

Re: File Removal after Database Comparison

Source: <http://coding.derkeiler.com/Archive/PHP/alt.php/2008-04/msg00018.html>

- *From:* Preventer of Work <not_this@xxxxxxxxxx>
 - *Date:* Fri, 04 Apr 2008 02:43:24 GMT
-

Chris Owens wrote:

Hi all,

I'm trying to figure the logic (and therefor the code) to enable the clean-up of a site I've inherited. Basically I have a folder that has around 6,666 images in it that are loosely related to the information held in a database. I say loosely because the images are uploaded (via FTP) manually. There is no direct reference to the JPEGs in the database, as the code on the site references them via:

```
  
  

```

Each record in the database has three related images:

Image of the front – recordid_f.jpg
Image of the side – recordid_s.jpg
Image of the back – recordid_b.jpg

If the image file names really contain the database id like

2177_s.jpg

2177_b.jpg

.... etc

It would be a matter of (think unix but similar in windows shell)

```
ls -1 *_s.jpg > fnames.txt
```

to capture all file names.

After that, scratch some code code together to collect the numbers from fnames.txt entry (easy in any language).

Create new folder nearby.

write script that does this:

open database

for each number in list

.. query database for the number

.. if in the db:

.. stringify the number to recreate filenames

Re: File Removal after Database Comparison

.. copy files to new dir

When this is done, ONLY the images you want to keep will be in the new folder. Once you are sure it worked, delete all originals. This is the safe way.

Faster (but test carefully first):

open database

for each number in list

.. query database for the number

.. if NOT in the db:

.. stringify the number to recreate filenames

.. delete files

Now over time records have been removed from the database, but the images haven't been removed and as such, short of removing all images, there is no easy way that I can think of to clear the folder (several GB of data) whilst keeping the images that have related records in the db.

My question is: How can I through code remove any orphaned images (that aren't needed anymore), whilst keeping the images that still relate to records in the database?

Any assistance/pointers/snippets would be much appreciated.

Kind regards,

Chris