

File indexing and full-text searching

Web File Share configuration

You can enable the full-text file searching from “Control Panel” » “System configuration” » “Files” » “Indexing”. It is not enabled by default because it requires third-party software (Apache Tika).

If you are running Tika in command line mode, simply provide the path to the “tika-app-1.12.jar” file. That's it! Click the “Check path” to make sure it works. If Java is installed on the server and the file path is correct, you should see the Apache Tika version displayed as a result of the test.

If you are running Tika in server mode, provide the hostname and port number of the Tika server. Click the “Test server” to make sure it works. If everything is in order you should see the Apache Tika version displayed as a result of the test.

Installing Apache Tika

Installing Tika is as simple as uploading a file to your server.

Download the “jar” file from here: <https://tika.apache.org/download.html>

Download the “tika-app[*].jar” if you want to run Tika from the command line and “tika-server[*].jar” if you want to run it as a server. Running Tika in server mode speeds up the indexing process.

You can read more about Tika here: <https://tika.apache.org>

Set Index Queue Manager

As extracting the text from a binary file requires a lot of CPU processing, the files are queued and processed one at a time. This requires the script “cron/process_search_index_queue.php” to be executed frequently. We recommend running the script every 5 minutes or so, so you will not have to wait too long until an uploaded file will be found by the search engine.

On a Linux server this can easily be done by setting up a cron job like this:

1. Create a new text file at “cron/process_search_index_queue.sh” and write the following inside:

```
php -c php.ini process_search_index_queue.php
```

To find out the path of the “php.ini” used by Web File Share open <http://your-site.com/WebFileShare/info.php> file in your browser.

2. Open a command line console (SSH)
3. Open the crontab editor by running:

```
crontab -e
```

4. Write:

```
***** /path-to-WebFileShare/cron/process_search_index_queue.sh
```

5. Press “:wq” and “Enter” to save the changes and close the editor.

If your hosting service is running the cPanel administrative tool, it usually provides a web-based tool for setting up cron jobs easier.

On Windows this can be achieved by creating a Windows schedule event which calls a .BAT file containing something like this:

```
CD cron
C:/PHP/PHP.EXE process_search_index_queue.php
```

🔄Revision #1

★Created 22 December 2021 00:44:02 by Mahesha Damayanthi

♣Updated 24 December 2021 17:45:57 by Mahesha Damayanthi