

Helabasa ReadMe v1.0.1

Required Software and libraries (Please install these software and libraries otherwise Helabasa won't work)

Python 3.6.8+

Pyfasttext installation guide: <https://pypi.org/project/pyfasttext/>

Foma 0.9.18 alpha version installation guide: <https://fomafst.github.io/>

Morphy directory contains the required files to reinstall the Helabasa application. So copy only the **Morphy** folder to your Web Server first.

test	File folder	12/3/2019 12:11:34 PM	drwxrwxr-x	ftpnginx ft...
mofessor	File folder	11/25/2019 4:56:52 PM	drwxrwxr-x	ftpnginx ft...
rootextr	File folder	11/5/2019 11:35:47 AM	drwxrwxrwx	ftpnginx ft...
morphy	File folder	10/16/2019 12:12:47 A...	drwxrwxrwx	basa 1001
sinling	File folder	9/2/2019 12:28:42 PM	drwxrwxrwx	ftpnginx ft...

But some very important files are missing due to GitHub upload limitation. Therefore you have to download those important files using the following download links and upload back to your server. Also you have to modify some file paths in python codes and PHP codes. Most of the wrong file paths can be fixed by changing **/var/www/html/** to your web server public path. Please go through all the .php and .py file to fix file path issues. (We used **Nginx** as our web server there for file path is quite different than **apache** web server)

Example: **morphy/fasttext/service.py**

```
def hashing_spell checking dic():
    filepath = "/var/www/html/morphy/fasttext/word_dms_hash"
    a = {}
    with open(filepath) as fp:
        for line in fp:
            x = line.strip().split(' ')
            a[x[0]] = int(x[1])
    return a
```

Example: **morphy/foma/dicexecute.php**

```
$words = $_POST["words"];
$data = json_decode($words,true);
$data_string = json_encode($data);
$arrayis =array();
$data_string_arr = rawurldecode($data_string);
$data_string_arr =str_replace ("'", "&quot;", $data_string_arr);

$command = "/usr/bin/python3 /var/www/html/morphy/foma/client.py ".$data_string_arr."";
```

Note : Here you have to change python directory path based on you OS's python location.

Missing Files (Download and Upload to you web server's correct directory)

Directory: **morphy/fasttext**

1. Missing file **sinhala_all2.bin** (3.24GB)

Download link: http://helabasa.projects.uom.lk/morphy/fasttext/sinhala_all2.bin

2. Missing file **hashed_dic** (157.94MB)

Download link: http://helabasa.projects.uom.lk/morphy/fasttext/hashed_dic

3. Missing file **unique_word_frequency** (62.3MB)

Download link: http://helabasa.projects.uom.lk/morphy/fasttext/unique_word_frequency

(Note: If you can't download these file the use UOM VPN access or Download using UOM wireless network)

After completing above steps, finally you have to run **two different Linux background services**.
Run following python file as background services. (There are lots of tutorials available online)

Files to be run as a service

Service 1: **/morphy/foma/word_detector.py**

Service 2: **/morphy/fasttext/service.py**