

World Wide Web Home Page Guidelines

and

Best Practices

Prepared by the

World Wide Web Federal Consortium

Revised November 1996

World Wide Web Home Page Guidelines

and

Best Practices

Revised November 1996

Table of Contents

- I. Executive Summary
 - A. Statutes and Regulations
 - B. Checklist
- II. Content
 - A. Agency-Controlled Content
 - B. Linked Content
- III. Navigation/Organization
- IV. Style/Markup
- V. Additional Points
- VI. Emerging Standards
- VII. Home Page Template
- VIII. Disclaimers, Release Notes and Credits

Appendix A. Glossary of Terms

Appendix B. References

- 1. Specific Federal Agency Standards
- 2. HTML Style Guides and Tools
- 3. Accessibility to Users with Disabilities
- 4. Some of the More Vexing Legal and Policy Concerns
 - a. Questions and Answers
 - b. References
- 5. Examples of Best Practices

Appendix C. WWW Security Considerations

Your comments are invaluable in gathering and publishing effective guidelines.

Please direct your questions and/or comments to:

Carlynn Thompson [cthompson@dtic.mil
Chair, WWW Federal Consortium
(703) 767-9170 Fax: (703) 767-9174

Valerie Gregg [vgregg@census.gov]
Vice-Chair, WWW Federal Consortium
(301) 457-4102

I. Executive Summary

The World Wide Web (WWW) Federal Consortium was founded in 1994 by the National Science Foundation in collaboration with the U. S. Nuclear Regulatory Commission. The Consortium was created to provide a forum for collaboration and information exchange among the participating agencies, resulting in a successful leveraging of WWW technology resources. The Consortium seeks to advance the goals of its member agencies, provide advice and peer review to NSF, and foster collaborative research and development.

In late 1995, the Consortium posted WWW guidelines for use and/or consideration by Federal agencies on our server (<http://skydive.ncsa.uiuc.edu/>). Users from many sectors used the guidelines, asked for help and offered suggested enhancements. The nature of user inquiries and suggestions changed as the rapid pace of innovation swept through the WWW community and the information technology paradigm continues to change almost daily. With the proliferation of Federal WWW sites, civil servants and citizens can get more information than ever before and are doing so in increasing numbers on a more frequent basis. Federal WWW sites are trying to keep pace, taking advantage of Internet innovation, improving services to their user constituencies, and doing so with fewer resources. In many instances technological change far out-paced strategic planning and business process redesign, raised implementation concerns and policy issues.

Perhaps the most significant change from a year ago is that now the Web is far beyond its initial use as a way of putting existing hard-copy documents on-line. "Publishing" in the traditional sense of distribution of static printed documents is no longer the only paradigm. Increasingly, the Web is being used for much more dynamic purposes, e.g. delivery of an on-the-fly-constructed result of a user-initiated query, a collaborative chat session, submission of forms-based information to the server from the client, and custom documents constructed dynamically based on the interests of the user. The variety of modes of interaction is much different than the "broadcast" one-to-many mode of early WWW sites, and requires those with WWW managerial and technical responsibilities to continually revisit policy and technical designs with an open and flexible mind. "That's the way we've always done it" is not an appropriate philosophy these days. The Web and the Internet is revolutionary, not evolutionary, and re-inventing our practices and processes is a necessity, not an option. That is why the Consortium intends to regularly update and revise these guidelines and provide examples of "best practices".

Another change is that the Federal government is making a major commitment to electronic networking as a primary means of providing information and services to customers. The WWW is a key enabling technology for the vision of "one-stop shopping." If Federal agencies recognize a reasonable set of WWW guidelines for the organization and presentation of information, customers will be able to enter the system at any point and tap not only the specific server they first contact but also the full

resources of the entire Federal presence on the Internet. The rapid growth of WWW services also intensifies the need to establish standards and guidelines to help users find, retrieve, and use the information they need. If the growing number of Federal agency servers is to coalesce and become a true "on-line library," we need to apply proven principles and practices from the publishing world, traditional libraries, and non-traditional, even experimental contributors to the Web.

The guidelines now include policy considerations which Federal Agencies should review as they update and/or make new use of the Internet and expand WWW sites to conduct agency business. The Consortium is committed to providing technical advice as agencies develop policies that not only encourage use of WWW technologies but comply with law and regulation. The Office of Management and Budget's (OMB), Office of Information and Regulatory Affairs provided guidance to the Consortium in matters related to statutes, regulations and policy. Agencies need to involve not only their information systems and services experts (Webmasters and Webauthors), but also public affairs officers, records managers, FOIA or Privacy Act Officers, and the General Counsel when establishing Internet and WWW policy.

In conclusion, the Consortium offers these guidelines as suggestions to help the Federal community accomplish agency missions, while not inhibiting the creativity necessary to improve services to customers. The Consortium tries to avoid placing unnecessary constraints on individual Webauthors or on the design and operation of individual servers and services in developing these guidelines. There is no simple set of rules for agencies to follow. Every organization faces the special circumstances of its particular mission, goals, content, customers, technical capabilities, and organizational culture. These guidelines provide suggestions to: create user friendly sites; ensure high quality and consistent content; and encourage logical organization and presentation of information.

I. A. Statutes and Regulations

The following statutes and associated regulations may be pertinent for consideration in complying with legal and regulatory requirements in the collection, dissemination, access, preservation and privacy of Government information resources. A more detailed discussion relating to some of these statutes can be found in Appendix B.4.

A. Americans with Disabilities Act of 1990 (42 U.S.C. 12101 note) and the Rehabilitation Act Amendments of 1992 (29 U.S.C. 794) and General Services Administration Regulation.

B. Computer Security Act of 1987 (40 U.S.C. 759 note).

C. Copyright Act of 1976 (Title 17, United States Code, Sections 101-810.) and Copyright Basics, Circular 1, Copyright Office, Library of Congress, Washington, DC, January 1991.

D. Establishment of Government Information Locator Service, OMB Bulletin No. 95-01.

E. Federal Depository Library Program (44 U.S.C. Section 1902).

F. Federal Records Act (44 U.S.C. Chapters 29, 31, 33, 35), National Archives and Records Administration Regulations (36 CFR Chapter 12, Subchapter B, "Records Management."

G. Freedom of Information Act (5 U.S.C. 552).

H. Information Technology Management Reform Act of 1996 (40 U.S.C. Chapter 25), Executive Order 13011.

I. OMB Circular A-130, "Management of Federal Information Resources."

J. Paperwork Reduction Act of 1995 (44 U.S.C. Chapter 35), 61 CFR 6428 (February 20, 1996).

K. Privacy Act (5 U.S.C. 552a).

I. B. Checklist

The Consortium developed "The Home Page Checklist" for guiding the development, maintenance and enhancements of agency WWW sites.

Home Page Checklist

CONTENT

- The content of a page should be related to the organization's function and mission.
- The specific purpose for each top-level home page should be stated.
- WWW content should be reviewed before release.
- Each home page should identify and display a responsible party and provide contact information.
- Information within a document should be accurate.
- A home page should be syntactically correct.
- Spelling and grammar should be correct within a document.
- Information within a document should remain current.
- Document content or data should comply with embargoes.
- Documents should comply with organizational templates.
- External links should be germane to the purpose of the document.
- External links should display the title and the URL.

NAVIGATION/ORGANIZATION

- Each page within a collection should have a link to the collection's home page.
- Each home page should link back to the sponsoring home page.
- Links should be checked periodically to avoid "dead links."
- When the home page is moved to a new location, leave a note at the old location.
- Use graphical clues, such as color, icons, wallpaper, consistency of fonts, etc. to give Web visitors a sense of location or local environment.
- Provide search capabilities for large or complex WWW sites.

STYLE/MARKUP

- Every home page should have a title.
- The title and top level heading for a home page should be the same.
- Include appropriate header comments, such as HTML version.
- Every title/heading should reference the theme of the home page.
- Paragraphs should be clear and concise.
- Horizontal rules can be used to separate dissimilar paragraphs.
- Text highlighting (italics, bold, underline) should be used sparingly.

- ___ Appropriate links should be provided to glossaries, indexes, footnotes, external documents and table of contents.
- ___ Footers should be separated from the body by a horizontal rule.
- ___ Footers should provide logical navigational aids.
- ___ The home page footer should display the URL for the home page.
- ___ The home page footer should include the 'mail to' for the responsible party.
- ___ The home page footer should include the last date the document was updated.
- ___ Navigational icons should be standardized across the agency.
- ___ Agency logos should be used where appropriate.
- ___ Large documents should be divided logically.
- ___ MIME types should be restricted to GIF, TIFF, JPEG, MPEG, AU, and SND.
- ___ Thumbnail images should be used to link to large images.
- ___ Images should not be wider than 472 pixels.
- ___ Avoid long, thin images such as specialty horizontal rules.
- ___ The width and height of each image should be specified in the HTML tag.
- ___ Every graphic should have an associated and meaningful (ALT) text.
- ___ Image maps should have alternate text-based selection mechanisms.
- ___ Audio clips should have text transcripts or descriptions.
- ___ Every page should be tested with multiple viewers (e.g. Lynx, Mosaic, Netscape, Microsoft Internet Explorer).
- ___ Write to the current HTML standard.
- ___ Proprietary markup should be avoided.
- ___ Tables should provide alternate preformatted text markup.
- ___ Consider on-the-fly generation of standard page elements, such as footers, images, etc. through the use of CGI scripts to ensure easier maintenance and consistency of pages.

II. Content

II. A. Agency-Controlled Content

Statement of Purpose

Use of the WWW should be incorporated into the agency's strategic planning process and should support an identifiable business purpose. The content of all pages on agency WWW servers should be related to the function and mission of the organization. Each organization's home page should include, or link to, a specific statement describing the purpose and content of the WWW site.

