

# How to Interpret Accenture Link Integrity Reports

---

## Purpose

The purpose of this document is to provide high level guidance on how to access, interpret, and identify issues found on the Accenture link integrity reports generated for an individual web site. Link integrity reports are automatically generated monthly for sites that are on the HHS leaderboard and point to broken links both internal and external to the scanned site.

## How to Access

Link integrity reports for individual sites that are on the HHS monthly leaderboard can be accessed in one of two ways:

1. Through the public leaderboard <http://128.231.74.134/HHSLeaderboard.html> (You will need to know the location of the site and drill down to the report).
2. Through the Accenture client (Those familiar with 508 Issue Management can access the link integrity reports in the same manner that the 508 custom reports are accessed)

The reports are available by expanding the 'Quality' option from the left navigation of the individual Accenture site report.



- Skip to main content
- ▼ Quality
  - **Link Integrity** [csv]
    - Rating
    - Summary
    - Graphs
    - Pages With Broken Links
    - Missing Pages
  - External Link Integrity [csv]
- ▶ HHS 508 Custom Reports

## Reports Types

### Link Integrity (Internal Page Links)

The report titled link integrity is based on links internal to the scanned site.

### Scores

The overall score is in the upper right of the page.

Scan Date: February 19, 2014



It is based on a weighted formula where broken links closer to the home page count higher against the score. Although the score gives a representation of the condition of the site, it is better to concentrate on the summary statistics and actual issues flagged.

Report Date:	February 19, 2014
% Unscanned:	28.8% (1360 of 4726 objects)
Site:	www.womenshealth.gov
File:	www.womenshealth.gov_2014-02-19_0415.max
Total web objects:	3366 (includes pages/images/documents)
Total HTML pages:	1761
Pages with broken links:	26
Total broken links:	67
Missing web objects:	18
Pages with unlinked references:	0 (unlinked http URLs contained in page copy)
Average broken link depth:	4.4 (distance from the home page of reachable pages that contain broken links)

#### Issues

- Missing web objects refer to issues with internal non html links such as CSS or JS references

## Views

### Screen view

On the screen, the information appears in two categories:

1. Pages Containing Broken Links: Shows each page on which there are broken links/missing web objects with each page expandable to show specifics on the links/objects.
2. Non-downloadable Pages and Where They Are Referenced: Shows each broken link/missing web object with each item expandable to show specifics on the page they were found

It is recommended to first use the information in Non-downloadable Pages where each item contains:

1. A clickable link to the URL of the broken link
2. The error encountered at the time of the report
3. On how many pages the broken link appears on the site
4. A link to each page where the link appears with:
  - a. Page Title
  - b. Link text used for the broken link
  - c. Line number of the link in the source

14. for-your-heart/ (Err 403 Forbidden)

is linked by 7 pages

[aging/diseases-conditions/heart-disease.html](#) (Heart disease | womenshealth.gov), link text "For Your Heart", line 620

[heart-health-stroke/](#) (Heart Health and Stroke | womenshealth.gov), link text "For Your Heart", line 379

[heart-health-stroke/government-in-action/programs-targeting-minorities.html](#) (Programs targeting minorities | womenshealth.gov), 2 links, link text "For Your Heart", line 602

[heart-health-stroke/government-in-action/website-based-activities.html](#) (Website based activities | womenshealth.gov), link text "For Your Heart", line 602

[heart-health-stroke/index.html](#) (Heart Health and Stroke | womenshealth.gov), link text "For Your Heart", line 379

[minority-health/american-indians/high-blood-pressure.html](#) (High blood pressure | womenshealth.gov), link text "For Your Heart", line 570

[minority-health/latinas/high-blood-pressure.html](#) (High blood pressure | womenshealth.gov), link text "For Your Heart", line 575

### CSV view

The report can also be viewed as a CSV file which users may find easier to manipulate and sort with Excel's built in features. To evaluate via CSV:

1. Select 'CSV'



2. Save and Open the file in Excel
3. When you open the CSV spreadsheet, 6 columns appear:

	A	B	C	D	E	F
1	From	To	Text	Line	Error	Count
2	<a href="http://www.heartdisease.com">http://www.heartdisease.com</a>	<a href="http://www.globalheal.com">http://www.globalheal.com</a>		12	Err 404 No	1
3	<a href="http://www.heartdisease.com">http://www.heartdisease.com</a>	<a href="http://www.globalheal.com">http://www.globalheal.com</a>		1	Err 404 No	1

- a. From: The page where the broken link can be found
- b. To: The link that is broken
- c. Text: The hyperlinked text on the "from" page
- d. Line: The line of the from page where the link may be found
- e. Error: The type of link issue
- f. Count: The number of times this link issue occurs on the page

### Evaluation Tips

- If using the CSV file, sort the spreadsheet rows by the "To" column. This will allow you to see if any of the link issues are reoccurring
- It's better to prioritize the broken links that appear on multiple pages since fixing the link will have a greater impact on the overall score
- Evaluate each unique broken link to determine the actual issue through both the report and navigating to the source page and finding the link on the page and in the source
- For any broken links to internal pages, a site owner is likely to diagnose with further investigation
- For any broken links that are template-related (ends in CSS, JS, etc.), a developer is likely to diagnose with further investigation

## External Link Integrity

The report titled External Link Integrity is based on links external to the scanned site (though in some cases, links may end up being internal to the site).

