

## NARA Guidance on Managing Web Records

*January 2005*

Web site operations are an integral part of an agency's program. Managing web records properly is essential to effective web site operations, especially the mitigation of the risks an agency faces by using the web to carry out agency business. This guidance will assist agency officials in this regard, including agency program staff, webmasters, IT staff, and other agency officials who have a role in web site management and administration.

**This guidance is comprised of four sections:**

- **GENERAL BACKGROUND, RESPONSIBILITIES, AND REQUIREMENTS** – This section outlines the various ways in which agencies use web sites, the roles played by different agency staffs in web operations, and the basic statutory requirements that govern web sites, especially the Federal Records Act. This section also identifies the types of records agencies typically accumulate in connection with their web sites.
- **MANAGING WEB RECORDS** – This section is intended to assist agency staff in managing their web records. Aimed particularly at webmasters, IT staff, and those program officials responsible for web content, this section outlines the steps an agency must take to ensure trustworthy web records and mitigate the risks associated with web site operations.
- **SCHEDULING WEB RECORDS** – This section is intended to assist agency staff, especially records officers, in developing disposition schedules for web records. It addresses such matters as the types of records that should be covered in web schedules, how these schedules might be structured, and the factors an agency should consider in determining how long records should be retained.
- **APPENDIXES**

## PART 1 - GENERAL BACKGROUND, RESPONSIBILITIES, AND REQUIREMENTS

---

### *Table of Contents*

---

#### Introduction

1. How are agencies currently using the web?
  2. Who has records management responsibilities for agency web sites?
  3. What statutory and regulatory requirements apply to agency web operations?
  4. What are Federal web site-related records?
  5. What web site records must be managed?
  6. Does managing agency web sites as Federal records mean that I must keep all page changes for a long time?
- 

#### **Introduction**

Web site operations are an integral part of agency programs. This guide provides an initial, high-level framework that you can use to manage both the content records on an agency web site and the records documenting web site operations. This guidance, which is based on statutory requirements, also provides principles that form a sound basis for agency web site management operations. It will help you reach agency program objectives by relating good records management practices (e.g., maintaining trustworthy web sites) and existing tools for meeting agency objectives (e.g., risk assessments).

While the examples in this guidance primarily discuss public Internet web sites, the guidance is equally applicable to web sites that may be on agency intranets, virtual private networks, and security-classified web sites.

This guidance does not address unique issues relating to web portals that are managed by multiple agencies. For example, when agencies jointly sponsor web sites, one party must claim ownership of records management responsibilities. Those unique web issues will be dealt with in subsequent guidance.

---

#### **1. How are agencies currently using the web?**

Federal agencies are currently using the web in a variety of ways. Table 1 lists these various activities, in order of increasing complexity. This presentation is only an example and is not meant to be exhaustive or prescriptive about how agencies use the web.

**Table 1. Types of Federal Agency Web Site Operations**

<b>Description of Agency Web Site Activity</b>
<p><b>Web site maintained as a relatively stable repository for agency publications</b></p> <ul style="list-style-type: none"> <li>• <b>How used:</b> Web site serves information that changes infrequently, such as electronic versions of printed agency publications.</li> <li>• <b>Change frequency:</b> Updates add information that rarely is changed.</li> </ul>
<p><b>Web site maintained as a more fluid repository for agency publications and information</b></p> <ul style="list-style-type: none"> <li>• <b>How used:</b> Web site serves information that changes frequently.</li> <li>• <b>Change frequency:</b> Updates include revision, removal, and addition of information on an as needed basis.</li> </ul>
<p><b>Web site as a limited communications tool</b></p> <ul style="list-style-type: none"> <li>• <b>How used:</b> Web site allows minimal interaction with the agency, primarily by the public (e.g., end users can submit questions or comments via public mailboxes referenced by web pages).</li> <li>• <b>Change frequency:</b> Pages may be relatively stable or fluid; messages sent via web mailbox generally are captured through the agency e-mail system and not as part of the web site.</li> </ul>
<p><b>Web site as search or query-based access to agency information</b></p> <ul style="list-style-type: none"> <li>• <b>How used:</b> Web site is used as an application interface for agency services, including databases or forms used to submit requests for off-line services.</li> <li>• <b>Change frequency:</b> Forms for displaying information may change infrequently, but databases that provide information posted to the forms updated continuously.</li> </ul>
<p><b>Web site as interface to multiple applications and information services</b></p> <ul style="list-style-type: none"> <li>• <b>How used:</b> Web site is a significant tool delivering agency's major services online, on-demand, e.g., <ul style="list-style-type: none"> <li>○ Conducting public, electronic town meetings</li> <li>○ Using the web site to solicit comments on proposed regulations.</li> </ul> </li> <li>• <b>Change frequency:</b> Information both supplied by the end user and displayed on the web sites is varied in nature and changes frequently.</li> </ul>

In reality, web sites serve multiple purposes, and most agencies conduct more than one of these activities. Web sites can be seen as an automated "customer service window" through which customers interact with a business on many issues. They can ask questions, get directions or information, make complaints, seek refunds, apply for jobs, order merchandise, or make payments. Each of these activities generate different records and must be analyzed to determine the appropriate level of management.

Records created from conducting program activities over web sites have a variety of characteristics, depending on how the web site is used. These characteristics influence the

selection of an appropriate unit of analysis for a risk assessment (e.g., web site in toto, specific program portions of web site, etc.), as is discussed in [MANAGING WEB RECORDS, section 2](#). Knowing the purpose, work processes, and type of content is basic to understanding how to manage the web site.

---

## 2. Who has records management responsibilities for agency web sites?

Responsibilities for managing Federal agency web sites and related records are assigned to different offices and individuals within agencies, depending on how each agency is organized.

The **head of the agency**, who is responsible for protecting the integrity of agency programs and trustworthiness of agency information (see [44 U.S.C. Chapter 31](#)) has statutory responsibility for the agency's records management program. The **agency Records Officer** is responsible for ensuring adequate management and control over agency records, including agency web site-related records, and development of an overall records management program.

Other records management-related responsibilities for web records are diffused throughout the agency to the programs and functions that create web content and web site operations records. Just as an agency's web presence requires a collaborative effort, so does managing the web records that are created. NARA believes agencies must take a team approach to implement this guidance effectively. NARA recommends that the team should include the individuals who have the following responsibilities relating to web content and web site operations:

- Developing management procedures for agency web site operations and the records that document those operations
- Ensuring that required web management and operations records are created and maintained
- Coordinating the overall presentation of an agency web site
- Approving what program-related content material gets posted
- Contributing to the content pages that are posted on an agency's web site
- Contractor provision of additional content creation, graphics design, editorial, technical, or other services supporting web site development and operations
- Legal counsel advice on or review of potential legal issues (e.g., incorporation of copyrighted materials in web site design or e-FOIA)

Agencies are assisted in these responsibilities by:

**The National Archives and Records Administration (NARA)**, which is responsible for issuing policies and guidance to assist Federal agencies to comply with statutory and regulatory recordkeeping requirements related to web site operations; and

**The Office of Management and Budget (OMB)**, which has responsibilities under [44 U.S.C. Chapter 35](#) for coordinating Federal information policy.