Approvals

Each organization within an agency should designate responsible parties for reviewing and approving WWW content. Given the dynamic nature of the Web, it may be appropriate for an agency to establish a policy that delegates the release of WWW content to individual Webauthors. In the case where the material will be made available to the public through an external home page, release procedures and/or authorizations comparable to those for any external publication should be considered.

Responsible Party/Contact

Every home page or grouping of related pages should list an e-mail address, or include a link to an e-mail form which can be used to contact a responsible party regarding the content of the page(s). The e-mail address should not necessarily be that of an agency employee and may be that of the Webmaster for the WWW server on which the document resides.

Where personal service is not important, it is acceptable to establish generic e-mail addresses for agency public points of contact instead of using personal addresses (e.g., library@agency.gov rather than john_doe@agency.gov). However, many people prefer responding to real people rather than impersonal mail locations.

Accuracy

Accuracy of all on-line information is an important goal. This is true for content of WWW documents. The construction of HTML syntax for content should comply with standards, network expectations and netiquette. In terms of content, care should be exercised in the following areas:

- Information on which users may base important decisions (e.g., grant application deadlines, policy guidance, etc.).
- Inaccuracies which compromise meaning (e.g., missing text, misaligned table cells, etc.).
- Cosmetic flaws (e.g., titles not italicized titles, missing dashes, etc.).

Before releasing an HTML document to a public server, the syntax and spelling should be checked and all links verified. Syntax can be checked by using a HTML validation tool such as the WebTechs HTML Validation Service [<http://www.cuitrain.com/lisa/htmls/valid.html>]. Many WWW development tools offer spelling checkers. Web-based spelling checkers are also available including the Webster's Dictionary 2.0 [<http://people.goldendome.com/~spectre/>] with update notes at [<http://www.qualcomm.com/~phenness/webprojects.html>.]

HTML does not precisely mimic the printed page. The formatting adjustments which are required to present information in HTML are not considered to compromise accuracy if they faithfully convey the information in a document. Further, user controlled browser preferences may also alter the document format. If the document format is important, consider delivering the content in Adobe's Portable Document Format (PDF) or in Postscript files.

Netiquette

There are many standard practices that have developed over the years on the Internet. As a developer using the Internet, it is the Webmasters and Webauthors responsibility to learn the informal rules known as netiquette. There are many references available on the Internet. The Net: User Guidelines and Netiquette - by Arlene Rinald is a useful source of information. [<http://rs6000.adm.fau.edu/rinaldi/net/>]

Timeliness/Currency

Effective customer service and the credibility of an agency's public access Internet sites depend on providing information which is up-to-date. Information, particularly time-sensitive information, such as grant announcements and press releases, must be posted as promptly as possible. Out-of-date information must be removed or updated promptly. It is recommended that agencies provide dates within home pages so users are aware of the currency. The party responsible for the document or collection should determine whether to retain an electronic version corresponding to the outdated version.

Embargo/Release

Information release must comply with any embargoes. Whenever possible, electronic release should be coordinated with release of printed publication so that press releases and other publicity can mention electronic availability and include the Uniform Resource Locator (URL).

Organizational or Agency Home Pages

Each organization within an agency, that has both the authority and responsibility to serve WWW pages, should maintain a home page describing the organization, its structure and its activities. The organizational home page should include appropriate links up the organizational hierarchy. It should also include appropriate links to subordinate home pages (e.g., other organizations, individual employee home pages,

programs, projects, etc.). Organizational entities that maintain home pages should retain ownership of and responsibility for their information content.

Agency Sponsored Home Pages

Information on WWW servers operated under agency funding, but outside of direct agency control, should be related to the purpose of the award under which the project is funded. The sponsoring agency should be identified prominently and a link should be provided to the sponsoring office's or program's home page. In the case of organizational servers where agency-sponsored information co-resides with other information, agency-sponsored information should be clearly identified and distinguished from non-agency-funded information.

Program and Project Home Pages

Program and project home pages allow organizations to highlight specific efforts or functions within an organization. These collections are optional, but highly recommended. If appropriate, they should link to the sponsoring organization's home page. Links from appropriate topical forums and a program/project index should be maintained on the main agency WWW server to guide users to these special interest pages.

Individual Employee Home Pages

Individual employee home pages can be very useful for those who deal directly with a specific user community, the public, or hold high profile positions. Employee home pages should not be viewed as 'personal' home pages. Content and external linkages should relate to the employee's role as they support the agency's mission. The following is one possible approach to establishing agency policy:

Individual employee home pages are permitted if they relate to and support the functions the individual performs at this agency. All information and external links on an individual home page should support this purpose. A link to the standard disclaimer must be included near the top of the page. Individuals are responsible for making sure the content of their home pages and related documents are appropriate and approved by organizational management. An individual home page should not be considered as an "electronic office" or an "electronic desktop." Information that might be harmlessly posted on an office wall or bulletin board is not necessarily appropriate to the purpose of an individual employee home page.

II. B. External Linked Content

Approval

The decision to include a link to an external source should be consistent with sound public policy, in support of the agency's mission, and based on the WWW site's "statement of purpose." (See section II. A.)

Context

The ability to link to sources external to your organization is a fundamental part of the World Wide Web, and can add significant value to the functionality of your WWW site. It is advisable to establish objective and supportable criteria or guidelines for the selection and maintenance of external links on WWW pages for which your organization holds responsibility.

For example, it may benefit users of your site to link to commercial resources that complement and enhance the value of your organization's WWW presence, or to WWW sites operated by government contractors, universities or others with whom your organization has established an affiliation or agreement. Alternatively, in some cases it may be decided that links to external WWW pages, or specific categories of external links, do not add value to your site or are not appropriate, and will not be permitted. In any case, it is good practice to provide a statement that establishes basic guidelines for selecting and maintaining external links.

In some instances, it may be necessary or appropriate to place external links in context for the user by including statements explaining the purpose of the link. It is important to avoid giving a user the impression that an agency is endorsing a commercial product. It is also important to avoid giving the impression that an agency is linking to frivolous (or objectionable) sites.

Prior to establishing links to external WWW pages or sites, their purpose and content should be reviewed carefully to ascertain if they are consistent with stated guidelines. External links should be reviewed periodically to ensure their continued suitability. If the content of external site becomes questionable or objectionable, remove the link and reassess its value to your organization's WWW site.

In summary, practical due diligence, common sense, and a stated guideline are essential to establishing and maintaining external links that tap appropriately diverse resources on the Web. All external links should be clearly identified as such by including the page title (or appropriate description) as well as the URL. Displaying the URL provides the added benefit of allowing users who are working from a printout the ability to locate the referenced source.

III. Navigation/Organization

Presenting a Unified Picture

An important goal for each agency's collective World Wide Web services is to offer each user full access to the entire expanse of the distributed collection, regardless of the point at which the user enters the system. Reaching the goal of providing the convenience of "one-stop shopping" in a widely distributed system may require that each server sponsored by the agency, in addition to serving its own particular constituency, provide links to other organizational servers. In some cases it is useful for "virtual servers" or servers which knit together disparate resources, to make a strong attempt to appear integrated, through common styles, buttons, environments, tools for the user, etc.

Agency servers should be designed to:

- Include links to each organization and agency program.
- Accommodate cross-links among organizations, programs, projects, and individuals.
- Highlight agency initiatives.
- Provide a comprehensive catalog of the agency's publications and products.
- Host a collection of links to agency-sponsored Internet sites and related external resources.
- Sponsor topical forums on important agency initiatives and subjects through the use of discussion groups and provide collections of documents or resources.
- Help users identify agency resources and services available to them by subject, role or geographical location.
- Provide a key-word searching capability on large sites.

Organization and agency-sponsored WWW materials should be coordinated with the main agency server, in any of the above areas that apply, in order to ensure that users who access the main agency WWW server gain access to the complete range of agency's information and services.

Home Page Links

Documents should be designed to minimize users' reliance on the navigational aids in WWW clients (e.g., back and forward buttons, history lists). The back button, particularly, tends to retrace a path through every page the user has visited rather than logically backing out of a collection. Therefore, it is often useful to provide links which move logically forward and backward through a document or collection, as well as to the table of contents or index.

There should be explicit links on each page within a document or collection back to the entry home page for that document or collection. The home page for each document or collection should include an explicit link back to the sponsoring organization or

program. Each organization home page should include an explicit link back to the agency home page.

Dead Links

Dead links inevitably occur on WWW servers as pages are modified, moved, or deleted. However, dead links can quickly damage a WWW server's credibility. The Webmaster for each server and the responsible party for each organization, program, or project should be jointly responsible for monitoring and maintaining the WWW site. They should correct or remove dead links as soon as possible. External links should be monitored using automated dead link detectors.

When changing URLs on agency servers, coordination may be required with external servers which point to those URLs. When moving a home page to a new URL courteous Webauthors often leave a note providing directions to the new location.

Restricted Access

Sensitive, confidential, or privacy information should not be placed in publicly available directories. In some instances, there may be a need to place documents that are not officially public (i.e., discussion drafts, prototypes, content which is in development, etc.) in a non-private directory for access by a geographically distributed workgroup, test group, or team. The sponsoring organization is responsible for determining whether to have the Webmaster password-protect the materials to prevent access by unauthorized individuals.

Documents and collections that are not public (i.e., not yet published, not fully marked up or tested, internal working group notes, etc.) should not be linked to publicly accessible documents or placed in publicly available directories without information on the restrictions and an explicit notice such as: "Coming soon expanded information about..." or "Internal working documents..." Pre-release information that is available on the open Internet should be restricted by domain, IP address, or password.

