

Fintest Pro Link Checker

What is a link checker?

A link checker is a software tool that is used for testing hyperlinks in a website or a web page. The test determines whether a link is active or broken. It allows website owners to identify dead or broken links and inform webmasters and developers on invalid Web pages.

A link checker is not a browser. Although the link checker scans web pages and follows links within web pages, it does not render web pages and may not be able to detect all links. Note that some websites generate dynamic content and thus results can vary between scans.

Link checkers can be used in multiple ways, including exploratory testing of web applications, monitoring web sites for integrity and operability, as well as integrations between multiple sites.

How does it work?

The link checker contains two main components, a **crawler** and a **link tester**.

Crawler

A crawler scans (or “crawls”) web pages for links based on the specified crawling mode, and compiles a list of links to test.

Link Tester

The link tester uses the list that the crawler creates and checks each link for validity. If the test results in an error, the link tester saves that response. Certain errors, such as HTTP 429 errors indicate “server Busy” state and may need to be repeated.

Crawling Modes

Selected HTML Page

Once a user specifies a destination, the crawler fetches the first page, scans it for HTML links, and adds these links to the test list. Links embedded in internal styles or scripts, as well as external CSS files or JS files can also be scanned. Only the first page is scanned.

Same Domain

Once a user specifies a destination, the crawler scans the first received page, and then follows links within that page to additional pages **ONLY** if they contain the same domain. For example, if a user scans www.fintest.ca only links with this domain will be crawled.

Please note that all links within the crawled pages are tested and those may include both links with the same domain and different domains. The same domain criteria applies to crawling, not link testing.

Same Root Domain

Similar to the **Same Domain** crawling mode, this mode widens the crawling criteria by following links that contain the specified root domain.

All Links

In this crawling mode, ALL links are scanned without domain limits.

Crawling Options

HTML

This option is always ON and includes a scan of links in the HTML excluding style sections or files and Javascript code.

Internal CSS and Scripts

If users wish to also scan internal style and Javascript sections, simply check that option in the drop-down menu.

External CSS and Scripts

If users wish to also scan external style and Javascript sections, simply check that option in the drop-down menu.

Crawl Mode

All Links (total)

HTML + ☒ Internal CSS and scripts + ☒ External CSS and scripts

Selected HTML Page

Same Domain (www.fintest.ca/*)

Same Root Domain (*.fintest.ca/*)

All Links (total)

Settings

Max links	"429" re-scan	Exclusions
500	Off	Exclusions (Comma delimited)

Max links

This setting is an upper limit on the total number of links that will be tested. Once the crawler reaches this limit, it stops the scan. All links fetched during the scan within this limit will be tested.

"429" re-scan

The HTTP 429 Too Many Requests response status code indicates the user has sent too many requests in a given amount of time. In this case, users may opt to retry such links after waiting a specified elapsed time.

"429" re-scan

Off

Off

20 sec

30 sec

40 sec

50 sec

60 sec

Exclusions

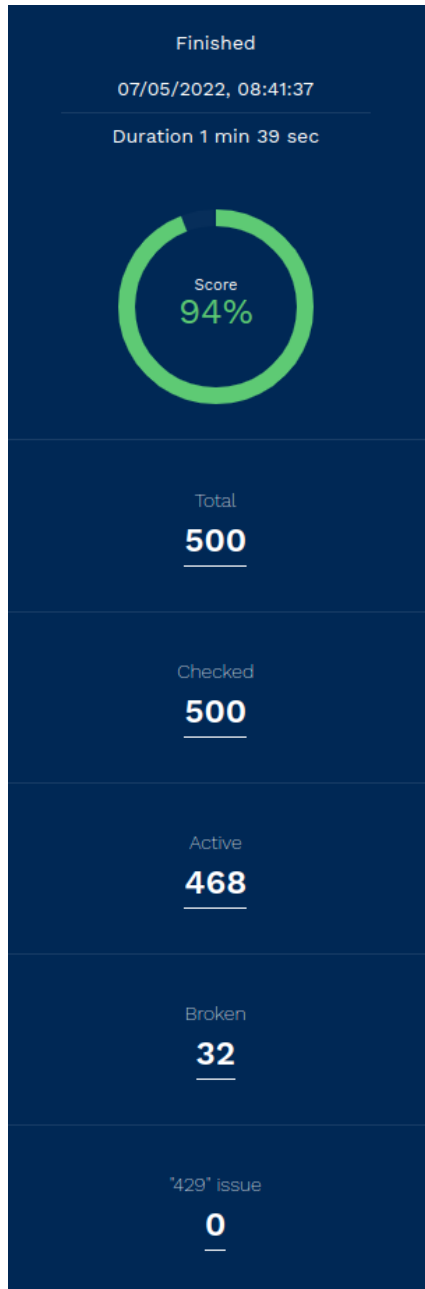
When a user wishes to exclude certain links from being crawled and tested, they can be added in the exclusion field. Multiple exclusions can be added, separated by a comma. Please note that users do not need to specify a full link for exclusion, keywords or partial links may also be used as matching is done via text comparison.