---

## 3. What statutory and regulatory requirements apply to agency web operations?

The Federal Records Act applies to all agency records, including web records (see [section 5](#)). *The E-Government Act of 2002* (Public Law 107-347), places a number of requirements relating to web sites on OMB, NARA, and agencies. Agencies must comply with a variety of other statutes and regulations for their agency web site operations, most of which are described in the [Federal Web Content Manager's Toolkit](#). Also, NARA has issued [guidance](#) for transferring

permanent web content records to the National Archives.

---

#### 4. What are Federal web site-related records?

Most Federal agencies use web-based technologies to assist in carrying out their mission. They may simply disseminate information also available in other forms or conduct business (e.g., e-Government initiatives). Agencies must document all of their agency programs ([44 U.S.C. 3101](#)), including web sites, that are part of its overall public message.

Web site-related records include (1) web content records, which represent information presented on a web site, and (2) web site administrative records, which provide evidence of the management and operations of the web site.

Federal personnel, contractors, and partners supporting Federal web operations should understand that agency web content may meet the definition of a Federal record <sup>1</sup> [and therefore must be managed as such](#). Records relating to managing web sites, tracking use (metrics), file posting, and documenting decisions about web content and design are also Federal records.

Federal web sites are part of an agency's approach in serving the public. The agency determines, in accordance with [44 U.S.C. 3103](#), what records must be created and maintained to provide "adequate and proper documentation of the organization, functions, policies, decisions, procedures, and essential transactions of the agency and designed to furnish the information necessary to protect the legal and financial rights of the Government and of persons directly affected by the agency's activities." It is NARA's view that much, if not all, documentation related to agency web site operations should be managed as Federal records. As such, they must be scheduled and can only be deleted with a NARA-approved disposition authority.

---

#### 5. What web site records must be managed?

You must manage those web site related records that result from agency web operations, including those that are needed to ensure trustworthiness for each site and those needed to document agency programs. Use the concepts of reliability, authenticity, integrity, and usability, as discussed in [MANAGING WEB RECORDS, section 1.1](#), to establish criteria for identifying the web site records needed to document agency programs. To help you identify the types of web site related records your agency may have, the following list suggests *potential* series of records that might result from agency web site operations. These web site related records may exist in electronic or non-electronic format. NARA does not expect that your agency will produce all of these types of records, and you may have other types of records not on the list. This list also is not intended to instruct agencies about composition or management of their web sites.

- *HTML-encoded pages*: The content pages composing an agency web site, inclusive of the HTML markup.
- *Records generated interactively on the web site*: Records that are created dynamically in real time when a user interacts with an agency web site (e.g., on-the-fly, text-based page creation, forms filled out online, etc.).
- *Additional content files referenced within HTML-encoded pages*: Files having the ability to "self-execute" (e.g., CGI scripts, Java/ActiveX applets, customized programs that generate online sound or moving images) as well as files that are static (e.g., these include graphic files, multi-national character sets, etc.). Both self-executing and static pages are external to the HTML-encoded content pages but referenced in the HTML syntax.
- *Comprehensive list of URLs referenced by a web site's hyperlinks*: In effect, a

bibliography of all uniform resource locators referenced via hyperlinks embedded within a web site's content pages, along with a textual reference describing the destination of the hyperlink.

- *Web site design records*: Records such as graphic design layouts for pages or embedded image maps, and/or minutes of meetings documenting the production of such.
- *Copyrighted web content and records documenting the use of such content materials*: Many Federal web sites incorporate copyrighted content (e.g., works for hire such as custom produced graphics files) in individual web pages. Additional records may be deemed necessary for documenting the appropriate use.
- *Web site program operations records*: Program management files that document policies and procedures for agency web site operations, including those that
  - Provide detailed procedures for documenting how records are selected, created and approved for web posting, and how they will be revised or removed
  - Specify what records will be created and how they will be created for interactive sections of web sites
  - Document procedures used in the operation of the site
  - Specify the relationship of the webmaster and other staff involved in preparing and posting web documents to program officials and the agency records officer
  - Demonstrate the development of policies and procedures to ensure Section 508 compliance
  - Otherwise explain or illustrate site development and management procedures
- *Web site system software-related records*: Records related to the application software used in conjunction with operation of the web site. These include
  - Records produced in the analysis and selection of any commercial off-the-shelf (COTS) software
  - Records describing customization of COTS web-related software
  - Documentation relating to COTS web-related software
- *Web site logs and statistical compilations*: Records including raw data or summary logs of user access (frequency of "hits"), site posting logs (indicating when pages were posted, updated, or removed from the site), system load and search result statistics, and ad hoc reports containing such.
- *Web site map*: A linked, graphic or text-based display of a web site's hierarchy, similar to an organization chart.
- *COTS software configuration files*: Files used to operate the web site and establish its look and feel including server environment configuration specifications.

---

**6. Does managing agency web sites as Federal records mean that I must keep all pages and their changes for a long time?**

No. As [MANAGING WEB RECORDS](#) and [SCHEDULING WEB RECORDS](#) discuss in greater

detail, your agency business needs, including the risks to the agency programs and mission should the information not be available, are the major factors in determining how long you need to keep those pages. Your web site schedule specifies the length of time you need to keep pages.

---

Endnote for PART 1 - GENERAL BACKGROUND, RESPONSIBILITIES, AND REQUIREMENTS

<sup>1</sup> I.e., they constitute "machine readable materials, . . . made . . . by an agency of the United States Government . . . in connection with the transaction of public business . . . ([44 U.S.C. 3301](#))."

## PART 2 - MANAGING WEB RECORDS

---

### *Table of Contents*

---

#### Introduction

1. **What are trustworthy records?**
    - 1.1 What are the characteristics of trustworthy records?
    - 1.2 How can I maintain a trustworthy web site?
  2. **Risk and risk assessment**
    - 2.1 What are the records management risks associated with web sites?
    - 2.2 How can I conduct a records management risk assessment?
    - 2.3 How do I determine the unit of analysis for doing a risk assessment?
    - 2.4 Who is involved in assessing risk?
    - 2.5 What do I do with the results of my risk assessment?
  3. **Mitigating risk**
    - 3.1 What steps should I follow to help mitigate risk in managing web records?
    - 3.2 How should an agency manage web site content in order to mitigate risk?
    - 3.3 How frequently should I capture a snapshot of my site's web content records?
    - 3.4 How do I track changes to web site content pages between snapshots?
    - 3.5 When preserving long-term web content records, how can I treat hyperlinks?
  4. **Roles and responsibilities**
    - 4.1 Who is responsible for managing web content records?
    - 4.2 Who is responsible for managing web management and operations records?
- 

#### **Introduction**

This guide is intended to assist agency staff in managing their web records. It is particularly geared to the needs of program officials, who provide the information posted on web sites, and those staff who manage agency web sites, including webmasters and IT staff.