### Scores

Like the internal Link Integrity report, the score is located in the top right of the page and is based on a weighted formula where broken links closer to the home page count higher against the score. In addition, issues are categorized with unique names and error codes. However, only certain error types are used to calculate the overall score. Therefore, although the score gives a representation of the condition of the site, it is better to concentrate on the summary statistics and actual issues flagged.

Report Date:	February 19, 2014
% Unscanned:	28.8% (1360 of 4726 objects)
Site:	www.womenshealth.gov
File:	www.womenshealth.gov_2014-02-19_0415.max
Total external objects referenced:	9485
Non-connecting external objects:	467
Non-connecting external links:	3781

- Non-connecting external object refers to image, PDF, or other file types that cannot be found

### Views

#### Screen view

On the screen, information is categorized by error type, which varies for each site. Only errors in the “-,” “400,” and “500” series are used to calculate the link integrity percentage (refer to the section ‘HTTP Errors Definition’ at the end of this document for more information). Error type details can be quickly navigated to by clicking the title under the External Link Integrity report.

- [External Link Integrity](#) [csv]
  - Rating
  - Summary
  - Host not found
  - HTTP timeout
  - HTTP connection error
  - Unknown error
  - Service unavailable or host not found by proxy
  - Internal server error
  - Unsupported media type
  - Object is no longer there
  - Object not found
  - Request forbidden
  - Bad request
  - Temporary object redirection, keep same method
  - Redirection (use GET method)
  - Temporary object redirection
  - Object permanently moved
  - Multiple choices for location of object
  - No information to return
  - Request completed

Each error type follows basic the format of the ‘Non-downloadable Pages and Where They Are Referenced’ format of the Link Integrity report.

#### HTTP connection error

1. [http://www.militaryhomefront.dod.mil/portal/page/mhf/MHF/MHF\\_HOME\\_1?section\\_id=20.40.500.230.0.0.0.0.0](http://www.militaryhomefront.dod.mil/portal/page/mhf/MHF/MHF_HOME_1?section_id=20.40.500.230.0.0.0.0.0), 2 links (531ms)  
 is linked by 2 pages  
 repair link on page [illnesses-disabilities/women-veterans-with-disabilities/index.html](#) (Women veterans with disabilities | womenshealth.gov), link text "Heroes to Hometowns", line 587  
 repair link on page [illnesses-disabilities/women-veterans-with-disabilities/](#) (Women veterans with disabilities | womenshealth.gov), link text "Heroes to Hometowns", line 587
2. <http://www.militaryhomefront.dod.mil/tf/casualtyassistance>, 2 links (609ms)  
 is linked by 2 pages

#### CSV view

The report can also be viewed as a CSV file which users may find easier to manipulate and sort with Excel’s built in features. To evaluate via CSV:

1. Select ‘CSV’
  - **External Link Integrity** (csv)
2. Save and Open the file in Excel
3. When you open the CSV spreadsheet, 7 columns appear:

	A	B	C	D	E	F	G
1	Error Code	To	Time	From	Link Text	Line	Is Embedded
2	-1	http://ans	94	http://ww	http://ans	21	
3	-1	http://usi	0	http://ww	U.S. deleg	21	

- a. *Error Code: The type of link issue*
- b. *To: The link that is broken*
- c. *Time: Time in milliseconds before timeout*
- d. *From: The page where the link can be found*
- e. *Link text: The hyperlinked text on the “from” page*
- f. *Line: The line of the from page where the link may be found*
- g. *Is Embedded: Disregard this column*

#### Evaluation Tips

- If using the CSV file, initially sort the spreadsheet rows by the ‘Error Code’ column
- Only error codes of the 400 series (404, 410, etc.), 500 series (500, 503, etc.), and - series (-1, -5, etc.) affect your score
  - The 401 code is an "authentication required" code and normally corresponds to a page that requires a username/password; this also does not count against the score
- Eliminate rows containing other error codes then resort by the ‘To’ column, this will allow you to see if any of the link issues are reoccurring
- Evaluate each unique link to determine the actual issue, often you may begin to see patterns
  - If a broken link shows on all or most pages, it is likely in the header/footer or an attached file like a css or js, so look there first
- Sometimes an external site may be down when the report is run; if the link appears to be working, the issue may not appear on a subsequent report