



April 11, 2013

SUBJECT: GUIDELINES FOR MAINTAINING AND PRESERVING RECORDS OF WEB-BASED ACTIVITIES

PURPOSE

The Delaware Public Archives (DPA) is issuing these “Guidelines for Maintaining and Preserving Records of Web-Based Activities,” for use by all agencies in state and local government in Delaware. The purpose of these guidelines is to provide agencies with a set of uniform best practices and statutory requirements for managing an agency’s website resources and the preservation, retention and access to public records.

Due to the highly technical nature of capturing website resources, it is highly advisable for Records Offices and Authorized Agents to seek assistance from the Information Technology (IT) professionals assigned to support their organization.

BACKGROUND

The use of Internet websites to distribute information, provide services, and transact business with the citizens of Delaware has become a preferred method of delivery of state and local governments. In many cases, the information provided is the electronic version of a previously published hard copy document (brochure, report, form, etc.); however, certain records published by government agencies on their websites are “born digital” and exist solely in electronic form for their entire lifecycle. These records may be stored in a multitude of locations and formats. The increasing use of websites adds yet another challenge to the myriad issues concerning electronic records management.

All provisions of the Delaware Public Records Law apply to web-based activities. Website source files meet the definition of a record in accordance with Title 29 §501-526, as this information, regardless of format, is “made or received pursuant to the law or ordinance in connection with the transaction of public business.” The key concept to recognize is that it is the record, its content, and its function that determine retention and disposition decisions, not the media format.

Delaware Public Records Law and Delaware Public Archive’s policy require state and local government agency heads to insure that the retention and disposition of these source files are planned, managed, and executed accordingly. Officials should also note that the requirement to manage an agency’s website resources is the same as for other recordkeeping systems - to ensure compliance with Delaware statutes concerning the creation of, retention of, and continuing access to public records. Since the charge of the DPA is to provide for long-term preservation and access to records of long-term historical value, these guidelines were created to provide government agencies with guidance in managing their website records.

The Guidelines consist of the following sections:

- I. Definitions
- II. Accountability Exposure Analysis
- III. Acceptable media
- IV. Procedures for Submission Web-Site Captures to DPA
- Appendix A. Accountability Exposure Analysis
- Appendix B: Website Description Form

I. DEFINITIONS

ASCII: American Standard Code of Information Interchange; it is the built-in binary code for representing characters in all computers except IBM mainframes.

Electronic record: means a public record that is stored, generated, received or communicated by electronic means for use by, or storage in, an information system or for transmission from one information system to another. The retention policy governing the maintenance of electronic records is the same as that used for records maintained in paper formats unless otherwise specified by an authorized retention schedule.

HTML: Hyper Text Markup Language; a non-proprietary file format for describing the structure of hypermedia documents - plain text (ASCII) files with embedded codes for logical markup, using tags to structure text into tables, interactive forms, headings, paragraphs, lists, and more. It can be created and processed with a wide range of tools from simple text editors to sophisticated authoring software.

Public Record: Delaware Code Title 29 §502(7) *“public record” means any document, book, photographic image, electronic data recording, paper, sound recording or other material regardless of physical form or characteristics, including electronic records created or maintained in electronic information systems, made, used, produced, composed, drafted or otherwise compiled or collected or received in connection with the transaction of public business or in any way related to public purposes by any officer or employee of this State or any political subdivision thereof.*

Snapshot: Used to save copies of a web page’s current memory state, along with the contents of all hardware registers and status indicators. It is periodically taken in order to restore the system in the event of failure. It is also used to save copy of a file before it is updated. It is part of a storage management program; snapshots enable previous versions of files to be brought back for review or to be placed back into use.

The Internet: a system of inter-connected computers using Transmission Control Protocol/Internet Protocol standards for transmitting and receiving digital data. The Internet consists primarily of the collection of billions of interconnected webpages that are transferred using Hypertext Transfer Protocol (HTTP), and are collectively known as the World Wide Web. The Internet also uses File Transfer Protocols and Simple Mail Transfer Protocols to transfer files and e-mail.

Websites: are collections of information, documents, and databases that are provided to a user community utilizing World Wide Web formats and protocols.

Web page: is a source file, maintained as ASCII text, provided by a file server, and subsequently executed on a local computer that uses HTML/XML markup languages and external software tools to produce a representation and provide meaning.

XML: Extensible Markup Language; it is used for defining data elements on a Web page and business-to-business documents. XML uses a similar tag structure as HTML; however, whereas HTML defines how elements are displayed, XML defines what those elements contain.