Postings to agency WWW sites are official agency disclosures and must be consistent with other agency disclosures of the same or similar information. For example, if requests for draft agency documents are routinely denied from Freedom of Information Act requesters as pre-decisional to protect the integrity of the agency's deliberative process, then it would be more appropriate to post draft documents on the agency intranet server rather than on the public WWW site.

It may be necessary to coordinate with the Webmaster to explicitly exclude restricted access documents from site-wide full-text indexes.

IV. Style/Markup

Titles

Every page should have a title. It is generally displayed in the title bar of the browser and is also displayed in the browser's saved hotlist or bookmarks. The title should be as short as possible but fully informative and specific (e.g., "FY 1996 Agency Budget" is preferable to "Budget"). By convention, the title is the same as the heading for the page.

Headers

Every home page should have a top-level <H1> header near the top of the first screen which clearly identifies the theme of the home page. An <H2> header should be used for continuation pages. Like the title, the header should be as short as possible but fully informative and specific. By convention, the top-level header and the title for each page should be the same. Lower-level headers (e.g., <H3>, etc.) may be used if appropriate to the document. Header markup should not be used to emphasize entire paragraphs.

Documents which are divided into multiple pages should reference the home page's theme in the header of each continuation page. This will help identify the document to users who may arrive at the page without knowing its context, (e.g., as the result of a full-text search) and will make the hotlist meaningful. For example, this document is part of the collection belonging to the World Wide Web (WWW) Home Page Guidelines. It should be referenced in an <H2> heading (and in the title) as IV. Style/Markup

Body

Paragraphs within the body of a document should be clear and concise. Where the audience has a limited knowledge of the subject being addressed, it is often desirable to hyperlink explanatory information. Hyperlinks to a glossary, footnotes, and external documents provide additional information to less informed readers. Other effective uses of hyperlinks include graphics, tables, surveys, and indexes.

Care should be taken in separating and emphasizing content within a page. Horizontal rules <HR> can often be used effectively to separate themes within a page. However, the overuse of italics and bold can make text difficult to read.

Consider using functional markup, such as , rather than appearance markup, like <bold>, to give the user more control over the way the final document is viewed.

Talking about mechanics rather than substance is considered to be in poor taste. For example, the sentence "You can read more about this product in the tutorial which is linked here to this home page" is better stated "The Tutorial will help you learn about this product."

Standard Footer

At a minimum, most home pages have footers which are separated from the body (usually by a horizontal rule<HR>) and contain the following information:

- The URL (for the home page footer).
- The last date the document was updated.
- An e-mail address of a responsible party for the home page.

In addition, the footer most often contains navigational aids such as a mapped bar or buttons which allow the user to move logically through the document. For documents, arranged like a book, the end of page links are to the table of contents, to the next chapter, to the previous chapter, and to the entry home page. These links are frequently tied to icons (see Standard Icons below) or mapped graphics.

Standard Icons

Pages on the main agency WWW server should use standard icons for navigational buttons through standard references. Icons, like all images, should specify the alignment, width, and height to facilitate faster loading in advanced browsers. The standard icons used in documents might be:



Previous chapter. This icon returns to the top of the previous logical (and, most often, physical) page (e.g., to the previous chapter).



Next chapter. This icon goes to the top of the next chapter.



Table of Contents. This icon returns to the beginning of the table of contents.



Back to Home Page. Located only on the first page of the document, this icon links to the entry home page.

The standard agency seal should be used whenever a seal is necessary. If a different size or color is needed, it should be included in the standard icon directory on the WWW server for all Webauthors to use.

An agency Banner can be used to set aside the start of each home page.

Offices and programs are encouraged to develop graphic banners and icons (within the guidelines defined in this document for sensible and economical use of graphics) to help give a graphical identity and consistent look-and-feel to their pages.

File Formats

The choice of file formats used should be based on the following considerations: (1) the intended use of the material by the target audience; (2) the accessibility of the format to the target audience; and (3) the level of effort required to convert the material to the format.

In the interest of making information readily available to as wide an audience as possible, WWW servers should avoid making information available only in proprietary file formats (e.g., WordPerfect, Microsoft Word, Microsoft PowerPoint, SAS, Adobe Acrobat Portable Document Format, etc.), except in cases where the target audience commonly has access to such formats. Links to files in proprietary or unusual formats should be explicitly noted.

Material intended to be viewed, read, or browsed on-line should be prepared in HTML format (for text and tables) and GIF (for graphics). JPEG format may be used instead of GIF for photographic material where there is a need to preserve a large number of colors. In such cases, JPEG produces smaller files with minor loss of image precision.

Portable document formats, such as Adobe Acrobat, should not be used as the primary format unless converting the material to HTML is not feasible. Although it is easier in many instances to create PDF than HTML, there are drawbacks: the contents of PDF files are not included in site-wide full-text search indexes, PDF viewers are not embedded in most WWW browsers, and PDF viewers require more powerful hardware for on-line viewing than a WWW browser alone.

Material intended to be downloaded for off-line print or display should be prepared in one of the following formats, which are listed in descending order of preference:

1. HTML and GIF or JPEG -- Same as materials for on-line viewing.
2. Adobe Acrobat (.PDF), Envoy (.EVY), or Common Ground Digital Paper (.DP) -- Include link to downloadable free viewer.
3. Postscript (.PS) -- Postscript produces excellent print quality but cannot be displayed on screen by most users and cannot be printed by many users with low-end printers.
4. Microsoft Rich Text Format (.RTF) -- RTF is easily created from most word processors and is more widely usable than native word processor formats such as Microsoft Word or WordPerfect. However, its reproduction of fonts and page layout can vary depending on the user's font set.

5. Proprietary formats (e.g., WordPerfect, Microsoft Word, Excel, PowerPoint, Freelance, etc.) should only be used if: (a) conversion to one of the above formats is not feasible; (b) the intended audience is known to have ready access to software which can handle the proprietary format; or (c) the intended use is data analysis or manipulation (see below). If use of a proprietary format is unavoidable, use an earlier, more widely available version if possible (e.g., WordPerfect 5.1, not 6.1).

Material intended to be downloaded by the user for off-line analysis or manipulation should be prepared in one of the following formats, which are listed in descending order of preference:

1. Raw data files:
 - comma separated values, text in quotes
 - column formatted (mainframe data)
 - space delimited (purely numerical data)
2. Proprietary formats (SAS, SPSS, etc.)

Data files should be accompanied with adequate documentation of the file content and structure in a widely usable format such as ASCII or PDF. Documentation for executable programs should include instructions for installation and use, as well as specifications for the platform needed to run the package (i.e., memory, disk space, operating system, etc.).

Material in formats other than HTML should be linked to an HTML page which describes the material in such a way that users of site-wide full-text search facilities can find material of interest.

Large or Complex Documents

Large documents (greater than five pages) should be organized into sections or chapters and linked together. If the material is meant to be read consecutively, then a table of contents and division by chapter may be most appropriate. If the material is meant to be accessed randomly, then a division by section with key word links to appropriate sections may be best.

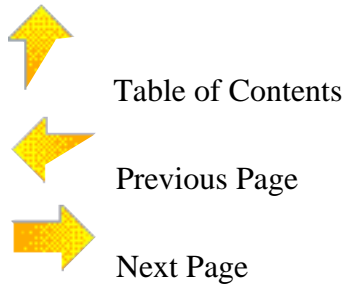
To assist users in navigating sectioned documents, each page should include a navigation mechanism which allows the user to logically progress through the document.

The following HTML code used at the end of each section will aid in logical navigation:

```
<A href="hmpggl.html#toc"><IMG SRC="icons/contents.gif" ALIGN="bottom"
  ALT=" " WIDTH="30" HEIGHT="30" >Table of Contents</A>
<A HREF=hmpggl03.html#Sec3><IMG SRC="icons/LeftArrow.gif"
ALIGN="bottom"
```

```
ALT=" "WIDTH=30 HEIGHT=30>Previous Page</A>  
<A HREF=hmpgg105.html#Sec5><IMG SRC="icons/RightArrow.gif"  
ALIGN="bottom"  
ALT=" " WIDTH="30" HEIGHT="30">Next Page</A>
```

appears as:



Links to Large Files

Links to files larger than 100 kilobytes should include an explicit note of the file size.

Multimedia

Multimedia is supported through the set of standard file types specified by filename extensions. The browsers associate files with these extensions. The most frequently used multimedia file types are:

GIF [.gif] The Graphics Interchange Format was developed by CompuServe in 1987. It is the only image type that can be used for inline images on all platforms. It supports 256 colors. GIF also supports the creation of transparent backgrounds through the 89a standard.

TIFF [DOS .tif, UNIX .tiff] The Tagged Image File Format was designed by Microsoft and Aldus as a desktop publishing standard.

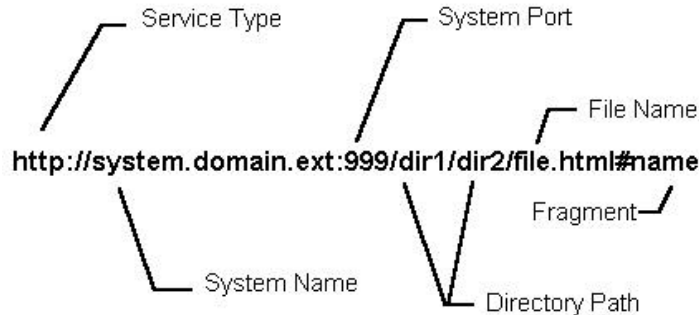
JPEG [DOS .jpg, UNIX .jpeg] The JPEG image standard was developed by the Joint Photographic Experts Group. JPEG is a bit map standard with compression. Not all browsers support JPEG images.