---

### **1. What are trustworthy records?**

Trustworthy records are essential for an agency to meet its legal and internal business needs. Reliability, authenticity, integrity, and usability are the characteristics used to describe trustworthy records from a records management perspective. How these terms apply to web sites and web records is discussed more fully in [section 1.1](#).

Creating and maintaining trustworthy records requires resources. Agencies need to conduct a risk analysis to balance the level of trustworthiness of records against costs and risks. The level of resources used to ensure these characteristics depends on the agency's business needs and perception of risk. (See [section 2](#) for a discussion of risk assessment.) Web site operations that are critical to agency business need a greater assurance level that they are reliable and authentic, maintain integrity, and are usable over a longer period of time than less critical operations.<sup>1</sup>

---

#### **1.1 What are the characteristics of trustworthy records?**

**Reliability.** A reliable web site is one whose content can be trusted as a full and accurate representation of the transactions, activities, or facts to which it attests and therefore can be



depended upon in the course of subsequent transactions or activities.

**Authenticity.** An authentic web site is one that is proven to be what it purports to be and to have been created by the agency with which it is identified.

Web site-related records should be created by individuals who have direct knowledge of the facts or by instruments routinely used within the business to conduct the transaction.

To demonstrate the authenticity of a web site, agencies should implement and document policies and procedures that control the creation, transmission, receipt, and maintenance of web site records to ensure that records creators are authorized and identified and that records are protected against unauthorized addition, deletion, and alteration (e.g., via hacking).

**Integrity.** The integrity of a web content record refers to it being complete and unaltered.

The agency's web management policies and procedures for routinely updating and modifying their web sites help ensure integrity. As stated in the ISO Technical Report 15489-2<sup>2</sup>, sec. 7.2.4, "records systems should maintain audit trails or other elements sufficient to demonstrate that records were effectively protected from unauthorized alteration or destruction." The web management policies should prescribe how changes to the web site are to be documented.

Another aspect of integrity is the structural integrity of a web site's content-related records. The structure of a web site, that is, its physical and logical format and the relationships between the pages and content elements composing the site, should remain physically or logically intact. Failure to maintain the web site's structural integrity may impair its reliability and authenticity.

**Usability.** A usable web site is one that can be located, retrieved, presented, and interpreted. In retrieval and use, you should be able to directly connect the web site to the business activity or transaction that produced it. You should be able to identify both the site and its content within the context of broader business activities and functions. The links between content, contextual, and structural web site-related records that document agency web site activities should be maintained. These contextual linkages should provide an understanding of the transactions that created and used them.

---

## 1.2 How can I maintain a trustworthy web site?

For web site records to have integrity and remain reliable, authentic, and useable for as long as they are needed, you must maintain the content, context, and sometimes structure of the site. A trustworthy web site includes not only the content pages but also information about the web site that relates to the context in which it was created and used. Specific contextual information varies depending upon the business, legal, and regulatory requirements of the business activity. Structural information on the organization of the web site supports its long-term integrity.

Federal web site-related records that support content, context, and structure are:

**Content:** The actual HTML-encoded pages themselves and additional content files referenced therein or content created by end users interacting with the web site. Maintenance of these web content records is necessary to support all of the characteristics of trustworthiness: reliability, authenticity, integrity, and usability.

**Context:** Administrative and technical records necessary for or produced during the management of an agency web site. Maintenance of these records provides a context for web operations, which attests to the reliability, authenticity, and integrity of an agency's web site.

**Structure:** For those web sites (or portions) that have been appraised as permanent and for high-risk temporary sites, a site map indicating the arrangement of a web site's

content pages and software configuration files of content management systems. Maintenance of this record provides a structure for content records and thereby enables the integrity and usability of both current and preserved versions of an agency web site.

Records in all of these categories contribute to the adequate documentation of agency web site operations. A risk assessment of web site operations advises which records are necessary to ensure the operation's trustworthiness, how the records should be maintained appropriately, and how long those records are to be retained (see [section 3](#)).

---

## 2. Risk and risk assessment

Typically, agencies conduct risk assessments in order to establish appropriate levels of management controls prior to undertaking new program initiatives. NARA assumes that such risk assessments have been conducted for development of agency web site operations. These risk assessments can also be used to establish records management controls.

Agency records management practices are based on operational needs and perceptions of risks. Operational needs (e.g., providing public information, documenting transactions with the public) determine the way agencies address the trustworthiness of web site operations (see [section 1.1](#)). Risk assessment and risk mitigation, along with other techniques, are used to establish both management controls for and documentation requirements of agency activities. The emphasis in this guidance on risk assessment relates to Clinger-Cohen requirements for incorporation of risk management into program activities, particularly for those that are dependent upon information technology (e.g., web site operations).

---

### 2.1 What are the records management risks associated with web sites?

From a records management perspective, risk relates to (1) challenge to the trustworthiness of the records (e.g., legal challenge) that can be expected over the life of the record; and, (2) unauthorized loss or destruction of records. Consequences are measured by the degree of loss that the agency or citizens would suffer if the trustworthiness of the web site-related records could not be verified or if there were unauthorized loss or destruction.

Examples of records management-related risks associated with agency web sites are mainly technical risks. Loss of information could result from:

- an inability to document or validate transactions that occur via an agency web site front end;
- an inability to reconstruct views of web content that was created dynamically and existed only virtually for the time that they were viewed;
- compromise of e-Government transactions; and
- an inability to track web-assisted policy development or document agency decisions relating to agency web operations.

A variety of negative programmatic consequences can result from any of these technical risks:

- litigation or liability if an agency is unable to verify what was on its site at a given point in time;
- impairment of program operations or an inability to detect or punish fraud, false statements, or other illegal behavior because of a lack of valid or probative records;

- an inability to produce records that document accountability and stewardship of materials posted to the agency web site; dissemination of misinformation;
  - financial losses due to compromising the citizens' or government's rights;
  - compromise of the agency's mission;
  - negative reactions of agency stakeholders (e.g., the Executive or Legislative branch); and
  - unfavorable media attention.
- 

## 2.2 How can I conduct a records management risk assessment?

A risk assessment should address the possible consequences of untrustworthy, lost, or unrecoverable records, including the legal risk and financial costs of losses, the likelihood that a damaging event will occur, and the costs of taking corrective actions. Agencies may have formal risk assessment procedures that may be applied to agency web site operations.

The assessment factors may include *records management threats*, *visibility*, *consequences*, and *sensitivity*.

*Records management threats* relate to the likelihood of experiencing technical risks discussed in section 2.1 (e.g., risks of unauthorized destruction of web site-related records, litigation risks associated with inability to reconstruct views of web sites at specific points in time, risks associated with inability to document web site policy decisions, etc.).

*Visibility* is the level of active public awareness of an agency's web site operations.

*Consequences* describes the level of negative organizational, economic, or programmatic impact if web records are untrustworthy, lost, or unrecoverable.

*Sensitivity* characterizes the agency's assessment of the importance of web site operations.

The results of an assessment will support agency programs by providing a basis for determining what types of web site records should be created, how they should be maintained, and how long they should be maintained. The assessment will help agencies ensure that the level of risk is tolerable and that resources are properly allocated. Assessment results can also aid in the development of web site records schedules.