Scan Results

Scan results are updated on the page as they become available. If a scan is canceled mi-way, partial results will be shown.

Summary

The report includes a synopsis of all results in terms of active links, links returned with errors and an overall score. The score is the % of active links compared to the total number of links.



Dashboard

The detailed results section includes information about the links that were found and analyzed. It includes,

Link Schemes

Link Schemes		
Links in HTML		500
Links in CSS		0
Links in scripts		0
Redirects		0

A list of where links were found within web pages.

Reference Schemes

Reference Schemes		
Canonical		1
Alternate		15
Lang href		16
Nofollow		0

A list of links with special attributes that help webmasters with search engine optimization.

Error Types

Error Types		
4	403 (Forbidden)	8
4	404 (Not Found)	21
5	503 (Service Unavailable)	1
5	515 (No answer)	2

A list of links that returned an error code organized by the type of error status code.

Links with Errors

Links with Errors	
< HTML links >	32
< XHTML >	0
< Redirects >	0
< E-mails >	0
< Style sheets >	0
< Fonts >	0
< Images >	0
< Media >	0
< Scripts >	0
< Other >	0

A list of links that returned an error code organized by link type.

Active Link Types

Active Link Types	
< HTML links >	395
< XHTML >	0
< Redirects >	0
< E-mails >	0
< Style sheets >	2
< Fonts >	8
< Images >	46
< Media >	0
< Scripts >	11
< Other >	6

This list summarizes all working links by link type.

Hosts



Hosts	
cnn.com	319
dynaimage.cdn.cnn.com	39
www.cnn.com	13
edition.cnn.com	12
cdn.cnn.com	12
bleacherreport.com	12
edition.i.cdn.cnn.com	9
money.cnn.com	8
facebook.com	5
twitter.com	5


A list of links organized by link host (domain).

Details


Details by Link Category

Users can view additional details about each link grouping by clicking on the link type or link category. A list of the selected links with more details will be displayed. Information includes the number of occurrences of a given link, status code, where a link was originated, and the link type.

Additionally, users can search for keywords to find specific links. And a PDF summary of the given category can be created and downloaded.

<div>  <div> <div>< HTML links ></div> <div>Active Links</div> </div> <div> <input type="text" value="Input text to Search"/> <div> <div>SEARCH</div> <div>PDF</div> <div>RETURN</div> </div> </div> </div>				
Link	Count	Status	Link From	Link Type
http://bleacherreport.com/	2	200	https://cnn.com/world https://cnn.com/style	< HTML links >
http://bleacherreport.com/college-football	1	200	https://cnn.com/style	< HTML links >
http://bleacherreport.com/mlb	1	200	https://cnn.com/style	< HTML links >
http://bleacherreport.com/nba	1	200	https://cnn.com/style	< HTML links >
http://bleacherreport.com/nfl	1	200	https://cnn.com/style	< HTML links >
http://bleacherreport.com/world-football	1	200	https://cnn.com/style	< HTML links >
http://cnnnewsroom.com/	8	200	https://cnn.com https://cnn.com/search https://cnn.com/world https://cnn.com/politics https://cnn.com/business https://cnn.com/health https://cnn.com/entertainment https://cnn.com/style	< HTML links >


Details of Individual Links



Links with issue

Code 404

SEARCH



RETURN

Link details :

URL :

https://aax.amazon-adsystem.com/

Status :

404 (Not Found)

checked at 07/05/2022, 08:40:00

Link Type :

< HTML links >

Count :

7

Nesting level :

0

Link From :

https://cnn.com

1 source

https://cnn.com/search

1 source

https://cnn.com/world

Link From

https://cnn.com

Insight into individual links is available by clicking on links from the categorized view. Users can locate the link in the HTML page, and in cases where there are multiple instances of the link, users view all the locations in the page. Searching for keywords is also available on this details page, as well as viewing the full path of the current link from the start page.

Save or Download Results

From any screen, users can download a PDF summary of the details being viewed. Additionally, from the main dashboard, users can save the results in the extension for later viewing.