronic formats.
- To provide all libraries, regardless of depository status, greater access to state government information in electronic formats.

What's in the Wisconsin Digital Archives?

Web content selected for digital preservation is consistent with the types of publications and documents previously collected in the print format. The Wisconsin Digital Archives preserves content from state agency web sites that is intended for public use and provides public perspective on state government programs. It is not the intent of the Wisconsin Digital Archives to collect content that might be considered public records. The Wisconsin Digital Archives is not digitally archiving entire state agency websites

How Does the Wisconsin Digital Archives Work?

Portions of state agency websites are selected to be spidered by a web harvester. The spider captures all links needed to replicate portions of state agency webpages to maintain the overall look and feel of the webpage at the time it was crawled. Full functionality of the webpage is provided through a stable digital archives URL. Digital archives URLs are persistent and provide a snapshot of what the state agency webpage looked like at the time it was spidered.

Digital archives URLs are described and cataloged based on author, content and subject matter and are designed to reside in an OPAC (Online Public Access Catalog) to provide a full historical range of state agency publications regardless of format. Digital archives records are being made available through OPACs such as WISCAT, MadCat, WorldCat and LRBCat.

Monthly digital archives lists of fully cataloged digital archive records are distributed to designated depository libraries and are made available online, <http://dpi.wi.gov/rl/indship.html> . The first digital archives shipping list was released in Oct. 2005. There has been a monthly list of new records since the first list was released.

Benefits of OPAC Access

There are many advantages to providing access to digital archives records in an OPAC setting:

- Provides search benefits and convenient access to state government information in a controlled environment
- Provides advanced authority control, refined subject access, and multiple field searching that makes a library catalog a much more precise search tool than a web search engine
- Provides convenient access to digitally archived information one click away from the catalog record
- State government information co-exists regardless of format in a repository setting providing full historical range of state agency publications
- Provides 24/7 access and convenience to unlimited users

Using WISCAT

The Wisconsin Dept. of Public Instruction, Division of Libraries, Technology and Community Learning has facilitated resource sharing in Wisconsin for twenty years with the production of a statewide union catalog of library holdings called WISCAT. The current edition of WISCAT, one of the largest physical statewide union catalogs in the country, contains 7 million titles and more than 32 million holdings from over 1200 contributing libraries. The catalog contains material in all formats, including books, serials, videorecordings, sound recordings and other audiovisual formats. WISCAT also contains catalog records for both print and digitally archived state government publications providing a full historical range of state agency publications.

To see a complete listing of digitally archived content that has been cataloged, go to WISCAT at www.wiscat.net. Click **Advanced Search** and enter the term *digitalarchive* as one word under **All Headings**. Press [Enter] to get a complete listing. To narrow your search by agency, enter the state agency as the Author. Do not use the Scoping feature in WISCAT to narrow search by library. Digital archives records do not have holdings because the catalog record is for electronic information. There is nothing physical to hold.

The screenshot shows the WISCAT search interface. At the top, there is a blue header with the text "Search For". Below this, there are three search input fields, each followed by a dropdown menu set to "All Headings". The first field contains "digitalarchive", the second contains "public instruction", and the third is empty. Between the first and second fields, and between the second and third fields, there are radio buttons for "And", "Or", and "Not", with "And" selected. Below the search fields is a blue button labeled "Clear Search". At the bottom of the interface, there is a "Sort By" dropdown menu set to "Alphabetically by Title", and two buttons labeled "Search" and "Reset".

Limitations of Current Digital Preservation Technology

There are technical limitations to what digital preservation technology can preserve at this time. As preservation technology evolves, the list of accepted formats and preservation capabilities will continue to expand. If you have questions regarding web content compatibility with the Wisconsin Digital Archives please contact Abby Swanton, abbigail.swanton@dpi.wi.gov.

The spider used to digitally archive web content is dependent upon following the domain structure of a specified webpage. All content that is captured in a digital archives URL must be from the same domain. For example if the Wisconsin Consumer Guide to Health Care were to be digitally archived, the spider would start from the following web address <http://dhfs.wisconsin.gov/guide/index.htm> and would crawl down a specified number of clicks to capture all links that are part of the <http://dhfs.wisconsin.gov> domain from the designated starting point. The spider can drill down a

maximum of 5 levels. The spider does not crawl up to addresses higher/shorter than the starting address. Any links that direct users to external sites will not be captured as part of the digital archives URL. A redirection notice will notify users when they click on any link that is not part of the digital archives URL. The spider respects robot exclusions and does not capture global navigation webpage features.

Currently any content that is accessible via the http protocol can be digitally archived successfully. Examples include formats such as Microsoft files, ZIP files, GIF, JPG, TIF, TXT and PDF. File formats that are served up via a protocol other than http, such as FTP, audio and video files that require a streaming server and databases, are not supported. JavaScript and embedded PDFs are also problematic and results will vary. The content of each individual digital archive URL must be under 4 GB.

How to Participate in the Wisconsin Digital Archives

Contact the Wisconsin Digital Archives to submit web content from state agency websites for digital preservation. This includes new content, content that has been significantly updated or revised, or content scheduled to be removed from a state agency website.

- Send the URL (Internet Address) of the web content to abbigail.swanton@dpi.wi.gov
- Please be as specific as you can.

Library staff is available to assist state agencies analyze web content for compatibility with project guidelines. Due to limitations of digital preservation technology, not all content can be or will be digitally archived. The digital archives URL will be cataloged by library staff and made available to users worldwide in multiple OPAC databases. Reference librarians will assist users in the location and use of content that has been digitally archived.

More Information

Please visit our website to learn more about the Wisconsin Document Depository Program and the Wisconsin Digital Archives, <http://dpi.wi.gov/rll/inddep.html> . For more information or to schedule a demonstration of the digital archives, please contact:

Abby Swanton, Wisconsin Document Depository Librarian
Wisconsin Reference and Loan Library
2109 S. Stoughton Rd.
Madison, WI 53716
(608) 224-6174
Abbigail.Swanton@dpi.wi.gov