---

## 2.3 How do I determine the unit of analysis for doing a risk assessment?

One key aspect of conducting a risk assessment is determining the appropriate unit of analysis; i.e., whether the web site will be assessed as a single entity<sup>3</sup> or whether assessments will be conducted for different portions of the site. This is important because it affects the choice of management controls (see [section 3.1](#)) and scheduling (see [SCHEDULING WEB RECORDS, section 5](#)). The concept of the appropriate unit is flexible to allow you to adapt it for your particular site and management needs. Possible units of analysis include the entire site, portions of the site related to specific functions or organizations, clusters of pages on a specific subject, etc.

### Basic options for analysis

- Evaluate the web site *in toto*. Note that this option is not advisable if the web site has multiple types of content (e.g., e-commerce transactions and static publications) or functions served. Records management risk and required management controls vary for those different portions of the web site
- Evaluate groupings of web sites referenced by an agency's main portal entry page
- Evaluate the web site basically as a whole, minus one or two portions that exhibit substantially different characteristics
- Substantially break out clusters or groups of web site pages based on function or other characteristics. Note that this option does not anticipate a page-by-page risk analysis of your web site.

First consider whether the site has a single level of risk or varying levels of risk. Use the risk assessment factors. If the level you have chosen for analysis has more than one answer to any of the factors, you may need to consider breaking out those portions. Note that changes in any of the four factors could affect the risk level.

#### Example:

One means of portioning the NARA web site for risk assessment is dividing it by program areas; for example, the Records Management portion of the NARA web site (see [http://www.archives.gov/records\\_management/index.html](http://www.archives.gov/records_management/index.html)).

Another example of portioning is to separate out specific content pages; for example, the FOIA portion of the NARA web site (see [http://www.archives.gov/research\\_room/foia\\_reading\\_room/foia\\_reading\\_room.html](http://www.archives.gov/research_room/foia_reading_room/foia_reading_room.html)).

A final example of portioning the NARA web site would be on the basis of the nature of the content pages; for example, those pages composing the Archival Research Catalog database (see [http://www.archives.gov/research\\_room/arc/](http://www.archives.gov/research_room/arc/)) or unique, one-time exhibits such as those on the Charters of Freedom (see [http://www.archives.gov/exhibit\\_hall/charters\\_of\\_freedom/charters\\_of\\_freedom.html](http://www.archives.gov/exhibit_hall/charters_of_freedom/charters_of_freedom.html)).

Determine the unit of risk assessment in consultation with other agency staff associated with the web site (see [GENERAL BACKGROUND, RESPONSIBILITIES, AND REQUIREMENTS, Section 2](#)).

If you decide, for operational reasons, to evaluate the web site as a single unit, all components will be treated the same in terms of risk. You will need to manage all parts of the site in accordance with the highest level of risk determined for any portion of the site. When applying this guidance to portals that are primarily federations of an agency's web sites, you must manage all of the agency web sites at the highest level of risk encountered in the aggregation.

---

## 2.4 What types of duties/functions should be involved in assessing risk?

Staff in various roles throughout your agency should contribute to your risk assessment. They will

bring knowledge and experience about these aspects of a web site:

- The nature of the information on the site, who uses it, and what problems might arise if information on a site is incorrect, out-of-date, or lost. An example is the program staff that produce content to be posted to the web site.
- How information is placed on the site, revised, and removed and, in addition, know what records are created or should be created when these actions take place. An example is the webmaster.
- Relating business processes to the records that result from those processes, developing procedures for ensuring the trustworthiness of those resulting records, determining appropriate retention periods for those records, and obtaining approved retention schedules reflecting such. An example is a records manager.
- Expertise in computer technology and the risks that its use can cause or mitigate. An example is the IT staff associated with web site operations.
- The legal requirements that the agency must follow, unique legal risk that might arise from web site operations, and an understanding of those types of records that may be required in legal proceedings. An example is the legal staff.

---

## **2.5 What do I do with the results of my risk assessment?**

After you have determined the level of records management risks for the site or portions of the site, you will need to protect the records appropriately. Review any web management policies and procedures that you already have in place to determine whether additional steps are needed. Develop a plan to address records issues (e.g., types of records needed to document the web-based activity, length of time they are needed to support the business purposes), as well as IT issues (e.g., security of the site and information exchanged over the site) and management/internal controls on the processes. The agency's program, web, and IT staff, the agency records officer, and the General Counsel should contribute to developing the plan.

---

## **3. Mitigating risk**

Risk mitigation issues are of particular relevance to program staff responsible for web content and to webmasters. These issues include how to mitigate risk by producing a web snapshot and other means of documenting web site content, how changes to sites between snapshots can be tracked, and how hyperlinks may be treated when preserving long-term web content pages. This section also addresses the roles of the web management and agency program staff in schedule implementation.

---

### **3.1 What steps should I follow to help mitigate risk in managing web records?**

Some of the steps outlined here are the same as for other kinds of records. You should address each of these steps:

- Document the systems used to create and maintain your web records.
- Ensure that your web records are created and maintained in a secure environment that protects the records from unauthorized alteration or destruction.

- Implement standard operating procedures for the creation, use, and management of your web records and maintain adequate written documentation of those procedures.
- Create and maintain your web records according to these documented standard operating procedures.
- Train agency staff in the standard operating procedures.
- Develop a retention schedule for your web records and obtain official NARA approval of that retention schedule. (See [SCHEDULING WEB RECORDS](#).) You will need to cite the official disposition authorities found in your schedule if your agency is faced with legal challenges to produce records that have been destroyed.

The results of your risk assessment will indicate the level of effort necessary to mitigate your risks.

---

### 3.2 How should an agency manage web site content in order to mitigate risk?

You must preserve the records as long as they are needed for business operations. Traditional records management techniques apply fairly easily to relatively stable contextual and structural web site records. Managing web content pages is more complex. Web content pages may be frequently changed or updated, and when updates or redesign of web site maps change the relation/organization of web content, it may be deemed necessary to set aside a new recordkeeping copy of web site content.

Agencies may preserve web content records by (1) producing a stand-alone copy or snapshot<sup>4</sup> of all content pages on the site at a particular time and (2) accompanying this snapshot with a [site map](#) that shows the relationship (i.e., directory structure) of those pages to each other. If your agency decides to take snapshots, you must decide:

- how frequently a new snapshot should be captured;
- if it is necessary to track changes in both the content pages and the site map that occur between snapshots; and
- if it is, how to track these changes (see [section 3.4](#) ). The answers to these questions depend on your risk assessment of web site operations.

Content management systems (CMS) can be used to manage the content of a web site. The system consists of a content management application (CMA) and a content delivery application (CDA). The CMA can relieve the webmaster of many of the decisions and actions required to manage the creation, modification, and removal of content from a web site. A CDA uses and compiles the content management information to update the web site. CMSs can be used to create audit trails associated with content that is created on-the-fly.

To ensure availability of current web content, you may use web server back-up software or an Internet-based service to preserve copies of files or databases to restore the content in case of equipment failure or other catastrophe.