MPEG [DOS .mpg, UNIX .mpeg] The most common movie format for the Web. It is based on the JPEG image standard. Viewers exist for all platforms.

Basic Sound [.au .snd] This audio standard was developed on UNIX platforms. Sound players exist on almost every platform.

Uniform Resource Locators (URLs)

A Uniform Resource Locator is a compact representation of the location and access method for a resource available via the Internet. The main component parts of a URL are:



URL Components

Service Type - Signifies to the browser the type of server to be contacted. WWW servers support the following types of services:

mailto:	Electronic Mail
news:	USENET news
telnet:	TELNET protocol for Interactive Windows
gopher:	Gopher Protocols
prospero:	Prospero Directory Services
wais:	Wide Area Information Servers Protocol
file:	Host specific (local) Files
ftp:	File Transfer Protocol
http:	Hypertext Transfer Protocol
nntp:	USENET news using NNTP access

System Name - The fully qualified domain name of the server of the data being requested, e.g., 'www.whitehouse.gov'.

System Port - The network socket address for specific services. HTTP defaults to port 80. Therefore, it does not need to be specified unless the receiving system is using a non-default value port.

Directory Path - Each level of the directory path is indicated by a '/'. The first '/' indicates the top (root) directory. An example of a directory is '/usr/smith/guide/'. Most browsers will display the contents of a directory when no File Name is specified.

File Name - The name of the data file itself. HTML files will end '.htm' if they originated in a DOS environment, or '.html' if they originated in a UNIX environment.

Fragment - While not properly a part of the URL, a fragment (indicated by the leading '#' character) specifies an anchored location within an HTML document.

For example, the beginning of the table of contents in a document can be anchored by the name toc:

```
<A NAME="toc"></A>
```

and referenced from other documents by

```
<A HREF="hmpggl.html#toc">...<A>
```

Within a server, relative URLs should to be used, for both anchors and images, in order to make documents and collections portable. URLs which reference other servers should always be fully stated (absolute). Common usage of relative URLs are:

```
<!-- to display the image home.gif from the sub-directory icons-->
```

```
<IMG SRC="icons/home.gif">
```

```
<!-- an anchor to the glossary term "hype" which resides in the current directory -->
```

```
<A HREF="glossary.html#hype">hype</A>
```

URL names should use simple, understandable words and should be kept as short as possible without becoming cryptic. The intent is to use names which are easy to read, remember, and type. URL names are most often created in lower case. This is easier for users since the URL is case sensitive.

Images

The appropriate use of images is to help convey information or to create a consistent and recognizable "look and feel" for a collection as well as to convey meaningful information which is not easily conveyed by words. The judicious use of images will help users remember your home pages and will attract frequent usage by the community. The following hints will help make pages effective:

- Images should be as small as possible. Use thumbnail images or text to link to pages with large images. Image resolution may often be reduced without compromising the information conveyed. An example of the typical use of a thumbnail image is:



(Mona Lisa, Leonardo da Vinci, 306 kbytes)

- As a general rule, images should be no wider than 472 pixels, in order to display on the typical WWW browser's 500 pixel wide viewing window on a 640 by 480 monitor.
- Image file size can often be significantly reduced (without compromising the information conveyed) by reducing the color depth, especially for non-photographic material such as charts and graphs.
- Limit the number of images to less than three per page and keep the total size of the images to less than 15 Kbytes.
- Avoid long, thin images such as specialty horizontal rules. These are not effective for users who do not have image capabilities.
- Provide a brief textual alternative description for each image which can be used by text-only browsers.
- Specify the width and height of each image within the HTML <image> tag. This will speed up document formatting on many browsers.

Considerations for Accessible HTML documents

The following guidelines have been offered by the General Services Administration (GSA) in order to provide solutions which can communicate to the broadest possible constituency.

- Every graphic image should have associated text.
- When image maps are used, there should be an alternate method of selection options.
- Include detailed descriptive "comments" with JPEG images.
- Provide text transcriptions or descriptions for all audio clips.
- Make link text descriptive but not verbose.
- File formats other than HTML should be used only as alternatives to rather than replacements for HTML.
- Provide alternate mechanisms for on-line forms.
- All pages should be tested using multiple viewers.
- Do not use proprietary procedural format markup.
- For simple images, such as icons performing the function of bullets, use simple ALT attributes (e.g. "*" or "-"). It should be noted that many users find the use of image bullets annoying, since they take up space and time, and add very little to functionality.

Further explanation along with examples is available in the GSA's document on Writing Accessible HTML Documents [<http://www.gsa.gov/coca/WWWcode.htm>].

Browser-specific HTML

WWW pages should be usable by all major clients to ensure equitable access to the information. Browser-specific HTML should be avoided.

Tables

Tables, like images, can be an extremely effective way to present information. However, like images, they can hamper access to information by visually impaired individuals or those with character-only browsers. <TABLE> markup should be used when it significantly enhances the effectiveness of information presentation. It should be accompanied by an alternative presentation for those whose browsers or disabilities prevent them from using table markup.

An appropriate use of <TABLE> markup would be to present a statistical table, accompanied by a version of the same information as formatted text(<PRE>) that is eighty (80) or fewer characters in width.

V. Additional Points

Reinventing the Wheel

Anyone who is about to implement a WWW site is encouraged to explore the Internet to identify sites which might have already created a similar resource. Webmasters and Webauthors are also encouraged, before writing code, to consider whether to generalize processes and procedures for future reuse.

Webauthors should consult the main agency WWW server and its project staff for advice on WWW design, implementation, and maintenance, as well as WWW-related tools and techniques.

Security

Operation of a WWW server opens up numerous potential security issues which must be addressed. The Webmaster has primary responsibility for the server's security. But Webauthors must be aware of security considerations, particularly in areas, such as Common Gateway Interface (CGI) script development, where potential risks abound. See the WWW Security Considerations in Appendix C for additional information.

Cookies

Cookies are used by the server to track user information across several WWW pages or WWW sessions. WWW site users should be notified of the use of Cookies and the purpose for their use. (See Cookies monitoring in section VII.) An example where cookies might be used to track usage is an end user training service. Cookies can be used to track student progress and automatically connect them to the next lesson.

User Information Collection

Organizations should be careful in collecting information from users. WWW site users should be notified of any user information collection activities and the purpose for its use. There are numerous laws and regulations that govern this activity. Two examples:

- Conducting electronic surveys is subject to the provisions of the Paperwork Reduction Act.
- E-mail addresses in many cases are personal identifiers and their collection and potential retrieval may be subject to the Privacy Act.

Usage Monitoring

Webmasters and Webauthors should review and analyze the usage reports generated by the server for their documents and collections, and use this information to improve their services and public information access. WWW site users should be notified of usage monitoring and the purpose for its use. Information which contains individual identifiers such as e-mail addresses should not be retained for long periods.

WWW Evolution

Webmasters and Webauthors are reminded that the WWW environment is changing rapidly. The number of servers and users is skyrocketing. New software, both commercial and freeware, for browsing, authoring, converting, serving, and searching becomes available every week. Developers should from time to time evaluate the needs of their projects against the changing environment and make appropriate adjustments.

Continuous Improvement

Webmasters and Webauthors should take advantage of the "live" nature of WWW services and the steady stream of technical enhancements by continuously improving their services. While frivolous and arbitrary changes are distracting to users, improvements to services based on user feedback, new ideas for organization, new HTML features, or newly available functionality keep services fresh, provide users additional functionality and reflect favorably on the organization's commitment and expertise.

Copyright and Multimedia Documents

A copyright is the 'rights' of an author or publisher to the 'copy' (text of an article) which that author or publisher produced. This has come to mean the right of intellectual property, whereby authors obtain, for a specific time, certain exclusive rights to their work.

In the United States, copyright protections are exclusively granted under federal law, which derive from Article 1, Section 8, Clause 8 of the U.S. Constitution which provides Congress with the power "to promote science and the useful arts, by securing for a limited times to authors ...the exclusive right to their...writings".

In the United States, and most other countries, a work is copyrighted automatically upon creation. No notice is required nor is registration required with a government agency.

Works which do not enjoy copyright privileges are considered to be in the public domain. Common examples of public domain works are:

- Works for which the copyright has expired. Expiration of a copyright depends on a number of criteria and can run from 28 to 100 years.
- Works of the U.S. Government. These works cannot be copyrighted. However, it appears that works which have been created for the Government by a commercial entity may have some copyright protection from commercial use.
- Non-copyrightable works such as titles, names, short phrases and slogans. (However, these may be trademarks.)
- Works for which the copyright has been forfeited or abandoned. The most common form of copyright forfeiture is the lack of specific copyright notice on materials published

before March 1, 1988. (After that date posting of notice was no longer required to effect a copyright.) Abandonment requires specific language and intent to place copyrighted works in the public domain by the author.

"Fair Use" of a Copyrighted Work.

Copyrighted works can be "fairly used" without fear of copyright infringement for such purposes as criticism, comment, news reporting, teaching, scholarship, or research. Whether the use of a work is fair is determined by balancing these factors:

- The purpose and character of the use.
- The nature of the copyrighted work.
- The amount and substantiality of the portion used in relation to the work as a whole.
- The effect of the use on the potential market for, or value of, the copyrighted work.

Incorporating Works into Multimedia (Internet) Documents.

For multimedia works the following should be considered:

- Extreme caution should be exercised in using digital material downloaded from the Internet because there is a mix of works protected by copyright and works in the public domain on the Internet. Access to these works on the Internet does not automatically mean that these works can be reproduced and reused without permission and/or royalty payment.

- Please note that proper credit should be given for all copyrighted material. When in doubt, credit should be given as if the material was copyrighted.