II. ACCOUNTABILITY EXPOSURE ANALYSIS

- A. **ACCOUNTABILITY EXPOSURE ANALYSIS.** DPA has included an Accountability Exposure Analysis (AEA) tool to assist agencies in determining the appropriate level of legal and/or managerial risk associated with the website. The outcome of this assessment should guide the agency’s record and information management actions. See Appendix A: Accountability Exposure Analysis
- B. **WHY TAKE WEBSITE SNAPSHOTS.** Snapshots will capture the look and feel of active agency websites at particular points in time, and ensure that DPA, per its charge in the Delaware Public Records Law, documents agency use of the Internet over time.
- C. **WEBSITE DESCRIPTION FORM.** DPA has prepared a standard description form (see attachment) that permits government agencies to easily capture information about the contents, format and technical characteristics of their websites. Submission of this descriptive information, along with copies of all active source files and either electronic or hard copy versions of relevant log files, will allow DPA to provide continuing access to and an historical perspective for the provision of government information and services. See Appendix B: Website Description Form.
- D. **FREQUENCY OF CAPTURE.** The frequency of website changes and revisions will vary greatly from agency to agency. Agencies that publish enduring versions of policies, publications or images may not see major changes to their website design for months at a time. However, agencies with more advanced services (interactive forms, streaming video, etc.) might undergo major changes a number of times each

year. Recognizing those differences, DPA recommends that website snapshots be taken at the time of each major version change to the website (different look, additional features, etc.) or at least once a year, whichever occurs first.

- E. WHAT TO INCLUDE IN THE WEBSITE SNAPSHOT. You should include all active documents available to the public that are located on the agency's web server, including copies of agency documents that exist in another form elsewhere EXCEPT:
1. Databases
 2. Files located on a web server external to the agency (e.g., another agency's website)

III. ACCEPTABLE MEDIA

To ensure ease of transfer and standardization of the media being submitted for preservation, DPA suggests that agencies adhere to the following media specifics when capturing their websites:

- A. Use a fresh Compact Disc-Read Only Memory (CD-R) or Digital Versatile Disc- Read Only Memory (DVD-R) with capacities of no more than 700 MB.
1. A gold reflective surface is required for CD and DVDs
 2. Source files should not be compressed.
 3. The CD/DVD's case should be appropriately labeled, but the CD/DVD itself should not be. A volume label is automatically created during initial use of the CD/DVD, and may be edited to reflect your agency's nickname/acronym, date, and/or disc number within the allotted 11 characters. Place this number on the case and all related transfer documentation.
 4. The CD/DVD writing should comply with the Joliet modifications to ISO 9660 specifications. These settings are available in your CD/DVD creation software options.
- B. External Hard Disk Drive
1. Source files should not be compressed.

IV. PROCEDURES FOR SUBMISSION OF WEBSITE SNAPSHOTS TO DPA

The media/documentation from the website snapshot may be transferred to DPA using the standard DPA Transfer Document (Form RM-3). Complete the form and submit it to DPA for approval. Once it has been approved, you will be contacted to make arrangements for pickup of the media/documentation and its transfer to DPA. A copy of the completed Transfer Document, showing the date of receipt by DPA, will then be returned to you for your file. Please see the [User's Guide for Transfer Documents](#) located on the DPA website, under For Government, Records Policies.

Effective: April 5, 2013

References:

[Delaware Public Records Law: Title 29, Chapter 5](#)
[Delaware Uniform Electronic Transactions Act: Title 6, Chapter 12A](#)
[DPA Policy User's Guide for Transfer Documents](#)
Appendix A. Accountability Exposure Analysis
Appendix B. Website Description Form

Delaware Public Archives

ACCOUNTABILITY EXPOSURE ANALYSIS

Accountability Exposure Analysis (AEA): AEA is the appraisal of the extent to which an agency may or may not be fulfilling its legal responsibilities under public records statutes or other obligations vs. the likely risk of litigation. The outcome of this assessment determines appropriate management and recordkeeping responses.

Low Level	Moderate level	High Level
<p>Characteristics:</p> <ul style="list-style-type: none"> ➤ Single website ➤ Copies of official publications ➤ Controlled postings ➤ Publications not controversial ➤ Publications never litigated and little or no legal risk ➤ Publications do not generate bad press ➤ Constituents are satisfied with website ➤ Little legislative interest 	<p>Characteristics:</p> <ul style="list-style-type: none"> ➤ Number of complex websites ➤ Many offices posting to the websites ➤ Overlapping and conflicting information ➤ Both static and interactive websites ➤ Unique original materials on websites ➤ Legal counsel was not consulted ➤ Some adverse and controversial reaction ➤ Variety of management controls 	<p>Characteristics:</p> <ul style="list-style-type: none"> ➤ More numerous & complex websites with sub-sites ➤ Contain bulletin boards, publications, unique original materials, hearings, and real time business ➤ Different administrative arrangements ➤ Public scrutiny; publications are controversial ➤ Substantial liability; frequently litigated ➤ Frequent press coverage; public interest groups monitor sites ➤ Active and vigilant legislative interest ➤ Public is clamoring for more
<p>Records Management Analysis: Satisfactory; no further measures necessary; periodic review of website for changes.</p>	<p>Records Management Analysis: Additional measures should be taken to reconstruct prior versions</p>	<p>Records Management Analysis: Requires precise reconstruction of exact copies of past contents</p>
<p>Recordkeeping Response:</p> <ul style="list-style-type: none"> ➤ Adequate documentation ➤ Records already in recordkeeping systems (paper and electronic) ➤ Snapshot of entire website to DPA at least once every year. 	<p>Recordkeeping Response:</p> <ul style="list-style-type: none"> ➤ Maintain directory listings of websites ➤ Snapshot of entire website to DPA at every significant version change or at least once every year, whichever occurs first ➤ Snapshots into the agency recordkeeping system and retention schedules applied 	<p>Recordkeeping Response:</p> <ul style="list-style-type: none"> ➤ Maintain regular sequential directory listings of website changes ➤ Save periodic snapshots of entire website (weekly, monthly) ➤ Snapshots of entire website transferred to DPA quarterly/biannually ➤ Snapshots into the agency recordkeeping system and retention schedules applied