Instead of snapshots or preservation of web content in a records management application (RMA), agencies may decide to manage the **live** versions of web site content pages **while the pages are up on the web site**. For low risk web sites, the current posted version of a site plus the standard operating procedures in place used to manage the site and a log of changes may be sufficient for business purposes. Please note that because this option does not set aside recordkeeping copies, it may not be appropriate for medium- and high-risk sites .

---

### 3.3 How frequently should I capture a snapshot of my site's web content records?

Determine the frequency of snapshots of a site's web content records and site map by using the risk-profiling factors described in [section 2.2](#) . The unit(s) of analysis for the risk assessment would correspond to the unit considered for the snapshot. Portions of a web site considered of higher records management risk are likely to require more frequent snapshots. The stakeholders discussed in [GENERAL BACKGROUND, RESPONSIBILITIES, AND REQUIREMENTS, section 2](#) should cooperate in deciding how frequently snapshots should be taken.

---

### 3.4 How do I track changes to web site content pages between snapshots?

Four types of changes can occur to a web site's content between snapshots:

1. Changes to the content of an individual page without changing its placement in the overall organization of the web site
2. Wholesale replacement of an individual page (or sections of pages) without changing its placement in the overall organization of the web site
3. Changes in location of a page (or groups of pages)
4. Combinations of changes of these first three types.

Changes of the first two types (i.e., changes to content without changing the page's placement in the overall organization of the web site) can be treated as a version-control issue. You must decide how to best keep track of the versions of content pages.

The most fundamental, non automated approach to tracking web site content, particularly for relatively stable sites, is to "print and file" a recordkeeping copy in the manual recordkeeping system. Another non automated approach to version control is to annotate changes of content pages as a comment in the HTML coding. The comment, which will not appear when the page is displayed in a browser, could indicate when the page was changed (e.g., **<!--Updated by MDG on 03/02/03-->** ) or could reference the page which it wholesale replaced (e.g., **<!--This page replaced content page Introduction\_1.html on 09/10/02-->** ). Another manual approach would be to maintain a log file of content changes of the first two types of changes. (Keep in mind that neither of these approaches would allow you to actually reconstruct views presented at a particular time. This may be found acceptable per your risk assessment).

Alternatively, you may use content management software (CMS) to track versions of web content in the first two cases. CMS would also offer limited page view reconstruction capabilities-default settings for the databases that support most CMS software would retain only recent changes.

You can handle major changes to the site's directory structure by producing a new site map at the time of major revision. This could be accomplished in a manual or automated manner.

One automated way to track changes is to manage the web content records with a DoD 5015.2-certified records management application (RMA). A DoD 5015.2-certified RMA allows you to impose version control over changed copies of documents. If you use the RMA to store iterative copies of individual web pages as they are changed, you will be able to see how many times and when each page was changed. Web content may be added to an RMA's repository manually or via any of the automated tools discussed below. Please note that DoD-certified RMAs have been endorsed by NARA for civilian agency use because they comply with records management regulations. However, none of the other tools described below were designed for records management.

Another tool is a type of search engine called "web harvester." Also called a "spider" or "crawler,"

to create entries for a search engine index. You can use harvester software to identify changes to web site content and to gather content related to specific site (sub)units.

When justified by risk assessment, you may want to be able to closely reconstruct the content and structure of a site by combining records of updates to web content pages with snapshots of web sites. The degree of exactness to which a web site may be reproduced depends on whether changes to all static and dynamic files referenced within HTML-encoded content pages were also tracked between snapshots.

### 3.5 When preserving long-term web content records, how can I treat hyperlinks?

Web content pages use hyperlinks to: (1) jump to another location within the page, (2) jump to a location on other pages within the web site, or (3) jump to a page on another web site. Depending on the preservation strategy chosen, it is possible, and in many cases likely, that these hyperlinks will not continue to function in the preservation copy of the web content records. If the site does not follow external-link-liability-transference policies such as those employing pop-up window notifications, agencies might want to use the following suggestions, to enhance the usability of preservation copies of long-term web content records. For hyperlinks within web content records appraised as permanent, agencies must adhere to [NARA's Transfer Instructions for Permanent Web Content Records](#) when transferring the records to NARA.

#### Suggestions for Managing Hyperlinks in Web Content for Long-Term Preservation

<p><b>Internal target hyperlinks</b></p>	<p>For hyperlinks that simply send the user to a different location within the same page (aka internal target), no additional work is required, as the link will continue to function when the content page is interpreted by a browser application.</p>
<p><b>Hyperlinks not under local records management control</b></p>	<p>For hyperlinks that send the user to either a different page or another web site that is not under the agency's records management control, NARA suggests that agencies consider requiring web site content developers to modify the HTML syntax of web content pages containing such hyperlinks on a day-forward basis. This modification would include the insertion of an HTML comment after the hyperlink that described, in plain English, name of the site (and perhaps portion of site) or page to which the hyperlink transfers. For example, the hyperlink in the records management portion of the NARA web site discussing DoD 5015.2-STD that links to the Joint Interoperability Test Command's web site, expressed in HTML (emphasis added) as <code>&lt;a href="http://jitic.fhu.disa.mil/recmgt/#standard"&gt;DoD Standard 5015.2&lt;/a&gt;</code> would be modified, inserting an appropriate title attribute, per accessibility requirements, that describes it, as follows: <code>&lt;a href="/global_pages/exit.html?link=http://jitic.fhu.disa.mil/recmgt/#standard" title="Joint Interoperability Test Command's 5015.2 - STD Records Management Application Design Criteria Standard"&gt;DoD Standard 5015.2&lt;/a&gt;</code></p>
<p><b>Hyperlink to new page within same web site</b></p>	<p>When a page includes a hyperlink that sends the user to another page <i>in the same web site</i>, it would be necessary to insert comments describing the hyperlink only when the site was not being scheduled <i>in toto</i> for the same retention (and those comments could reference the series containing the destination of the hyperlink).</p>



Another alternative would be to produce what is in effect a bibliography for all of the hyperlinks referenced within the content pages composing a site. List all of the URLs referenced by hyperlinks, along with a description of the hyperlinked page (much as in the comment used in the previously suggested method).

---

## 4. Roles and responsibilities

### 4.1 Who is responsible for managing web content records?

Content pages on the web site may originate in many program areas within the agency and may be created by agency staff or contractors. The agency should establish clear guidelines for managing records on web sites. The guidelines should specify whether the program office or the webmaster's office (or other office responsible for the web site) is responsible for implementing the agency's records management policies for these records. The program office and the personnel responsible for agency web operations may each have specific responsibilities in this area. Ideally, agencies should have a team of individuals, including program staff, web management staff, and records management staff, who develop records management plans for the web site. Among their responsibilities are both the development of records schedules and managing the retention and disposition of the web records.

Web content pages may be scheduled as records of the (program office) content owner or of the web services providers. The schedule decision should be based on which office is assigned the responsibility for keeping content current, setting security levels, and identifying access requirements (See [SCHEDULING WEB RECORDS](#) for additional information on scheduling web records.)