- In general no more than 10% of copyrighted textual, motion, music, or collections of illustrations or photographs should be included. In the case of independent illustrations or collections, no more than 5 images of an artist or photographer should be included.

- If there is a possibility that multimedia content (e.g. image, movie) may become part of a commercial product in the future or will become widely disseminated, then permissions should be sought before publication of the product.

- If any alterations are made to copyrighted material, then care should be taken to explain the specific changes.

Disclaimers

Agency servers and most agency multimedia documents should carry a Disclaimer of Endorsement and a Disclaimer of Liability. These disclaimers address references to commercial products and services, as well as merchantability and fitness for purpose. Sample disclaimers can be found on the disclaimer template page in section VII.

Electronic Public Disclosure

Postings to agency WWW sites are official agency disclosures and must be consistent with other agency disclosures of the same or similar information. For example, if requests for draft agency documents are routinely denied from FOIA requesters as pre-decisional to protect the integrity of the agency's deliberative process, then it would be

more appropriate to post draft documents on the agency intranet rather than on the public WWW site. To facilitate release of information, agencies may want to develop a WWW statement of responsibility that reminds Webauthors of release criteria and then rely on the professionalism of content developers. An example of this statement of responsibility can be found at [<http://www.cise.nsf.gov/pub/responsibility.html>].

The Government Information Locator Service (GILS)

GILS is an information processing standard and comprehensive indexing scheme that will identify, describe and help find electronic and non-electronic Federal government information resources. Not only will it point the user to the source of the information; as it evolves, GILS will also provide linkages to assist in its delivery. GILS supplements other agency information dissemination mechanisms and commercial information sources. GILS uses network technology and international standards for information search and retrieval so that information can be retrieved in a variety of ways, and so that GILS users can find other information resources worldwide. Agencies should ensure that a GILS record is created for each agency WWW site. Agencies also should assure that all GILS records which identify WWW-retrievable information dissemination products include linkage to that product. See the DefenseLINK GILS for the DoD implementation of at <http://www.dtic.mil/defenselink/locator/morein.html>.

VI. Emerging Standards

The nature and character of the Web is constantly changing. This page is dedicated to standards which are widely accepted, but limited in availability. Soon, many of these will be available to the masses. Even now, some may be appropriate to utilize for specialized groups. All merit consideration and investigation as the future is just around the corner.

Frames

Frames give HTML authors more control over their WWW pages and are an effective way to organize content. When frames are used, the same information should be made available in a way that is usable by browsers which are not frames-capable.

Java

Java is a computer programming language developed by Sun Microsystems. It is Object Oriented, fully threaded, and patterned after C++. Java is interesting to those who work on the Internet because it can open and access objects across the net via URLs, is architecturally neutral, and it runs in a safe (secure) environment.

HotJava is a WWW browser written by Sun Microsystems in the JAVA programming language. Many other commercial browser producers have licensed Java and have incorporated Java capabilities.

To find out more about Java and HotJava try:

Java TM Programming for the Internet (<http://www.javasoft.com>)

Java FAQ
(<http://cuiwww.unige.ch/db-research/java/doc/javafaq/javafaq.html>)

Beyond the animated graphics that are typically seen today, Java will give WWW developers the opportunity to add greater functionality and capability to their WWW services making them much more interactive for the end user.

VRML

Virtual Reality Markup Language is an evolving markup language that promotes a 3-D graphical page format which uses intuitive navigation and interaction rather than text and icon-based interaction. For example, a VRML page may display a picture of a room. Querying objects in the room such as clicking on a television or a book, may start a movie, or display text. Moving from link to link could be accomplished by opening the door to another room. Current estimates are that this technology will become mature enough to use and a hot technology over the next year or so.

To find out more about VRML try:

On the Net: VRML Resources

(http://www.hitl.washington.edu/projects/knowledge_base/compendiums.html)

NCSA VRML Home Page

(<http://www.ncsa.uiuc.edu/General/VRML/VRMLHome.html>)

VII. Sample Home Page Templates

Table of Contents

A table of contents structure with links to each chapter.

Table of Contents Template. WWW Guidelines



Home Page Heading

[Release Notes](#) ● [Disclaimers](#)

Table of Contents Template

- I. Introduction
- II.
 - A.
 - B.

Sample Chapter

A sample page from a chapter.

V. Chapter Template. WWW Home Page Guidelines

Chapter text



[Table of Contents](#)



[Chapter 4](#)



[Chapter 6](#)

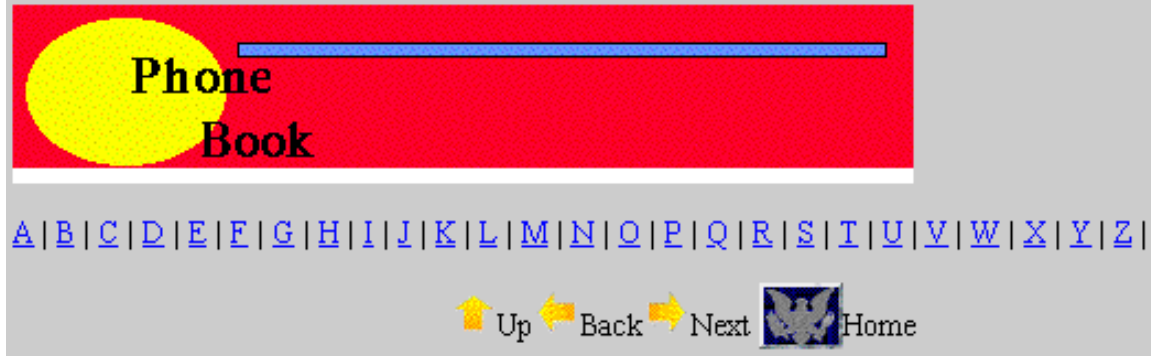


[Comments](#)

Mapped Phone Book

A mapped alphabetic locator at the top of the page takes you to each letter
...each letter returns to the mapped locator
...off-page navigational buttons are built into the locator

Phone Book Template. WWW Guidelines



Cookies Monitoring

This notice has been designed to warn WWW site users that the cookies feature is being employed. One possible approach to a cookies notice:

Cookies Notice:

This site uses a technique known as cookies to provide better services to our users. Cookies allow us to keep a record of your activities while visiting our WWW site. If you object to this monitoring, you may wish to exit the WWW site at this time.

Disclaimers

These disclaimers are styled after many government disclaimers on the Internet... they are not intended to be used as they are... but rather a starting point for thinking about your own disclaimers.

Copyright Status: The U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce these documents, or allow others to do so, for U.S. Government purposes. These documents may be freely distributed and used for non-commercial, scientific and educational purposes. Commercial use of the documents available from this server may be protected under the U.S. and Foreign Copyright Laws. Individual documents on this server may have different copyright conditions, and that will be noted in those documents.

Disclaimer of Endorsement: Reference herein to any specific commercial

products, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government. The views and opinions of authors expressed herein do not necessarily state or reflect those of the

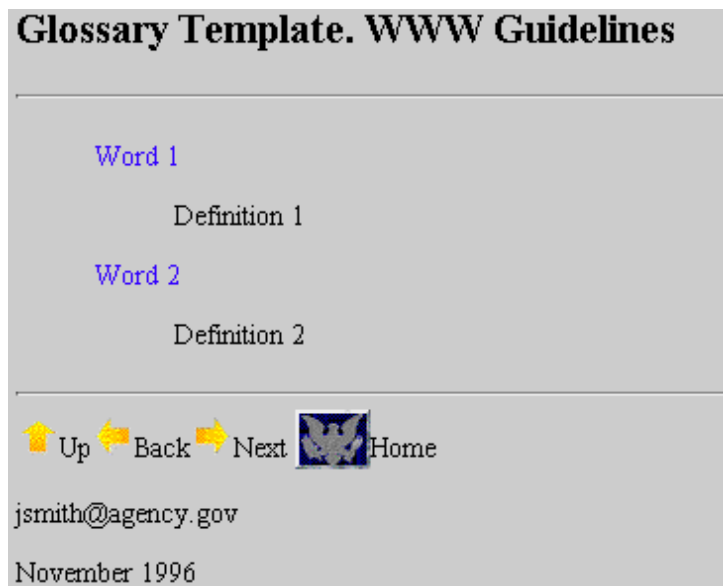
United States Government, and shall not be used for advertising or product endorsement purposes.

Disclaimer of Liability: With respect to documents available from this server, neither the United States Government nor any of its employees, makes any warranty, express or implied, including the warranties of merchantability and fitness for a particular purpose; nor assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed; nor represents that its use would not infringe privately owned rights.

Notice: Information from this server resides on a computer system funded by a U. S. Government agency. The use of this system may be monitored for computer security purposes. Any unauthorized access to this system is prohibited and is subject to criminal and civil penalties under Federal Laws including but not limited to Public Laws 83-703 and 99-474.

Glossary


In a typical glossary a link should be generated from the first occurrence of the word in the document to the glossary entry and a reciprocal link is provided from the word in the glossary to its first occurrence in the document.



Glossary Template. WWW Guidelines

[Word 1](#)
Definition 1

[Word 2](#)
Definition 2

[Up](#) [Back](#) [Next](#)  [Home](#)

jsmith@agency.gov

November 1996

Moved Notice

When a WWW site moves, it is good practice to leave a pointer to the new location.

Moved Template. WWW Guidelines



We have a new home!!!!

Please update your hotlist with our new address:

World Wide Web Home Page Guidelines [<http://skydive.ncsa.uiuc.edu/>]



Up



Back



Next



Home

jsmith@agency.gov

VIII. WWW Home Page Guidelines Disclaimers, Release Notes and Credits

Disclaimers

WWW Guidelines Disclaimer of Endorsement: Reference herein to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government, and shall not be used for advertising or product endorsement purposes.

