

Web Audit Step-by-Step Guide

The goal of a web audit is to clean up a website by removing redundant, old, and outdated web items. The purpose of this specific web audit is to prepare for migration to a new content management system.

When you perform a web audit, you determine what action to take with existing content, become aware of content gaps, and plan future content. Your web audit will answer these three questions:

1. [What content do I have?](#)
2. [What content should I keep?](#)
3. [How do I update this content?](#)

See the [complete web audit checklist](#).

AOC provides resources and deadlines for each phase of the content review. If you have any questions, contact the [AOC web management team](#).

Web Audit Step-by-Step Guide

Question 1: What content do I have?

Before you begin to make decisions about your content, you need to make a list of your web pages.

1-1 Create a Content Inventory

Use a spreadsheet or database application to categorize, sort, and describe the information on every page of your site. You should include:

- the overall topic or area to which the page belongs
- the page title and URL
- a short description of the information on the page
- when the page was created/ last updated
- the name and contact of page owner or reviewer
- the page decision status

1-2 Delegate Sections to Review

Update content owners and sites (be sure to note additions and edits)

Work with your team of content owners to divide and conquer

Resources:

- AOC will provide you with a list of web pages and registered owners and developers where available.

Due: TBD

Web Audit Step-by-Step Guide

Question 2: What content should I keep?

After you itemize your content, you will need to decide what to do with it. If you have a lot of content, prioritize what to review first.

Removing old content decreases the number of pages you must manage and improves search results. A goal of removing 15% of your content should be in mind as you clean house.

2-1 Prioritize the Content
Review the “low-hanging fruit,” such as old web pages and top web pages
Sort the content inventory spreadsheet by last updated/last reviewed; these will be your old pages
Use web analytic reports (Omniure/SiteCatalyst, ForeSee/ACSI, top search terms); these will be your top pages
Use web analytic reports to identify low-traffic pages; these may be candidates for removal
2-2 Decide if Current, Outdated, Historical, or Obsolete
Is it redundant? Does it exist elsewhere on the site? If it exists elsewhere on the site, mark the Decision column “DELETE”
Is the web item old? Is it out-of-date and no longer relevant? Mark the Decision column “DELETE”
Is the web item older than 5 years and needs to be saved for historical purposes? Mark the Decision column “ARCHIVE”
Is the web content good “as is” and can be moved to the new content management system with no changes or updates? Mark the Decision column with “KEEP”
Is the web item still relevant, but maybe out-of-date, inaccurate or needs to be edited? Mark the Decision column “UPDATE”

Resources:

- Use [Attachment A](#) to help you determine to keep, delete, or archive pages
- [Contact your records management officer](#) for answers to specific questions
- [General records management](#)
- [Web site records management](#)

Due: TBD

Web Audit Step-by-Step Guide

Question 3: How do I update this content?

The checklist below helps you to review web page items designated “UPDATE.”

3-1 Main Topic/Section Landing Pages
Are short, concise, and easy to scan
Make sense when found out of context via searches, links, or bookmarks
Provide a brief introduction to content in the section
Provide links to top levels of content in the section
3-2 All Pages
If for the public, do not provide information meant for employees
Have a purpose for existing; don't duplicate pages found elsewhere
Avoid use of “welcome” and “under construction”
Are short, concise, and easy to scan
Are well-organized with most important information for the reader at the top
Make sense when found out of context via searches, links, or bookmarks
3-3 Title of Page
Is concise
Is unique (does not repeat other page titles)
Describes the page content
Contains keywords that are used in searches, helps the page rank high in search results, and is important to your users
Provides a meaningful caption for search results
Does not contain acronyms or abbreviations
Does not include ampersands unless they are embedded in terms, e.g., R&D
3-4 Introductory Text
Describes/introduces page content
Contains keywords that are used in searches, helps the page rank high in search results, and is important to your users
Spells out or defines acronyms or abbreviations
Provide a brief, meaningful summary in search results
Does not include ampersands unless they are embedded in terms, e.g., R&D
3-5 Subheadings
Are used to break up long blocks of text
Contain keywords that are used in searches, help the page rank high in search results, and are important to your users
Describe paragraph/section content
Do not include ampersands unless they are embedded in terms, e.g., R&D

Web Audit Step-by-Step Guide

3-6 Body of Content
Is short, concise, and easy to scan (uses white space so pages are easy to scan; uses bullets or enumerates steps, etc.)
Contains keywords that are used in searches, helps the page rank high in search results, and is important to your users
Does not include ampersands unless they are embedded in terms, e.g., R&D
Spells out acronyms before they are used
Uses “you” and other pronouns to speak to the reader
Contains everyday words instead of legalese and jargon; grammar is in current usage
Uses active voice, e.g., “she creates web pages” instead of “web pages were created by her”
Omits excess words
Uses current mention of people, places, and things
Limits use of personal contact information in favor of generic contact information, e.g., “contact us” instead of “contact Chris Smith”
Uses lists and tables to simplify complex material
Use simplest tense, e.g. “We issue a report every quarter” instead of “We will be issuing a report every quarter”
3-7 Links
Tell users where they are going (e.g., avoid “Click here”)
Point to current web addresses, are not broken, and avoid becoming a repository of old, unnecessary content (e.g., newsletters, calendars of events)
Downloadable files are necessary, current, and in appropriate format
Avoid reposting files that already exist elsewhere (e.g., link to Federal Register notices, link to orders and forms)
3-8 Images & Videos
Are current and necessary
For charts: Are adequately described in the text or have descriptive captions
For images: Have descriptive alt text
For videos: Have closed captioning
3-9 Tables
Are used to display data
Are not used for layout
Have appropriate scope & caption in the code

Resources:

- [Plain language tool kit](#) (brochure, PDF, 2 pp.)
- [Plain language training](#)
- Accessibility – [Section 508 resources](#)
- Accessibility – [WCAG resources](#)

Due: TBD

Web Audit Step-by-Step Guide

Attachment A: Decide if Current, Outdated, Historical, or Obsolete

As you review content for relevance, use these basic categories to determine content status and the most appropriate action to take.