Each agency decides which office will be responsible for implementing records schedules. If the web content pages are scheduled as records of the relevant program office, that office also implements the schedule. The program office will need to establish procedures to ensure that the schedule is properly implemented, including notification to the web operations staff when web content records need to be destroyed.

---

### 4.2 Who is responsible for managing web management and operations records?

Agency personnel who manage the web site are responsible for managing the contextual and structural records necessary to adequately document agency web site operations. Web management records are managed the same as other program records in the agency. Web management records provide context and structure for web sites and do not present the same complexities as content records, which are frequently revised or replaced. Hence, standard records management techniques should be sufficient.

---

Endnotes for PART 2 - MANAGING WEB RECORDS

<sup>1</sup> For guidance on whether records are trustworthy for legal purposes, consult your Office of General Counsel.

<sup>2</sup> ISO/TR 15489-2:2001, Information and documentation - Records management - Part 2: Guidelines. See <http://webstore.ansi.org/ansidocstore/find.asp>.

<sup>3</sup> In cases when assessing a portal web site as a single entity, it is necessary to manage all sites in the portal to the highest level of risk encountered by any individual site in that portal.

<sup>4</sup> NOTE: A snapshot captures a web site as it existed at a particular point in time (e.g., by harvesting, exporting to an image format, simple device backup). For web content records

appraised as permanent, agencies must use capture method(s) that retain hypertext functionality (e.g., harvesting) as described in [NARA's Transfer Instructions for Permanent Web Content Records](#).

## PART 3 - SCHEDULING WEB RECORDS

---

### *Table of Contents*

---

#### Introduction

1. [What records should be covered in a web schedule?](#)
  2. [What purposes does a web schedule serve?](#)
  3. [How is a web schedule developed?](#)
  4. [How does risk assessment help with developing a records schedule?](#)
  5. [What is the structure of a web schedule?](#)
    - 5.1 Single Item Schedule for Web Content and Site Management and Operations Records
    - 5.2 Multiple Item Schedule for Web Content and Site Management and Operations Records
    - 5.3 Web Snapshots
    - 5.4 Examples of Scheduling Options
  6. [How are retention periods for web site-related records determined?](#)
  7. [Can the GRS or existing agency schedules be used for web records?](#)
- 

#### **Introduction**

This guide is intended to assist agency staff, especially records officers and webmasters, in developing disposition schedules for the records relating to agency Internet and intranet web sites. This guidance will discuss such matters as

- the types of records that must be scheduled,
  - how risk assessments may be used in making scheduling decisions,
  - how web schedules should be structured, and
  - the factors records officers should consider in determining retention periods for web records.
- 

#### **1. What records should be covered in a web site schedule?**

A web schedule should cover web content records that document the information on the site itself. A web schedule should also include web site management and operations records, which provide the site's context and structure.

More detailed information concerning the types of records agencies accumulate in connection with their web sites is available in [GENERAL BACKGROUND, RESPONSIBILITIES, AND REQUIREMENTS, section 5.](#)

Web content records include:

- the content pages that compose the site, inclusive of the HTML markup;
- records generated when a user interacts with a site; and
- if the agency chooses to document its site this way, lists of the URLs referenced by site's hyperlinks.

Web management and operations records that provide context to the site include:

- web site design records,
- records that specify an agency's web policies and procedures by addressing such matters as how records are selected for the site and when and how they may be removed,
- records documenting the use of copyrighted material on a site,
- records relating to the software applications used to operate the site, and
- records that document user access and when pages are placed on the site, updated, and/or removed.

Web management and operations records that provide structure related to the site include:

- site maps that show the directory structure into which content pages are organized and
- COTS software configuration files used to operate the site and establish its look and feel, including server environment configuration specifications.

---

## 2. What purposes does a web schedule serve?

A web schedule fulfills an agency's statutory responsibilities as spelled out in the Federal Records Act. In addition, a web schedule mitigates the [risks](#) associated with the agency web site by ensuring that records needed to prove its trustworthiness are maintained for an appropriate period of time. A web schedule also provides you the legal authority to destroy web records at the end of their NARA-approved retention period. Finally, the scheduling process will identify any web-related records that warrant permanent retention and eventual transfer to the National Archives.

---

## 3. How is a web schedule developed?

Developing web schedules involve several distinct steps. The most important of these are:

- determining the structure of the web schedule,
- describing the specific series to be included, and
- specifying retention periods for each series.

You will use the risk assessment as a key tool in performing these steps.

Developing a web schedule is the responsibility of the agency records officer (or his or her designee), who takes the lead in carrying out these steps. At all stages of schedule development, the records officer works closely with program staff (who are responsible for the site's content and

who are most familiar with the business needs and risks associated with the site) and with webmasters and IT staff (who are responsible for web operations). In the first stages, they provide the records officer with key data concerning the site and how it is used. They continue to work with the records officer in carrying out risk assessments, and at the end of the process, they must review the final schedule to ensure that it meets business needs and mitigates risks.

Scheduling of databases supporting content management systems is a sufficient means of addressing back-end, dynamically created content. Back-end programmatic databases for which a web page serves as the interface are normally scheduled as program records, separate from the web schedule(s).

---

#### 4. How does risk assessment help with developing a records schedule?

The analysis that is performed in a risk assessment helps you gather information about the web site, the agency programs that the web site supports, and the records that must be scheduled. Some of the information you will gather includes:

- how your agency uses its web site (see [GENERAL BACKGROUND, RESPONSIBILITIES, AND REQUIREMENTS, section 1](#), for examples),
- how often the site and specific portions of the site are changed or updated,
- the degree to which the information on a web site is unique or is readily available in other agency records, and
- whether the web site or portions of the site are considered high risk. More detailed information concerning risk and risk assessment is included in [MANAGING WEB RECORDS, section 2](#). Trustworthy records and their characteristics (i.e., reliability, authenticity, integrity, and usability) are discussed in [MANAGING WEB RECORDS, section 1](#).

---

#### 5. What is the structure of a web schedule?

Deciding on the level of analysis for risk assessment will resolve one aspect of schedule structure: whether the schedule describes records at the level of the entire web site or whether individual portions of the site are scheduled separately. If an agency chooses the latter approach, the schedule items should describe pages or groupings of pages broadly in terms of their content or function. Describing portions of the site too narrowly increases the likelihood that the schedule will become out of date as the site changes over time. Drastic changes in site content and/or function would likely require revisions to the schedule, just as significant changes in the content or function of a traditional record series typically warrant changes to a previously approved schedule.

There is no hard and fast rule as to what number of items is appropriate in a web schedule. Web management and operations records should be grouped, based on business needs and level of risk, into no more than three or four series for the entire site or applicable unit of analysis (i.e., portion of the site).

The following are two approaches to structuring a web schedule.

---

### 5.1 Single Item Schedule for Web Content and Site Management and Operations Records

You can use a single schedule item to describe the web content records (either the entire site or portions of it) along with the related records that pertain to site management and operations. This option would be appropriate if all of the records related to the site warrant the same retention period in order to meet business needs and mitigate risks. For the sake of simplicity and ease of management of the web site, an agency may also choose to use a single item and retention period for web records even if there are variations in business needs and risk. In this case, the retention period chosen must ensure that records are retained and remain usable for the appropriate period of time. This approach would require that some records with shorter business needs and lower levels of risk be maintained for a period of time longer than is necessary. See [APPENDIX C](#), option A, for an example.