Disclaimer of Liability: With respect to this document, neither the United States Government nor any of its employees, makes any warranty, express or implied, including the warranties of merchantability and fitness for a particular purpose; nor assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed; nor represents that its use would not infringe privately owned rights.

Release Notes

December 18, 1995 The initial draft release of this document has been reviewed by a small group of Consortium members. Substantive suggestions will be incorporated in the next release.

November 1996 A major revision of the Guidelines was made based on input from the WWW community.

Getting the most of this document.

Like any publication, the greatest benefit can be gained from this document by reading it in its entirety. However, for those who already possess a working knowledge of HTML, an efficient approach is to work from the templates and tools provided in Section VII and employ the checklist from the Executive Summary. The most important thing, of course, is that these guidelines be treated as suggestions to help, not inhibit, the creativity necessary to effectively communicate your message.

Credits

This document is indebted to the Nuclear Regulatory Commission for its vision in leveraging common interest to provide a generic document for use by the federal sector and others. Many people were involved in the formulation, content development and critique. We are indebted to the former NRC Chair, consultants and staff for its original compilation. Significant content and structure were drawn from the DTIC WWW Server Standards and Guidelines, the U.S. Department of Education World Wide Web Server Standards and Guidelines and to the National Science Foundation CISE Policy for Internet Information Dissemination Using the World Wide Web. Comments from the Consortium and other interested readers were also greatly appreciated.

Appendix A. Glossary of Terms.

Absolute Universal Resource Locator or Absolute URL

An absolute URL contains all of the information necessary to reach a resource, including the protocol, the system, the path, and data file name.

ASCII

ASCII is an acronym for American Standard Code for Information Interchange and is an international standard in which numbers, letters, punctuation marks, symbols and control codes are assigned numbers from 0 to 127. It is plain text without style or font specifications.

.aiff

.aiff is a sound file format.

.au

.au is a sound file format.

Browser

A browser is a program for reading hypertext. Since this is the primary function of a WWW client, they are generally called browsers. Common browsers include Mosaic, Netscape, Microsoft Internet Explorer and Lynx.

Byte

A byte is a common unit of computer storage. It is made up of eight binary digits. A byte holds the equivalent of a single character, such as the letter A or a dollar sign.

Client

Client software requests information from another computer or "server". Examples of client software used to browse the World Wide Web are Netscape Navigator, NCSA Mosaic, and Microsoft Explorer.

Common Gateway Interface or CGI

CGI is an interface for running external programs, or gateways, under an HTTP server. Together the HTTP server (httpd) and the CGI programs are responsible for servicing a client request by sending back responses.

Cookies

A "cookie" is a small piece of information which a WWW server (via a CGI script) can store with a WWW browser and later read back from that browser. This is useful for having the browser remember some specific information across several pages; for example, when you browse through a "virtual shopping mall" and add items to your

"shopping cart", a list of the items you've picked up is kept in your browser's cookie file so that you can pay for all the items at once when you're finished shopping.

Dissemination

Dissemination is government distribution of information to the public. Not considered dissemination are distribution limited to government employees or agency contractors or grantees; intra- or inter-agency use or sharing of government information; and responses to requests for agency records under the Freedom of Information Act (5 U.S.C 552) or Privacy Act. (OMB Circular A-130)

Document

A document is a set of information designed and presented as a discrete entity. A technical report is a good example of a document. It may contain logical sub units such as parts, sections, or chapters; but it is typically created, updated, and presented as a single unit. The WWW presentation of a document may consist of one or many WWW pages.

Document collection

A document collection is a set of documents that are logically related, usually by their content, target audience, or origin (e.g., a collection of studies produced by a program, project, or organization).

Federal Records

Federal records are "...materials, regardless of physical form or characteristics, made or received by an agency of the United States Government under Federal law or in connection with the transaction of public business and preserved or appropriate for preservation by that agency or its legitimate successor as evidence of the organization, functions, policies, decisions, procedures, operations or other activities of the Government or because of the informational value of the data in them." (44 U.S.C. 3301)

GIF or .gif

Graphic Interchange Format is commonly used to compress and transfer graphics files to and from on-line services.

Government information

Government information is information created, collected, processed, disseminated, or disposed of by or for the Federal Government. (OMB Circular A-130)

Government Information Locator Service

The Government Information Locator Service (GILS) is essentially an electronic card catalog. It identifies public information resources throughout the U.S. Federal Government, describes the information available in those resources, and provides

assistance in obtaining the information. Ultimately, GILS will consist of a decentralized collection of agency-based information locators and associated information services.

Government publication

A government publication is information which is published as an individual document at government expense, or as required by law. (44 U.S.C. 1901)

Home page

A home page is the entry point to a WWW site or a grouping of related pages commonly referred to as a “document collection.” It is the first page of information received by a visitor to the site or document collection.

Hotlist or Bookmark

A stored list of Internet URLs which an individual may save in order to expedite rapid access to that site. The name on the list is the Title associated with that URL.

HTML or HyperText Markup Language

The HyperText Markup Language (HTML) is a simple markup system used to create hypertext documents which are portable from one platform to another. HTML is the most frequently used language when creating documents for the World Wide Web.

HTTP

HyperText Transfer Protocol is the primary protocol used on the WWW. HTTP performs the request and retrieve functions necessary to display documents stored on remote computers.

Hyperlink

Highlighted text or images that contain links to other information or documents.

HyperText

Text that contains links to other documents, that can be chosen by a reader and which cause another document to be retrieved and displayed.

ICON

A small, pictorial, on-screen representation of an object (file, program disk, document, database, etc.) used in the graphical interface of the WWW.

Information

Information is any communication or representation of knowledge such as facts, data, or opinions in any medium or form, including textual, numerical, graphic, cartographic, narrative, or audiovisual forms. (OMB Circular A-130)

Information Dissemination Product

An Information Dissemination Product means any book, paper, map, machine-readable material, audiovisual production, or other documentary material, regardless of physical form or characteristic, disseminated by an agency to the public. (OMB Circular A-130)

Information Management

Information Management is the planning, budgeting, manipulating, and controlling of information throughout its life cycle. (OMB Circular A-130)

Inline Image

Graphic that appears as a part of the Web page; inline images are coded into Web documents through the HTML language, and load with the WWW page itself.

Internet

An internet is a collection of interconnected networks. The Internet is the largest of the internets. It has a tiered architecture and supports multiple protocols.

Internet Protocol or IP

The internet standard protocol that provides a common layer over dissimilar networks, used to move packets among host computers and through gateways.

Intranet

The intranet is the use of internet technologies within an agency deployed on an internal network based on open WWW technologies.

Imagemaps or Mapped graphics

Image maps are graphic elements that have embedded into them two or more hyperlinks, each hyperlink offering an individual jump to a linked document.

IP Address

The numeric address of a computer connected to the Internet; also called the Internet address.

JPEG

Joint Photographic Experts Group is an image compression format used to transfer color photographs and images over computer networks. Along with GIF, it is one of the most common ways photos are moved over the WWW.

Links

See Hyperlinks.

Mapped graphics

See Imagemaps.

MIME

Multipurpose Internet Mail Extensions is a messaging standard that allow Internet users to exchange messages enhanced with graphics, video and voice.

MPEG

Moving Picture Expert Group, an international standard for video compression and desktop movie presentation.

Netiquette

Network etiquette is prescribed social behavior and manners used on computer networks.

Organization

The internal organization of Federal Agencies varies slightly. However, it is generally hierarchical, such as Agency, Office, Division, Branch, and Individual, Program or Project. Home pages intended to serve the agency itself (i.e. internal servers) are probably best arranged in this hierarchical fashion. WWW sites intended to serve the public are probably better arranged by subject (e.g., mission, initiatives, etc.) with organization as merely one topic.

Page or Home Page

A page, in WWW parlance, is an individual computer file which can be addressed by a hypertext link. Documents and collections are constructed of linked pages.

Portable Document Format or PDF

PDF is used to convert visually rich documents, complete with scaleable graphics and type, for access via the World Wide Web. Documents in PDF can be displayed and printed from many computer.

PostScript

PostScript is a page description language for specifying the formatting of typeset documents or displays developed by Adobe Systems.

Public Information

Public information is “any information, regardless of form or format, that an agency discloses, disseminates, or makes available to the public.” (Paperwork Reduction Act)

Record (as defined by the Federal Records Act)

A record includes all books, papers, maps, photographs, machine readable materials, or other documentary materials, regardless of physical form or characteristics, made or received by an agency of the United States Government under Federal law or in connection with the transaction of public business and preserved or appropriate for preservation by that agency or its legitimate successor as evidence of the organization,

functions, policies, decisions, procedures, operations, or other activities of the Government or because of the informational value of data in them. Library and museum material made or acquired and preserved solely for reference or exhibition purposes, extra copies of documents preserved only for convenience of reference, and stocks of publications and of processed documents are not included. (Federal Records Act, 44 U.S.C. 3301)

Record (as defined by the Privacy Act)

A record is “any item, collection, or grouping in information about an individual that is maintained by an agency. (Privacy Act, 5 U.S.C. 552a)

Records Management

Records Management is “...managerial activities involved with respect to records creation, records maintenance and use, and records disposition in order to achieve adequate and proper documentation of the policies and transactions of the Federal Government and effective and economical management of agency operations. (Federal Records Act)

Relative URL

A relative URL contains only enough information to reach a resource within a server. It includes the path, and data file name. Relative URLs are used to make it easier to move documents and collections among computer systems.

Server

A computer that provides a resource on the network. Client programs such as WWW browsers access servers to obtain information.