DPA AEA Risk matrix based on the model developed by: McClure, Charles R. and J. Timothy Sprehe, Analysis and Development of Model Quality Guidelines for Electronic Records Management on State and Federal Websites, January 1998.

Delaware Public Archives

WEBSITE DESCRIPTION FORM

1. WEBSITE URL:	
2. AGENCY UNIT SUPPORTED BY WEBSITE:	
3. PURPOSE/FUNCTION OF WEBSITE:	
4. DATE SNAPSHOT WAS TAKEN:	
5. FILE FORMATS AND VERSIONS: <input type="checkbox"/> HTML: version(s) - <input type="checkbox"/> Portable Data Format (PDF): version(s) - <input type="checkbox"/> Text (ASCII) <input type="checkbox"/> Word processing software (Word, WP, etc.): version(s) - <input type="checkbox"/> Audio: format(s) and version(s) - <input type="checkbox"/> Video: format(s) and version(s) - <input type="checkbox"/> Image: format(s) and version(s) - <input type="checkbox"/> CGI Scripts - <input type="checkbox"/> Active X - <input type="checkbox"/> Java: version(s) - <input type="checkbox"/> Java Scripts - <input type="checkbox"/> Web browser version for which site is optimized - <input type="checkbox"/> Specialized formats (Please describe) - <input type="checkbox"/> Other (Please describe and note versions used) -	
6. TOTAL NUMBER OF MEDIA UNITS:	7. MEDIA UNIT VOLUME SERIAL NUMBERS:
8. TOTAL NUMBER OF FILES SUBMITTED:	
9. DOCUMENTATION INCLUDED: See instructions for guidance on what must be included. <input type="checkbox"/> Site map (e.g., hierarchical list of files, graphical representations, etc.) – MANDATORY <input type="checkbox"/> Other lists of files, if available (e.g., list of the forms and publications available on website)	
THIS DOCUMENTATION IS PROVIDED IN: <input type="checkbox"/> Paper <input type="checkbox"/> Electronic form (ASCII only) <input type="checkbox"/> Combination of paper and electronic	
10. RECORDS OFFICER: PHONE: EMAIL:	11. TECHNICAL CONTACT: PHONE: EMAIL:
12. FORM PREPARED BY: PHONE: EMAIL:	

INSTRUCTIONS FOR WEBSITE DESCRIPTION FORM

Block 1. Enter the top-level URL of the snapshot being taken [e.g., www.archives.state.de.us]

Block 2. Enter the name of the highest-level agency component the URL (Block #1) represents. If the snapshot is for the entire agency website, give the full name of the agency (e.g., Delaware Public Archives, State Treasurer, etc.).

Block 3. Describe the purposes of the website, the services provided, and any agency mission specific requirements met by the website, e.g.:

- to provide information to the public (note specifically if any information provided on the website is unique and therefore only obtainable or accessible from the website.)
- to serve as an interface for submissions of information from the public (describe – e.g., questions or general comments from users; submission of forms/data; rulemaking comments)
- to provide access to mission-critical databases (identify which ones)

Block 4. Provide the date the snapshot was taken.

Block 5. Place an “X” in the box of all the file formats present on the website. Identify the software and version used to create the files. This information is needed so that DPA can copy and preserve the files on appropriate media.

Block 6. Enter number of media units submitted.

Block 7. Enter the volume serial numbers that uniquely identify these media units. If the file is recorded on multiple volumes, enter the first volume serial number in this item and list others, in sequence. You do not have to enter this number if the media unit does not have a serial number.

Block 8. Enter the total number of files being submitted.

Block 9. You must provide a site map including URLs, or a similar comprehensive list of files including extensions. (A graphical representation, if already available, is desired.) Identify additional documentation that is included with your snapshot submission. You may print information from the website if it is already available there. Indicate whether the site map and other documentation is in paper or electronic form, or a mix of the two media. Electronic files must be in ASCII (text).

Block 10. The name of the records officer responsible for the website.

Block 11. The name, telephone number, and email address of the person who can provide an overview of the website (e.g., webmaster, director of communications, etc.).

Block 12. The preparer’s name, telephone number, and email address. If the records officer or website contact prepared the form, you may enter just the name.