---

### 5.2 Multiple Item Schedule for Web Content and Site Management and Operations Records

If business needs and the mitigation of risk mandate different retention periods for the site content records and the management and operations records, you can schedule them separately. Follow these guidelines:

- If all management and operations records associated with a site (or individual portions) are needed for a uniform period of time in order to mitigate risk, then a single item for all such records might be appropriate. Web content records would be covered by one or more separate items. [APPENDIX C](#), option B provides an example of multiple items for web content records and one item for all web management and operations records.
  - If the assessment of risk dictates that records are needed for different periods of time, then records that need to be kept for the same amount of time in order to mitigate risk should be grouped together and each grouping assigned an appropriate retention period. Such variation is likely in higher risk situations. Multiple items should then be developed for the site management and operations records, regardless of whether web content records are included in single item or in multiple items. See [APPENDIX C](#), option C and [APPENDIX C](#), option D for examples.
- 

### 5.3 Web Snapshots

When setting up the structure of a web schedule, decide if it should include an item for web snapshots that capture the content pages and related site map as they existed at particular points in time. Business needs and the need to lessen risk determine whether or not such snapshots are warranted and their frequency. However, in determining when snapshots should be taken, an agency should also consider how frequently the information on a site changes. A snapshot should be taken each time the site changes significantly.

---

### 5.4 Examples of Scheduling Options

[APPENDIX C](#) contains examples of the different options that agencies may employ in scheduling their web site records.

## 6. How are retention periods for web site-related records determined?

When determining retention periods for web site-related records, as with other records, the agency needs to assess how long the information will be needed to satisfy business needs and mitigate risk, taking into account Government accountability and the protection of legal rights. If specific web content is available in places other than the web, consider whether the existence of

the information in other records affects the retention needs for the web records. In the case of information unique to the web site, the web version is the only recordkeeping copy.

In many cases, particularly where the risk is low, the web content and the related site management and operation records should be assigned a retention period that allows disposal as soon as records are no longer needed in the conduct of agency business.

In instances where risk levels are higher, web content and the related web management and operations records would probably warrant retention for a period of time that exceeds the time needed to satisfy all business requirements. The extra time needed in order to mitigate risk would usually not be more than 3-5 years beyond the retention period mandated by business needs alone. However, the mitigation of risk may require an even longer retention period in selected instances.

As with other agency records, most web records do not warrant permanent retention and should be scheduled for disposal in accordance with the guidance provided above. In instances where NARA determines that a site or portions of a site has long-term historical value, NARA will work with the creating agency to develop procedures to preserve the records and provide for their transfer to the National Archives.

---

## **7. Can the GRS or existing agency schedules be used for web records?**

There are currently no items in the General Records Schedules that were developed to specifically cover web records. However, some items in the GRS may be applied to web site management and operations records. For example, GRS 14, items 1 and 2 (Information Service Records - Information Request Files and Acknowledgment Files, respectively) can be used to cover transactions of this sort that are generated via an agency web site.

Another GRS item that may be used for web records is GRS 21, item 6 (Audiovisual Records - Graphic Arts, Routine Artwork for Handbills, Flyers, Posters, Letterhead, and Other Graphics).

Likewise, records relating to training staff regarding the agency web site might be covered by subitems in GRS 1, item 29 (Civilian Personnel Records - Training Records).

A variety of items included in GRS 24 (Information Technology Operations and Management Records) may be relevant to web management and operations records. Examples include items 1, oversight and compliance records; item 3, IT asset and configuration files; item 5, files related to maintaining the security of systems and data; and item 6, user ID, profiles, authorizations, and password files.

As these examples demonstrate, a GRS item may be used for web management and operation records if the basic content and function of the records is consistent with the GRS and the retention period is appropriate to meet business needs and mitigate risk. This principle also governs whether or not an existing agency-specific schedule item can be used for web management and operation records. For example, a schedule item covering graphics design can be used for records generated in connection with this function as it relates to the agency web site. Likewise, an agency schedule item for internal committee records could be applied to records accumulated by a committee established to advise the agency webmaster.

## PART 4 - APPENDICES

---

### *Table of Contents*

---

- [APPENDIX A](#) - Definitions
  - [APPENDIX B](#) - Web Sites of Possible Interest
  - [APPENDIX C](#) - Web Schedule Options
- 

### APPENDIX A - Definitions

**ActiveX**—a set of "strategic" object-oriented programming technologies and tools developed by Microsoft allowing the creation of self-sufficient programs that can operate anywhere in an ActiveX network. A rival of the Sun's Java programming language. [from <http://whatis.techtarget.com/>]

**ActiveX control**—the rough equivalent of a Java applet. [from <http://whatis.techtarget.com/>]

**Applet**—an applet is a small program, or application module, that can be sent from a web site, along with browser-ready web content, to a user's workstation thereby allowing the client to perform a simple task without having to send a request back to the server. Applets are written in Java. [from <http://whatis.techtarget.com/>]

**Back end**—A term used to characterize program interfaces and services relative to the initial user of these interfaces and services. (The "user" may be a human being or a program.) A "back-end" application or program serves indirectly in support of the front-end services, usually by being closer to the required resource or having the capability to communicate with the required resource. [from <http://whatis.techtarget.com/>]

**CGI scripts**—A set of instructions that use the common gateway interface [CGI] standard to pass a web user's request from a web server to an application program and to receive data back from the application to the web server in order to forward it to the user. [from <http://whatis.techtarget.com/>]

**Content**—Any material that is available on a Federal public web site. [from <http://www.cio.gov/documents/ICGI/ICGI-June9report.pdf>] For agency intranet web sites, the term refers to any material that is available on the intranet site.

**Content management system/software**—Software that facilitates web site administration by providing suites of web-related functionality that may include template for web content construction, on-the-fly page creation from databases, versioning control, workflow, and import/export functionalities.

**Context**—The organizational, functional, and operational circumstances in which documents are created and/or received and used [Society of American Archivists Glossary]. The placement of records within a larger records classification system providing cross-references to other related records.



**Cookie**—Information that a web site puts on your hard disk so that it can remember something about you at a later time. [from <http://whatis.techtarget.com/>]

**DoD5015.2-STD (v. 2)**—A standard, developed by the Department of Defense Records Management Program and endorsed by the National Archives and Records Administration, that sets forth mandatory baseline functional requirements for Records Management Application [RMA] software, defines required system interfaces and search criteria to be supported, and describes the minimum records management requirements that must be met based upon current NARA regulations.

