

Instructions

Intro

404bypass is designed to help you avoid 404 errors by generating a redirect file for your website. This is especially helpful when a redesign leads to pages being assigned a new URL.

404bypass scans your old and new website, automatically compares the scan results and generates a list of redirects. It also allows you to manually match or add redirect pages.

This program assumes you have basic knowledge or htaccess redirection.

Quick start

- Enter the URL of your old and new site.
- Scan both.
- Click Smart Match.
- View and export the list of redirects

Advanced options

You can load the website information by entering a start URL and letting 404bypass scan it, or you can import it from a file like a sitemap or CSV file.

Scanning

Input the URL of your old website's home page in the start url field and click Scan. A progress bar is displayed below the field. Do the same for the new site.

Once scan is complete, results from the old site scan will be displayed. Results from the new site scan are stored in memory, ready to be matched.

Since one of your sites is likely to be hosted in a staging environment, we included options to adjust the paths in the output. See "Base" section below.

Loading file

Scanning a website will collect more data and yield better matches.

sitemap XML file: will load only the URLs.

CSV file: (need format requirements)

Matching

Click the Smart Match button for automatic matching.

Matching results will be displayed in the right column. ~~The arrow color indicates the accuracy of the match, with darker colors indicating a more accurate match. Click the arrow (optional) to confirm a match.~~ (feature under construction)

Click the X to remove incorrect matches.

Click the search icon to match a page manually. Then select the page from the pop-up window. You can also use the search field to narrow the results.

If the page you need to redirect to is not displayed in the results, use the Custom URL option. If the page is in a different website, check the External link box (this will prevent adding your website's base in the output).

View and Export

All items with a checkmark are included in the output.

Click View/Export button to see the output. You can select the template that suits your need (see Preferences below) and save the file.

If the urls displayed are not what you expect, you may need to adjust the base (see below)

Base

Since the website you scanned or uploaded may not be in the exact location as the real website, we included options to adjust the paths in the output.

For standard redirect, the request url ("old page") is relative to the website root and the destination url ("new page") is absolute.

- If the start url for the old site scan is a directory instead of the website root, input the path to that directory in Prepend old site path.
- If the start url for the new site scan is different location than the actual site, input the location (domain and, if applicable, directory) in Replace new site base.

Deployment

.htaccess

Copy the export to your server's root folder and rename it .htaccess. Open your browser and enter one of the old site URLs. Your site should redirect to the new page.

If your site already has a .htaccess file, open both files in a text editor and copy/paste the entries. You can also copy entries directly from the View window.

Other tools

If using another tool to manage redirection, please follow that tool's instruction. The templating system should give you enough options to export to a compatible format.

Toolbar

Right-click or Control-click the toolbar to open the toolbar options. You can then reorganize the items to your liking.

Preferences

Matching:

404bypass uses advanced processes to find likely match. Higher precision will result in fewer more accurate matches. Lower precision will result in more matches, but may include more errors. (under construction)

Templates:

While the basic htaccess template is the one you're most likely to need, we know that your needs may be different than ours. The templates allow you to leverage the power of 404bypass and output the matching results in a format that best suits your needs. See Templates section below for more details.

General:

By default, URLs that match perfectly don't need to be redirected and are excluded from the output. You can check (uncheck?) the box to include those matches in the output.

Scan Settings

Scan settings can be adjusted separately for old and new site by clicking the cog icon.

- The number of threads is the number of concurrent requests that the engine makes on the server while crawling. More threads will result in a quicker scan but some servers can produce timeouts if bombarded too heavily. Some will detect a script or interpret the requests as an attack and deliberately stop responding to your IP. The default setting should work for most uses.
 - Stop at and Crawl maximum: prevent excessive scanning
 - Ignore trailing slash:
 - Ignore querystrings: will ignore any URL parameters and treat a page such as *example.com/index.php?parameter=1* as *example.com/index.php*.
-

File Upload Settings

This feature is not documented yet.

Templates

Two templates are included by default and should satisfy most needs. You can build a template using the placeholders below. The most commonly used placeholders can also be input using the buttons.

- {status} is the redirect status of the page (usually 301)
- {old_page} is the url of the old page relative to the start url.
- {new_page} is the url of the new page relative to the start url. In the case of an external link, it is the absolute URL.
- {old_base} is the value of Prepend old site path (usually empty).
- {new_base} is the value of the new site Start URL or Replace new site base. In the case of an external link, it is empty.
- {new_title} and {old_title} are the page titles. (not yet functional)
- ↵ is the line break, which can be set in the dropdown.

Literal vs. sanitized URLs: by default, 404bypass will check if a URL is missing a protocol and will prepend http://. If you want your input to be used literally, use the placeholders {...} (TBD-under construction)

Escaping special characters: If you need to escape certain characters in your output, wrap the placeholder in an escape tag. For example {esc_js{new_title}} (under construction). Available escapers will be:

- esc_html: HTML
- esc_xml: XML
- esc_js: Javascript
- esc_json: JSON

Encoding: Files exported from 404bypass use UTF-8 encoding (???)

Template examples:

PHP:

```
<?php
    var old_page = "{old_page}";
    var new_page = "{root}{new_page}";
    var status = "{status}";
?>
```

JSON:

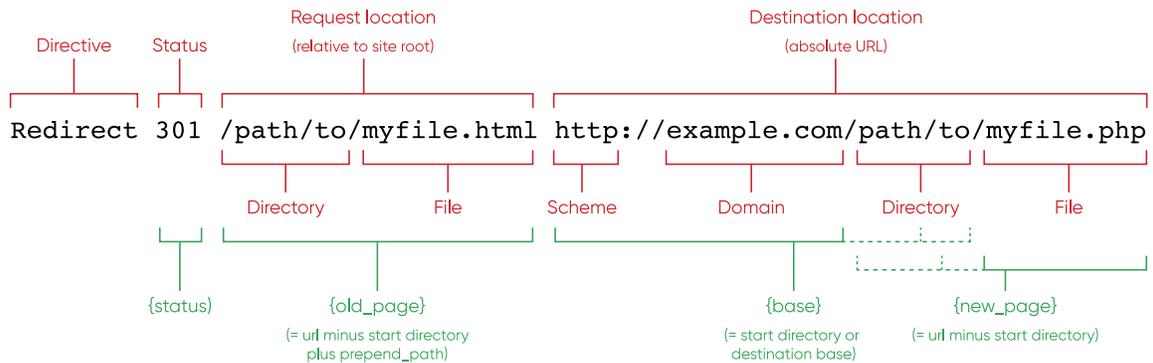
```
{
  "old_page" : "{old_page}",
  "new_page" : "{root}{new_page}",
  "status" : "{status}"
},
```

XML:

```
<entry>
  <old_page>{old_page}</old_page>
  <new_page>{root}{new_page}</new_page>
  <status>{status}</status>
```

</entry>

Terminology



Known Issues

- To do: option to save and manage multiple projects
- To do: Add weight to matches (by probability)
- Bug: search only searches url, should also search titles
- CSV upload is currently very limited. We are looking into allowing more extensive info in a future version.
- Only individual url redirect directives are currently output. We are looking into adding rewrite rules in a future version.