

Advanced 404 Pages

By Lauri Harpf

Advanced 404 Pages

Lauri Harpf
webmaster@apromotionguide.com

A Promotion Guide <http://www.apromotionguide.com/>

Advanced 404 Pages

=====

In a previous article, I discussed the advantages of using a customized 404 error page on your site and gave instructions on how to create one. As I explained back then, these pages are highly useful because they enable you to benefit from traffic that would otherwise be lost. They however do have one dark side that can make maintaining your site significantly more difficult than what it has been before.

What is this problem I'm talking about? Well, it is quite simple. Usually when a server encounters a 404 error, it records the details about the event into its log file. Should you suddenly notice that the log shows multiple 404 errors due to a page named "abutme.html" not being found, you can deduct that you've probably accidentally linked to "abutme.html" somewhere on your pages instead of using the correct filename "aboutme.html". As you can see from the above, the functionality offered by log files makes them a great help in tracking down such simple mistakes and enables you to keep your site relatively free of in-site broken links.

However, if you happen to have replaced the standard 404 error with your own HTML 404 error page, you can't utilize this useful feature. Yes, you will still see that a 404 error has occurred, but you won't get any further details that would help you figure out what exactly happened. Finding and fixing these errors becomes nearly impossible, causing them to pile up over time and deteriorate the professional image of your site. Now you're facing the tough choice of either showing your visitors a very unfriendly error message that drives them away or accepting the fact that your stylish custom 404 error page will become an all too familiar sight to those who click around your site.

All hope is not lost

=====

After reading the above, you must be feeling pretty down. I sure know I did after having installed my own 404 page only to notice that I had corrected one problem, but caused another one while doing it. Still, there is a solution to every difficult situation and this one is no exception. If you want to keep your 404 page and still get informed when you mess things up and create a broken link, you'll be pleased to hear that I happen to have just the thing for the job. Before we begin, please take into account that in order to use this fix, your host has to be running Apache with support for .htaccess files, Server Side Includes (SSI) and CGI. Contact your technical support for information on whether you have access to these valuable features or not.

Without further ado, let's roll up our sleeves and get to work. The first thing you will need is a CGI script that will log the errors and let you know about them. There are several ones out there that you can use, but I personally prefer Matrix Vault's free 404 Helper that can be found at <http://www.pixelwarehouse.com/cgi/404helper.shtml> . Download the source code to your hard drive and open it in a text editor. You can get free editors from the Net, but the old MS-DOS Edit supplied with just about every version of Windows will do just fine. To run Edit, go to Start, Run, type "edit" without the quotes into the box and click OK.

Before you start editing the file, you'll need to know where your host has installed the Perl interpreter and Sendmail. Once you have figured it out, check if the paths used in the CGI script match those your host uses. The Perl interpreter's location is set to /usr/bin/perl in the first row and the location of Sendmail is set to /usr/lib/sendmail in the 21st row. Make changes, if necessary.

After you have made sure that the paths are correct, modify the rest of the script to suit your needs. Be sure to replace the E-mail address in the \$email field with the one you want the error report to be sent to. You might also wish to use a smaller value in the \$mailon field than the default of 10, as it can take quite a while for a small site to generate enough 404 errors to fill up a 10K log. I suggest using a value of 1 or 2 at the beginning and raising it later if you feel that you are receiving the error reports more often than you'd like to.

You're now done with the CGI script. Save it as "404helper.cgi", without the quotes of course. However, there's still work to be done, so take a deep breath and prepare yourself for the next challenge.

Editing your custom 404 page and .htaccess file

=====
Just having the script will not be enough. In order for it to work, it has to be executed when an error is encountered. This is the part where the SSI's step into the picture. Open up your 404 error page in a text editor and add the following line into it:

```
!--#exec cgi="/your_CGI_directory/404helper.cgi"--
```

Because the script prints out a few rows of HTML after it has been run, the best place for that line is at the bottom of your 404 error page, but before the /BODY tag. After everything is safely in place, simply save the file, but instead of ending it with the usual ".htm" or ".html", use ".shtml". Do not forget to do this, as the SSI tag might not work if you fail to use the proper extension.

Finally, you will have to edit the .htaccess file you created when you built your custom 404 page. If you only have " ErrorDocument 404 <http://www.yourdomain.com/404page.html> " in it, modify the file so that it will contain the following:

```
Options Includes ExecCGI
AddType application/x-httpd-cgi .cgi .pl
AddType text/html .shtml
AddHandler server-parsed .shtml
ErrorDocument 404 http://www.yourdomain.com/404page.shtml
```

The new lines will enable Server Side Includes and CGI so that your script will work. Do not forget to change the ErrorDocument 404 line to point to the new .shtml page instead of your old .html version. After you are done, save your .htaccess file.

Upload, set permissions and launch!

=====
Connect to your host with an FTP program and upload the .shtml version of your 404 page and your new .htaccess file into your root directory. Then go to the directory you've reserved for CGI programs and send the 404helper.cgi file there. Make sure that you upload in ASCII, not in Binary mode! I nearly drove myself crazy by accidentally using Binary mode and then trying to figure out why the script refused to work.

Everything is now uploaded and all that remains is to set permissions so that the CGI program can be executed. You'll need to CHMOD 755 both the CGI file and the directory where it is in. The steps you need to take in order to accomplish this depend on what software you are using, but here are the instructions on how to do so with WS_FTP, a popular Windows FTP program which can be downloaded from Tucows.com.

First, navigate to the directory where you've uploaded the CGI program. Left-click it once to highlight the file, then right-click it. Select "chmod (Unix)" from the menu that appears. Give Read, Write and Execute permissions to the Owner and Read and Execute permissions to Group and Other. Then go into the root directory and repeat the same process with the folder of the directory where you placed the CGI program.

That's it. The work is finally done and you can now enjoy the luxury of receiving an E-mail report on all 404 errors, allowing you to quickly stomp out any broken links and improve the quality of your site. Congratulations!

Lauri Harpf runs the A Promotion Guide website, where he offers free information about search engines, directories and other promotion methods. His site can be found at <http://www.apromotionguide.com/>

[Get-Articles.com : 1000's of reprintable business and internet marketing-related articles.](#)

[Submit your article for reprint.](#)