If the content is ...	and it ...	then it is...	and you should ...
still relevant	accurately reflects current decisions and/or activities OR is valid until superseded	current	leave it alone
still relevant	requires updates to be current and accurate	outdated	update it and post the new version; if the revision process is lengthy, seek approval from supervisors
still relevant for reference or legal purposes	cannot be updated	historical	archive it
no longer relevant	is not required for reference or legal purposes	obsolete	remove it and take the appropriate action (e.g., back up to CD)

Web Audit Step-by-Step Guide

Complete Audit Checklist

1-1 Create a Content Inventory
Use a spreadsheet or database application to categorize, sort, and describe the information on every page of your site. You should include: <ul style="list-style-type: none">• the overall topic or area to which the page belongs• the page title and URL• a short description of the information on the page• when the page was created/ last updated• the name and contact of page owner or reviewer• the page decision status
1-2 Delegate Sections to Review
Update content owners and sites; note additions and edits
Work with your team of content owners to divide and conquer
2-1 Prioritize the Content
Review the “low-hanging fruit,” such as old web pages and top web pages
Sort the content inventory spreadsheet by last updated/last reviewed; these will be your old pages
Use web analytic reports (Omniture/SiteCatalyst, ForeSee/ACSI, top search terms); these will be your top pages
Use web analytic reports to identify low-traffic pages; these may be candidates for removal
2-2 Decide if Current, Outdated, Historical, or Obsolete
Is it redundant? Does it exist elsewhere on the site? If it exists elsewhere on the site, mark the Decision column “DELETE”
Is the web item old? Is it out-of-date and no longer relevant? Mark the Decision column “DELETE”
Is the web item older than 5 years and needs to be saved for historical purposes? Mark the Decision column “ARCHIVE”
Is the web content good “as is” and can be moved to the new content management system with no changes or updates? Mark the Decision column with “KEEP”
Is the web item still relevant, but maybe out-of-date, inaccurate or needs to be edited? Mark the Decision column “UPDATE”
3-1 Main Topic/Section Landing Pages
Are short, concise, and easy to scan
Make sense when found out of context via searches, links, or bookmarks
Provide a brief introduction to content in the section
Provide links to top levels of content in the section

Web Audit Step-by-Step Guide

3-2 All Pages
If for the public, do not provide information meant for employees
Have a purpose for existing; don't duplicate pages found elsewhere
Avoid use of "welcome" and "under construction"
Are short, concise, and easy to scan
Are well-organized with most important information for the reader at the top
Make sense when found out of context via searches, links, or bookmarks
3-3 Title of Page
Is concise
Is unique (does not repeat other page titles)
Describes the page content
Contains keywords that are used in searches, helps the page rank high in search results, and is important to your users
Provides a meaningful caption for search results
Does not contain acronyms or abbreviations
Does not include ampersands unless they are embedded in terms, e.g., R&D
3-4 Introductory Text
Describes/introduces page content
Contains keywords that are used in searches, helps the page rank high in search results, and is important to your users
Spells out or defines acronyms or abbreviations
Provide a brief, meaningful summary in search results
Does not include ampersands unless they are embedded in terms, e.g., R&D
3-5 Subheadings
Are used to break up long blocks of text
Contain keywords that are used in searches, help the page rank high in search results, and are important to your users
Describe paragraph/section content
Do not include ampersands unless they are embedded in terms, e.g., R&D
3-6 Body of Content
Is short, concise, and easy to scan (uses white space so pages are easy to scan; uses bullets or enumerates steps, etc.)
Contains keywords that are used in searches, helps the page rank high in search results, and is important to your users
Does not include ampersands unless they are embedded in terms, e.g., R&D
Spells out acronyms before they are used
Uses "you" and other pronouns to speak to the reader
Contains everyday words instead of legalese and jargon; grammar is in current usage
Uses active voice, e.g., "she creates web pages" instead of "web pages were created by her"
Omits excess words
Uses current mention of people, places, and things
Limits use of personal contact information in favor of generic contact information, e.g., "contact us" instead of "contact Chris Smith"
Uses lists and tables to simplify complex material

Web Audit Step-by-Step Guide

Use simplest tense, e.g. “We issue a report every quarter” instead of “We will be issuing a report every quarter”
3-7 Links
Tell users where they are going (e.g., avoid “Click here”)
Point to current web addresses, are not broken, and avoid becoming a repository of old, unnecessary content (e.g., newsletters, calendars of events)
Downloadable files are necessary, current, and in appropriate format
Avoid reposting files that already exist elsewhere (e.g., link to Federal Register notices, link to orders and forms)
3-8 Images & Videos
Are current and necessary
For charts: Are adequately described in the text or have descriptive captions
For images: Have descriptive alt text
For videos: Have closed captioning
3-9 Tables
Are used to display data
Are not used for layout
Have appropriate scope & caption in the code