Syntax

The syntax is the grammatical rules associated with a language (e.g., HTML).

System of Records (as defined by the Privacy Act)

A system of records is a group of any records under the control of any agency from which information is retrieved by the name of an individual or by some identifying number, symbol, or other identifying particular assigned to the individual.” (Privacy Act, 5 U.S.C. 552(a)(5))

TIFF

Tagged Image File Format is a graphic file format developed by Aldus and Microsoft.

Thumbnail Pictures

A thumbnail picture is a small version of a larger graphic. The typical size of these pictures is often in the range of 40 by 60 pixels. These pictures are placed within documents to provide the user with a flavor for the picture. Usually the picture is linked

to another page which has the expanded graphic. There are a number of shareware tools like Lview which facilitate resizing of images.

Uniform Resource Locator (URL)

URL is an acronym that stands for Uniform Resource Locator and is a reference (an address) to a resource on the Internet. URLs are the doorway to the Internet and the World Wide Web. You provide URLs to your favorite Web browser so that it can locate files on the Internet in the same way as you provide addresses on letters so that the post office can locate your correspondents. The URL includes the protocol, the system, the path, and data file name.

Webauthor

Webauthors are individuals with primary responsibility and expertise in developing home pages for the Web.

Webmaster

Each WWW server has a Webmaster--an individual with primary responsibility for the server. By convention, each WWW server maintains a Webmaster e-mail address (e.g., webmaster@www.ed.gov) to which suggestions and inquiries about the site may be directed.

WWW Server or Web Server

A WWW Server is a computer which provides access to one or more collections of information using World Wide Web formats and protocols. Each WWW Server has a main entry point or home page, although numerous files or pages are usually directly addressable.

WWW Site or Web Site

A collection of information, documents, or databases that is provided to a user community using World Wide Web formats and protocols.

WWW or World Wide Web

The World Wide Web is a network information delivery system that uses HTML as the authoring language and the Hypertext Transport Protocol (HTTP) as the transport protocol. It is used to find and access Internet resources.

WWW Clients or Web Client

Software applications, typically browsers, that access information which is typically distributed over the Internet.

WWW Federal Consortium

The World Wide Web Federal Consortium is composed of sixteen Federal agencies. Participating agencies have the opportunity to adopt and influence the development of NCSA World Wide Web technologies, are interested in the strategic use of these and related information technologies to better accomplish their assigned missions and work together to better use their pooled knowledge base in these areas. NCSA is becoming a national focal point for the fusion of High Performance Computing and communications technologies with the emerging National Information Infrastructure. The major focus for the Consortium for FY 97 is on collaborative work and tools.

Member agencies include:

- National Science Foundation
- Bureau of the Census
- Central Intelligence Agency
- Defense Technical Information Center
- Intelligence Services
- National Aeronautics and Space Administration
- National Biological Service
- National Cancer Institute
- National Institutes of Health
- National Oceanic and Atmospheric Administration
- Nuclear Regulatory Commission,
- National Security Agency
- U.S. Department of Agriculture
- U.S. Department of Education
- U.S. Geological Survey
- U.S. Department of Housing and Urban Development

Ex-officio members include

- The White House
- National Performance Review
- Office of Management and Budget

Appendix B. References

Appendix B. 1. Specific Federal Agency Standards

U.S. Department of Education World Wide Web (WWW) Server Standards and Guidelines

Keith Stubbs, U.S. Department of Education
[<http://inet.ed.gov/~kstubbs/wwwstds.html>]

DTIC WWW Server Standards and Guidelines (Final Draft: 4/24/95)

Carlynn Thompson, Defense Technical Information Center
[<http://www.dtic.mil/hovlane/standards.html>]

Policy for Internet Information Dissemination Using the World Wide Web (Draft: 8/9/94)
Larry Brandt, NSF Directorate for Computer and Information Science and Engineering

[<http://www.cise.nsf.gov/pub/CISEWebPolicy.html>]

Guidelines for Establishing and Maintaining a Department of Defense Web Information Service (12/28/95)

Rick Silva, Office of the Assistant Secretary of Defense (Public Affairs)
[<http://www.dtic.mil/defenseink/webguide.html>]

Appendix B. 2. HTML Style Guides and Standards

Putting Information onto the Web

a collection of documents on authoring hypertext

[<http://www.w3.org/hypertext/WWW/Provider/Overview.html#author>]

NSF/NCSA World Wide Web Federal Consortium Training Materials Page

[<http://skydive.ncsa.uiuc.edu/train/>]

Style Guide for On-line Hypertext

Tim Berners-Lee, World Wide Web Consortium

[<http://www.w3.org/hypertext/WWW/Provider/Style/Overview.html>]

NCSA HTML Style Sheet

National Center for Supercomputing Applications

[<http://www.ncsa.uiuc.edu/Pubs/StyleSheet/NCSAStyleSheet.html>]

Introduction to HTML and URLs

Ian Graham, University of Toronto

[<http://www.utoronto.ca/webdocs/HTMLdocs/NewHTML/intro.html>]

Guide to Writing HTML Documents

Dan LaLiberte, National Center for Supercomputing Applications

[<http://union.ncsa.uiuc.edu/HyperNews/get/www/html/guides.html>]

Composing Good HTML

Eric Tilton

[<http://www.cs.cmu.edu/~tilt/cgh/>]

Yale C/AIM WWW Style Manual

Patrick J. Lynch, Yale Center for Advanced Instructional Media

[http://info.med.yale.edu/caim/StyleManual_Top.HTML]

Yahoo collection of WWW-related sites and documents

[http://www.yahoo.com/Computers_and_Internet/Internet/World_Wide_Web/]

Names and Addresses, URIs, URLs, URNs, URCs

[<http://www.w3.org/pub/WWW/Addressing/>]

Appendix B. 3. Accessibility to Users with Disabilities

Design of HTML Pages to Increase Their Accessibility to Users with Disabilities

Trace R& D Center, University of Wisconsin - Madison

[<http://www.trace.wisc.edu/HTMLguide/htmlguide.html>]

Writing Accessible HTML Documents

Paul Fontaine, GSA Center for Information Technology Accommodation

[<http://www.gsa.gov/coca/WWWcode.htm>]

Appendix B. 4. Some of the More Vexing Legal and Policy Concerns

B. 4. a Questions and Answers

Answers were provided by the Office of Management and Budget, Office of Information and Regulatory Affairs.

What are the Paperwork Reduction Act of 1995 and OMB Circular A-130 and how do they apply to agency use of the World-Wide Web?

What are they?

The PRA addresses two primary areas: 1) agency information dissemination and access practices, under the umbrella of information resources management, and 2) the collection of information from the public and the relative burden that such collections might create. OMB Circular A-130 derives many of its authorities from the PRA and as such is the implementing guidance for the PRA.

How do they apply?

On the dissemination and access side, OMB Circular A-130 directs agencies to implement a management system for all information dissemination products and states that at a minimum, such a system would: 1) assure that information dissemination products are necessary for proper performance of agency functions, 2) would ensure that members of the public with disabilities have reasonable access, 3) facilitate availability of government publications to depository libraries through the Government Printing Office, 4) include as an integral part, communication with the public to include adequate notice when initiating, substantially modifying, or terminating significant information products.

A-130 also instructs agencies to avoid making electronic information dissemination the sole means of disseminating the product unless the agency knows that a substantial portion of the intended audience has ready access to the necessary information technology or that exclusive use of such method will not impose substantial acquisition costs on the users. Agencies are also reminded that their information dissemination practices are to be based upon equitable and timely terms and shall achieve the best balance between the goals of maximizing usefulness of the information and minimizing the cost to the government and the public. These are all common sense issues that may impact agency Web policy.

For information collections, 5 C.F.R. 1320 is the primary implementer of the PRA and it specifically addresses electronic information collection practices such as customer service surveys and questionnaires, including those that are wholly voluntary. In brief, the statute instructs agencies to collect from the public only that information necessary for the proper performance of agency functions and which has practical utility, to seek public comment prior to beginning a collection by publishing a notice of the proposed collection in the Federal Register. Then, agencies are further required to obtain approval for the collection from the Director, OMB and display the assigned control number on the

collection. However, WWW site suggestion boxes or non-specific requests for ideas and comments are not covered by the statute or need to be cleared by OMB. Consult with your agency Paperwork Clearance Officer.

What is the applicability of the Federal Records Act?

Federal records are defined in 44 U.S.C. 3301 as materials, regardless of physical form or characteristics, made or received by an agency of the United States Government under Federal law or in connection with the transaction of public business and preserved or appropriate for preservation by that agency or its legitimate successor as evidence of the organization, functions, policies, decisions, procedures, operations or other activities of the Government or because of the informational value of the data in them.

So, you see, whether information meets the threshold definition of agency records is independent of the medium in which it is produced, maintained, or disseminated. Whether its on a WWW site or not doesn't really matter. When information qualifying as a Federal Record is located on an agency WWW site, the record copy should exist in a format and location that is readily identifiable and appropriate for access and preservation in accordance with the Federal Records Act and NARA guidance. This can, of course, include electronic media as the access and preservation mechanism. Consult with your Records Officer and have them visit the WWW site on a regular basis.

What are the Privacy Act Issues?

We'll discuss the issues below but please go talk to your Privacy Act Officer and get some background on the Act itself. Find out about systems of records, and Federal Register notices, and Privacy Statements, and how your agency handles paper based correspondence and mailing lists, and so on.

Issue One: Electronic Mail Addresses

Because most electronic mail addresses are personally identifying or are at least associated with a specific individual, their compilation into a database or mailing list presents valid privacy concerns. To allay public fears of inappropriate use and loss of control, agencies should generally treat electronic mail addresses as personal identifiers and personal records under the Privacy Act.

