Name

degrotesque — A tiny web type setter.

Synopsis

```
degrotesque [( --input | -i ) <PATH>][( --recursive | -r )][( --no-backup | -
B ) ][( --unicode | -u )][( --extensions | -e ) <EXTENSION>[,<EXTENSION>]*][( --
encoding | -E ) <ENCODING>][( --skip | -s ) <ELEMENT>[,<ELEMENT>]*][( --actions
| -a ) <ACTION>[,<ACTION>]*]
```

```
degrotesque [--help]
```

DESCRIPTION

degrotesque is a small python script that replaces plain ASCII characters by their typographic counterparts.

OPTIONS

--input *\<PATH>*, --i *\<PATH>*	Defines the files to read; paths are accepted as well. *Mandatory, type:file- or pathname, default: none*
--recursive, --r	Whether a given path shall be processed recursively. *Optional, default: false*
--no-backup, -B	Whether no backup shall be generated. *Optional (backup files are generated if not given), default: false*
--unicode, -u	If set, unicode characters will be used in replacement instead of HTML entitites. *Optional, default: false*
--extensions *\<EXTENSION>[,\<EXTENSION>]** -e *\<EXTENSION>[,\<EXTENSION>]**	Defines the extensions of files to process. *Optional, default: html, htm, xhtml, php, phtml, phtm, php2, php3, php4, php5, asp, jsp, jspx, shtml, shtm, sht, stm, vbhtml, ppthtml, ssp, jhtml*
--encoding *\<ENCODING>*, -E *\<ENCODING>*	Defines the document encoding to use when reading. *Optional, default: utf-8*
--skip *\<ELEMENT>[,\<ELEMENT>]**, -s *\<ELEMENT>[,\<ELEMENT>]**	Defines the elements to skipt. *Optional, default: script, code, style, pre, ?, ?php, %, %=, %@, %--, %!, !--, !DOCTYPE*
--actions *\<ACTION>[,\<ACTION>]**, --a *\<ACTION>[,\<ACTION>]**	Defines the actions to perform. *Optional, default: masks, quotes.english, dashes, ellipsis, math, apostrophe*
--help	Prints the help screen. *Optional, default: false*

USAGE

degrotesque is implemented in [Python](). It is started on the command line.

The option `-i <PATH>` / `--input <PATH>` tells the script which file(s) shall be read - you may name a file or a folder, here. If the option `-r` / `--recursive` is set, the given folder will be processed recursively.

The tool processes HTML files, XML files, and their derivatives. You may change the extensions of files to process using the `-e <EXTENSION>[,<EXTENSION>]*` / `--extensions <EXTENSION>[,<EXTENSION>]*` option.

The files are read one by one and the replacement of plain ASCII chars by some nicer ones is based upon a chosen set of “actions”. You may select the actions to apply using the `-a <ACTION_NAME>[,<ACTION_NAME>]*` / `--actions <ACTION_NAME>[,<ACTION_NAME>]*` option. The default actions are masks, quotes.english, dashes, ellipsis, math, and apostrophe. Per default, HTML entities are inserted. If you rather wish to have unicode values, use the option `-u` / `--unicode`.

The files are assumed to be encoded using UTF-8 per default. You may change the encoding using the option `-E <ENCODING>` / `--encoding <ENCODING>`.

The script does not change the quotation marks of HTML elements, of course. As well, the contents of several elements, such as <code> or <pre>, are skipped. You may change the list of elements which contents shall not be processed using the option `-s <ELEMENT_NAME>[,<ELEMENT_NAME>]*` / `--skip <ELEMENT_NAME>[,<ELEMENT_NAME>]*`.

After the actions have been applied to its contents, the file is saved. By default, a backup of the original file is saved under the same name, with the appendix ".orig". You may omit the creation of these backup files using the option `-B` / `--no-backup`.

Please note that "masks" is a special action set that disallows the application of some other actions so that, e.g., the dividers in ISBN numbers are not replaced by –.

Known Actions

The actions that can be applied are:

- quotes.english

- quotes.french

- quotes.german

- to_quotes

- commercial

- dashes

- bullets

- ellipsis

- apostrophe

- math

- dagger

Please consult the online documentation (http://www.krajzewicz.de/blog/degrotesque.php or https://github.com/dkrajzew/degrotesque) for further information about the actions.

EXAMPLES

degrotesque -i my_page.html -a quotes.german

Replaces single and double quotes within the file "my_page.html" by their typographic German counterparts.

degrotesque -i my_folder -r --no-backup

Applies the default actions to all files that match the extension in the folder "my_folder" and all subfolders. No backup files are generated.

CAVEATS

You may combine php- and html-code in a way that the parser replaces wrong elements - especially quotes. The tool also fails if you embed HTML-elements within other HTML-elements.

You should thereby verify the results.

DIAGNOSTICS

degrotesque reports errors if something fails.

BUGS

Nothing is known besides the named caveats, yet. Please report bugs to daniel@krajzewicz.de.

SEE ALSO

Internet pages at http://www.krajzewicz.de/blog/degrotesque.php

git project at https://github.com/dkrajzew/degrotesque

COPYRIGHT

(c) 2020-2022 Daniel Krajzewicz

http://www.krajzewicz.de/blog/degrotesque.php