

## Guidelines for Managing Web Site Content: Section 6.0

### Records Management for Web-based Records

#### 6.2 Guidelines for Managing Web-based Records

A single web site may be made up of a variety of resources with differing degrees of technical complexity. Applying the concepts of records scheduling, management, disposition, and preservation will increase in complexity in tandem with the increase in the technical complexity of the web site.

There is no generic solution for creating and maintaining records of web-based activity. The best option will depend on the outcome of an analysis of the particular circumstances. Each agency should assess a number of factors, including:

- The type of web-based activity information dissemination or transaction or both;
- The complexity, diversity and sophistication of the web-based activity;
- The frequency and regularity of change to the web-based activity;
- The frequency of challenges to the validity of the information on the website;
- The level of risk and public visibility of the agency;
- The agency's record retention requirements;
- The agency's technological environment; and
- The availability of resources.

##### 6.2.1 Static Web Sites

In its most basic form, a web site may be nothing more than a collection of static documents sitting in folders on a server and tied together with hyperlinks. The only interactivity provided by static sites is in the links that enable movement from one document to another or from one part of the site to another.

##### Scheduling

Static sites are relatively simple to schedule. Steps included in the process include:

1. Identify distinct record series
2. Is the series included on an existing agency retention schedule or in the General Schedules?
3. If the series is included on an existing schedule, you should:

Determine if the retention period is appropriate. Revise as needed

Determine if the web-based format is included on the retention schedule. If it is not, revise the retention schedule to include web-based format.

4. If the series is not included on an existing schedule, you should:

Determine if the series is maintained in other formats (paper or microfilm)

Determine appropriate retention period

Create and submit retention schedule

### **Implementing the Retention Schedule Requirements**

1. Work with webmasters and content creators to create procedures for updating files and maintaining deleted files for the appropriate retention period. Include appropriate metadata in web pages that includes the date the page was made live; the date the page was removed from the web; and the proposed disposal date.

2. Move updated and deleted files to alternative storage medium for the length of the retention period.

3. If the web site has a high rate of change it may be helpful to take periodic snapshots of the entire web site

(See "Technical Guidelines for Scheduling and Managing Web Sites" section 6.3 for more information)

### **6.2.2 Static Web Sites with Form-based Interactivity**

[This section does not refer to web based systems that facilitate a financial transaction between government and client(s).]

Many web sites utilize forms and back end information systems. These sites consist primarily of static html pages with simple form based interactivity. There are essentially two types of form-based interactive sites:

- Sites that collect information such as comments and requests from visitors, and
- Sites that provide a search interface for back-end information resources

Agencies should manage these sites as static web sites, but should take the additional steps of scheduling and managing:

The "back end" information system

The form itself

The human readable source code of the script or program which enables the forms functionality (again, not necessarily scheduled, but "managed")

### **Scheduling**

Evaluate each of these additional three elements of the web page and determine the appropriate retention period for each. Create and submit retention schedule.

## Implementing the Retention Schedule Requirements

1. Work with webmasters and content creators to create procedures for updating forms and scripts and for maintaining deleted files for the appropriate retention period.
2. Include appropriate metadata in files that includes the date the page was made live; the date the page was removed from the web; and the proposed disposal date.
3. Move updated and deleted files to alternative storage medium for the length of the retention period.

(See "Technical Guidelines for Scheduling and Managing Web Sites" section 6.3 for more information)

### 6.2.3 Web Sites Based on Dynamic Data Access

Web sites are sometimes used as front ends, or user interfaces, for accessing an agency's database(s). Site users search prepared lists or put together their own searches that, in turn, query the content of a database. The information returned from these queries is displayed as an HTML document to the user.

In many cases, documents exist as objects in a database. Each document will have its own unique identifier, usually reflected in the URL. This means that a user can bookmark the particular document and return to it later without reconstructing the original search query (provided the document has not been deleted from the database).

Even if the site's main or top level pages are static, dynamic data access web sites raise some additional issues for records management, including:

Not all users "see" the same web site. At designated levels, the pages displayed on users' browsers are based on what they ask for, therefore user queries are an integral part of generating the web site and may need to be captured.

Information contained in databases behind the site may be continually changing.

#### Scheduling

Scheduling these types of web sites can be complex. Agencies should review the section, 6.3 "Technical Guidelines for Scheduling and Managing Web Sites" and evaluate which method(s) are appropriate.

### 6.2.4 Dynamically Generated Web Sites

An increasing number of web sites are being built which generate all of the pages "on the fly." This means that the component parts of each individual page -- its content, structure, and presentation -- are generated dynamically using a combination of databases and stylesheets based on:

- A stored set of user preferences
- A stored set of access profiles
- A user query and/or
- The capabilities of the user's browser

In these situations, the web site does not exist in any single or easily capturable form. Each

user sees a different "site" based on their stored preferences and access rights, current needs, and the capabilities or limitations of the technology they are using.

Although the end result for the user might be a set of static pages, the processes that build the pages involve the use of a number of software tools. This is the point at which web sites become more like software applications than electronic publications. Agencies need to consider how to manage and archive dynamically generated web resources in a fully functional state.

### **Scheduling**

Scheduling these types of web sites can be complex. Agencies should review the section, "Technical Guidelines for Scheduling and Managing Web Sites" and evaluate which method(s) are appropriate.

[\*\*Section 6.3: Technical Guidelines for Scheduling and Managing Web Sites\*\*](#)

[Go to Table of Contents](#)