The accumulation of electronic mail addresses presents at least two separate issues: 1) the intentional accumulation and compilation into a database, and 2) the incidental collection as a by-product of a software application or otherwise during the normal course of business.

The intentional collection of electronic mail addresses and compilation into a database presumes the intention on the part of the agency to maintain the database (mailing list) in some up-to-date condition and then retrieve and use it in the future. Such a list is a system of records covered by the Privacy Act and is subject to the Act's protections and requirements that your Privacy Act officer will tell you about, e.g., you need to notify the public that you are compiling them and why. But, please note that this discussion pertains to an agency list not your own Rolodex.

The incidental collection of e-mail addresses does not have the same Privacy Act implications as does the intentional collection and therefore you don't need to notify the public. For example, the automatic accumulation of electronic mail addresses by discussion group software does not have Privacy Act implications so long as the accumulation is maintained automatically by the software and not used by agency employees. Another example of incidental collection is the agency keeping the entire text and header of electronic mail messages where the sender's electronic mail address is located in the header. This type of accumulation is analogous to an agency's maintenance of routine paper correspondence and should be handled in a similar fashion.

Issue Two: Using cookies and accumulating a WWW site visitor's session information.

Anecdotal evidence suggests that the use of cookies and the collection of session information poses even greater privacy concerns to WWW site visitors than does the accumulation of electronic mail addresses. This concern is fitting as over time the accumulation of this data can reveal a visitor's personal preferences and particular interests. Agencies should always notify the public when they collect this type of data. Depending on the circumstances, notification may explicitly required by the Privacy Act or provide a beneficial service to the customer. For example, if the session information is associated to an Internet Protocol (IP) address (which is generally held as not a personal identifier) then notification isn't statutorily necessary. But, again, as service to the customer, notice should be given anyway. Consult with your Privacy Act Officer.

B. 4. b. References

A. Americans with Disabilities Act of 1990 (42 U.S.C. 12101 note) and the Rehabilitation Act Amendments of 1992 (29 U.S.C. 794) and General Services Administration Regulation

For text see <http://gopher.usdoj.gov/crt/ada/statute.html>

See also http://www.gsa.gov/coca/tech_act.htm

<http://www.gsa.gov/coca/SECT508.htm>

Americans with Disabilities Act Information on the Web

<http://gopher.usdoj.gov/crt/ada/ada-home.html>

Technology-Related Assistance for Individuals with Disabilities Act of 1988 as Amended in 1994

http://www.gsa.gov/coca/tech_act.htm

Code of Federal Regulations on Reasonable Accommodation (Public Laws 100-407 and 103-218)

http://www.gsa.gov/coca/reas_acc.htm

B. Computer Security Act of 1987 (40 U.S.C. 759 note)

For text see <http://www.net.ohio-state.edu/hypertext/csa-1987.html>

C. Copyright Act of 1976 (Title 17, United States Code, Sections 101-810.) and Copyright Basics, Circular 1, Copyright Office, Library of Congress, Washington, DC, January 1991

For text see <http://www.cni.org/docs/infopols/www/US.Copyright.1976-1.html>

<http://www.cni.org/docs/infopols/www/US.Copyright.Basics.html>

See also <http://www.cni.org/docs/infopols/www/index.html>

D. Establishment of Government Information Locator Service, OMB Bulletin No. 95-01

For text see <http://info.er.usgs.gov/public/gils/omb95-01.html>

E. Federal Depository Library Program (44 U.S.C. Section 1902)

For text see <http://www.law.cornell.edu/uscode/44/1902.html>

F. Federal Records Act (44 U.S.C. Chapters 29, 31,33,35), National Archives and Records Administration Regulations (36 CFR Chapter 12, Subchapter B, "Records Management"

For text see <http://www.law.cornell.edu/uscode/44/>

G. Freedom of Information Act (5 U.S.C. 552)

For text see <http://www.cni.org/docs/infopols/www/US.Freedom.Info.Act.html>

H. Information Technology Management Reform Act of 1996 (40 U.S.C. Chapter 25), Executive Order 13011

For text see <http://www.rdc.noaa.gov/~irm/div-e.htm>
See also <http://www.nismc.navy.mil/don-cio/ombitmra.htm>

I. OMB Circular A-130, “Management of Federal Information Resources,”
<http://www.whitehouse.gov/WH/EOP/OMB/html/circulars/a130/a130.html>

J. Paperwork Reduction Act of 1995 (44 U.S.C. Chapter 35), 61 CFR 6428
(February 20, 1996).
For text see <http://www.rdc.noaa.gov/~pra/pralaw.htm>
See also: <http://www.rdc.noaa.gov/~pra/>

K. Privacy Act (5 U.S.C. 552a,
For text see <http://www.law.cornell.edu:80/uscode/5/552a.html>

To search the Code of Federal Regulations see http://www.access.gpo.gov/su_docs

Appendix B. 5. Examples of Best Practices

Agency Controlled Content

The NASA Homepage [<http://www.nasa.gov/>]

Biological Resources Division - USGS [<http://www.nbs.gov/>]

DefenseLINK [<http://www.dtic.mil/defenselink>]

Timeliness/Currency

today@nasa.gov [<http://www.hq.nasa.gov/office/pao/NewsRoom/today.html>]

Air Force News [<http://www.dtic.mil/airforcelink/news/afns.html>]

THOMAS: Legislative Information on the Internet [<http://thomas.loc.gov/>]

Presenting a Unified Picture for Navigation/Organization

NIST WWW Home Page [<http://www.nist.gov/>]

US Agency for International Development [<http://www.info.usaid.gov/>]

U.S. Census Bureau [<http://www.census.gov/>]

Disclaimers

National Weather Service [<http://www.nws.noaa.gov/>]

United States Information Agency [<http://www.usia.gov/>]

Copyright

National Endowment for the Arts [<http://arts.endow.gov/AboutSite.html>]

URL on Home Page

US Environmental Protection Agency [<http://www.epa.gov/>]

Responsible Party Contact

Welcome to Goddard Space Flight Center [<http://www.gsfc.nasa.gov/>]

Links to external sources (explained)

DFAS Lane (Defense Finance and Accounting Service)
[<http://www.dfas.mil/links.htm>]

Restricted Access

NASA Automated Systems Incident Response Capability (NASIRC)
[<http://nasirc.nasa.gov/>]

Individual Employee Home Pages

Defense Technical Information Center [<http://www.dtic.mil/staff/cthomps/>]

Appendix C. WWW Security Considerations

The following list was gathered from several WWW systems administrators to help organizations focus on security concerns.

- Establish a WWW site security policy.
- Limit the number of people who can modify the WWW Server.
- Run analysis and or hacker tools against your system as often as possible (daily, weekly) to reveal the same weaknesses that hackers will find and exploit. Examples include:
 - Crack
 - the Computer Oracle and Password System (COPS)
 - Tiger
 - Tripwire
 - Ice-pick
 - the Internet Security Scanner (ISS)
 - Security Administrator Tool for Analyzing Networks (SATAN) or something better

Some overlap exists in these programs but each has its strengths and weaknesses. Using as many as possible provides a stronger overall security watch. Remember, if you don't run SATAN or ISS or something like them against your system, someone else will, but they probably won't be sharing the results with you!

- Install alarms so you will know if your server has been the victim of a hacker attack.
- Perform regular audits and review the log files.
- When monitoring or filtering IPs into and out of your system,
 - only local IP addresses should be outgoing
 - local IP addresses should not be incoming
 - private network IPs; 10.0.0.0; 127.0.0.0; 172.16; and 192.168.0.0 should not be incoming or outgoing
- Block disallowed traffic at the router. (example: no telnet connections from *.edu)
- Establish the server on a separate subnet of your network.
- Put the WWW server outside your firewall.

- Run the server as a non-privileged user (i.e., nobody).
- Run server in restricted file space (chrootuid).
- Preferred Network Security Philosophy - deny all, then allow specific services
- Stay current with software releases. Older versions of systems often have security vulnerabilities that are well known to intruders.
- Remove utilities and shells that are not needed.
- When using public domain software, examine the source code carefully for “back doors” and malicious code. Always verify the crypto-checksum of publicly available software.
- Encrypt sensitive systems/administrative files that must be stored on-line.
- Store system accounting records, system administration tool output and proprietary data off-line.
- Use CD-ROMs for storing and serving information and system files that rarely change.
- Be very careful in the use of Common Gateway Interface applications. Hackers often use it as the door to a WWW server. When accepting input from CGI scripts make sure your script checks the input before processing for special characters and systems calls.
- For systems administration access, consider using one-time passwords (such as SKEY or key cards).
- When appropriate, use of passwords to enter and use the WWW site.
- If you detect an intruder in your computer system, leave the system running and disconnect it from the network. Take a complete back-up of the system (Level 0) including the memory. Do not shut down or power off the system. Call in a security response team.
- Do contingency planning for back-up and recovery.
- Monitor Security Advisories such as those posted on the Computer Emergency Response Team’s WWW site: <http://www.cert.org>

- Other Sources of Security Information.
 - Electronic mailing lists
 - CERT advisories: cert-advisory-request@cert.org
 - CERT tools : cert-tools-request@cert.org
 - firewalls: Majordomo@GreatCircle.com
 - Usenet newsgroups
 - comp.security.announce
 - comp.security.misc
 - comp.security.(your system)
 - alt.security
 - WWW site
 - The World Wide Web Security FAQ
 - <http://www-genome.wi.mit.edu/WWW/faqs/www-security-faq.html>
 - Common Gateway Interface (CGI) Security
 - <http://hoohoo.ncsa.uiuc.edu/cgi/security.html>