**FAQs**—An increasingly common feature of the Internet (pronounced “Fak” or “Fax” in plural), it is usually provided on a web site as a list of “frequently asked questions” (and answers). The FAQ seems to have originated in many of the Usenet groups as a way to acquaint new users with the rules. Today, there are thousands of FAQs on the World Wide Web. [from <http://whatis.techtarget.com/>]

**Front end**—A term used to characterize program interfaces and services relative to the initial user of these interfaces and services. (The “user” may also be a program.) A “front-end” application is one that application users interact with directly. [from <http://whatis.techtarget.com/>]

**HTML**—Hypertext Markup Language, is the set of markup symbols or codes inserted in a file intended for display on a World Wide Web browser page. [from <http://www.w3.org>]

**Hyperlink**—(aka Hypertext Linking) The HTML syntax for expressing the ability for one Internet document to reference another document located either in another portion of the web site or at another web site on the World Wide Web. [from <http://www.w3.org>]

**Internal target**—The HTML syntax for expressing the ability for an Internet document to reference another portion of itself. [from <http://www.w3.org>]

**Java**—An object-oriented programming language developed by Sun Microsystems, expressly designed for use in the distributed environment of the Internet, that can also be used to build a small application module or applet for use as part of a web page. [from <http://whatis.techtarget.com/>]

**JavaScript**—An interpreted programming or script language from Netscape, which is used in web site development to do such things as automatically change a formatted date on a web page, cause a linked-to page to appear in a popup window, or cause text or a graphic image to change during a mouse rollover. [from <http://whatis.techtarget.com/>]

**Legacy system**—Applications and associated data that have been inherited from languages, platforms, and techniques earlier than current technology. [from <http://whatis.techtarget.com/>]

**Portal**—A term, generally synonymous with gateway, for a World Wide Web site that is or proposes to be a major starting site for users when they get connected to the web or that users tend to visit as an anchor site. [from <http://whatis.techtarget.com/>]

**Records schedule**—A document describing records of an agency, organization, or administrative unit, establishing a timetable for their life cycle, and providing authorization for their disposition [Society of American Archivists Glossary], i.e., destruction or transfer to the National Archives.

**Records series**—File units or documents arranged in accordance with a filing system or maintained as a unit because they result from the same accumulation or filing process, the same function, or the same activity; have a particular form; or because of some other relationship

arising out of their creation, receipt, or use. [Society of American Archivists Glossary]

**RMA [Records Management Application]**—Software used by an organization to manage its records. An RMA's primary management functions are categorizing and locating records and identifying records that are due for disposition. RMA software also stores, retrieves, and disposes of the electronic records that are stored in its repository.

**Structure**—The physical and logical format of a record and the relationships between the data elements. [from [http://www.archives.gov/records\\_management/policy\\_and\\_guidance/electronic\\_signature\\_technology.html](http://www.archives.gov/records_management/policy_and_guidance/electronic_signature_technology.html)]

**Site map**—A linked, graphic or text-based display of a web site's hierarchy, similar to an organization chart. Typically, site maps break down a web site's content into increasingly specific subject areas to help the visitor understand its structure, from the main navigation pages to their subordinate pages. [from <http://www.cio.gov/documents/ICGI/ICGI-June9report.pdf>]

**Web site administrative records**—Federal records generated by the web site program office in the course of its operations.

**Web site content records**—Web site content determined by the agency to meet the statutory criteria in 44 USC 3301 for being deemed a Federal record.

**Web site-related records**— Web site-related records include (1) web content records, which represent information presented on a web site, and (2) web site administrative records, which provide evidence of the management and operations of the web site.

---

## APPENDIX B - Web Sites of Possible Interest

Agencies may wish to supplement this initial NARA guidance with these other applicable resources:

- The Interagency Committee on Government Information's Web Content Standards Working Group's Federal Web Content Manager's Toolkit (<http://www.firstgov.gov/webcontent/index.shtml>)
  - National Historical Publications and Records Commission-funded research Analysis and Development of Model Quality Guidelines for Electronic Records Management on State and Federal Websites ([http://slis-two.lis.fsu.edu/~cmclure/nhprc/nhprc\\_toc.html](http://slis-two.lis.fsu.edu/~cmclure/nhprc/nhprc_toc.html) )
  - Australia's Archiving Web Resources: A policy for keeping records of web-based activity in the Commonwealth Government ([http://www.naa.gov.au/recordkeeping/er/web\\_records/intro.html](http://www.naa.gov.au/recordkeeping/er/web_records/intro.html))
  - Government of Canada Internet Guide (3rd ed.). A general guide outlining and understanding and growing appreciation of the structural complexities of delivering information on the Internet ([http://www.canada.gc.ca/programs/guide/main\\_e.html](http://www.canada.gc.ca/programs/guide/main_e.html))
  - U.S. Department of Justice's Guide for Federal Agencies on Implementing Electronic Processes (<http://www.cybercrime.gov/ecommerce.html#GFA>)
-

### APPENDIX C - Web Schedule Options

**NOTE:** The examples in this appendix do not recommend a specific disposition. They are provided only for illustration of the ways you may structure your schedule items. You must determine the appropriate retention period based on your agency's business and risk mitigation needs.

#### A. Single schedule item for all web records

See [SCHEDULING WEB RECORDS, section 5.1](#)

1. Agency web site content and all related web management and operations records -- Destroy when \_\_\_ old or when superseded, obsolete, or no longer needed for the conduct of agency business, whichever is later.

#### B. Single schedule item for web content and single item for web management and operations records

See [SCHEDULING WEB RECORDS, section 5.2](#)

1. Web site content -- Destroy when \_\_\_ old or when superseded, obsolete, or no longer needed for the conduct of agency business, whichever is later.
2. Web management and operations records -- Destroy when \_\_\_ old or when related web content records are deleted, whichever is later.

#### C. Multiple schedule items for distinct portions of the web site but a single item for all management and operations records See [SCHEDULING WEB RECORDS, section 5.2](#)

1. Web content:
  - a. Pages containing copies of agency issuances -- Destroy when superseded, obsolete, or no longer needed for the conduct of agency business.
  - b. All other pages -- Destroy when \_\_\_ old or when superseded, obsolete, or no longer needed for the conduct of agency business, whichever is later.
2. Web management and operations records -- Destroy when \_\_\_ old or when related web content records are destroyed, whichever is later.

#### D. Multiple schedule items for distinct portions of the web site and multiple items for web management and operations records See [SCHEDULING WEB RECORDS, section 5.2](#)

1. Web content:
  - a. Pages containing copies of agency directives and handbooks -- Destroy when superseded, obsolete, or no longer needed for the conduct of agency business.
  - b. Pages containing annual reports, IG semiannual reports to Congress, and Privacy Act reports -- Destroy when \_\_\_ old.
  - c. Monthly snapshot of web content -- Destroy when \_\_\_ old.
  - d. All other web content -- Destroy when \_\_\_ old or when superseded, obsolete, or no longer needed for the conduct of agency business, whichever is later.

2. Web management and operations records

a. Design records -- Destroy after \_\_\_ year[s].

b. Program management records, including policies and procedures -- Destroy when 4 years old or when superseded, obsolete, or no longer needed for the conduct of agency business, whichever is later.

c. Web site posting logs indicating when pages were posted, updated or removed -- Destroy when \_\_\_ old.

d. All other records -- Destroy when superseded or obsolete.