BH_Rebrand.py Manual

Rolf-Dieter Fuerste 2021-12-07

1 Introduction

1.1 Purpose

The purpose of BH_Rebrand.py is automated, “bulk” rebranding of MS Office documents (Word, Excel PowerPoint) from the former “BHGE” branding to the current Baker Hughes Company style and re-create the respective PDF files.

This applies to replacing the company logo in the document header, company name, document font, brand colors.

The replacement details are controlled by a configuration file.

1.2 Pre-requisites

The MS Office documents to be rebranded must be in the “Office Open XML (OOXML) format, i.e. docx, docm, xlsx, xslm, pptx, pptm

The files to be rebranded need to be stored in a single folder in the file-system.

Several folders may be created to perform rebranding of different sets of file and file-types with respective configuration files.

1.3 Out of scope

- Replacement of the company logo, font, colors in embedded in images, e.g. screenshots
- Export & refile of the affected files from & to a document management system.

1.4 Acknowledgement

Special thanks to Ole Gerlof who wrote the code for BH_Rebrand.py.

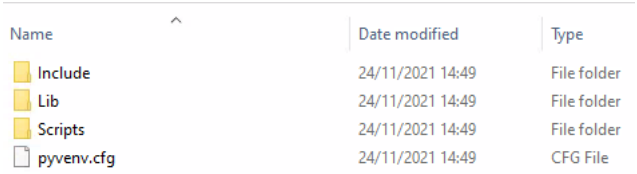
2 Setup

2.1 Python runtime interpreter

Unless already installed, request the package Python 3.9.1 from Baker Hughes TechHub.
https://techhub.ent.bhicorp.com/esd/Items/Details?PackageId=198

2.2 Python runtime virtual environment

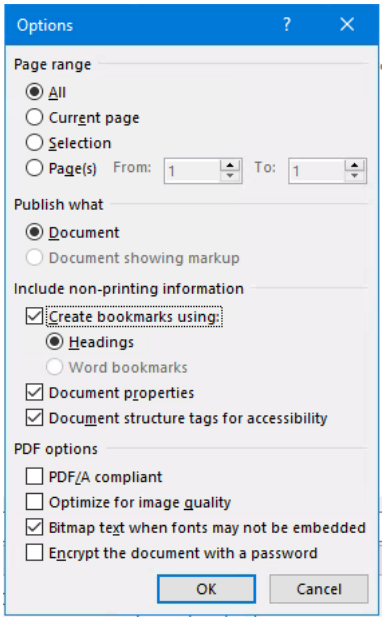
1. Open command prompt window
2. Execute command python –m venv [folder path]
   E.g.: `python –m venv d:\rebrand`
   This will create the folder “rebrand” and add required library files



3. Change directory to folder
   E.g.: `cd d:\rebrand`
4. Execute activation script
   `Scripts\activate.bat`
5. Copy files from BH_Rebrand.zip to [folder path]
6. Start installation of required modules
   `python –m pip install –r requirements.txt`
   This will download the required modules and add them to the library.

2.3 Word configuration

In File- > Export -> Create PDF/XPS -> Options:
Ensure that the option “Create bookmarks using headings is enabled.



3 Configuration file

3.1 InputFolder =

Path to folder with original files.

3.2 OutputFolder =

Path to folder with updated files.

3.3 LogFolder =

Path to folder with log-files per run

3.4 PDF =

Enable/disable automated conversion to PDF.

- `PDF=True`
  Will create a PDF file
- `PDF=False`
  Will not create a PDF file

Note that conversion to PDF may take several seconds per document.
Therefore, for testing the configuration file, use PDF = False at first.

3.5 ReplaceHeaderImage=

Controls replacement of the logo in the header or the entire document.

- ReplaceHeaderImage=True
  Searches **only** the header and replaces all images with the new logo, independent of the filename and size.
  This setting is suitable for input files which have only the former logo in the header.
- ReplaceHeaderImage=False
  Searches the **entire** document and replaces the image acc. to OldLogoPath =", see 3.7 or images with the size "OldLogoSize =". See 3.8.
  This setting may be suitable for input files which do *not* have the former logo in the header.

3.6 NewLogoPath =

Filepath & name of the new logo to use.

Example:

```
NewLogoPath = d:\Rebrand\replacementLogo.png
```

Notes:

- In order ensure that the aspect ratio of the logo in the document is retained, make sure that the "NewLogo" image has the same aspect ratio with respect to width/height in pixels like the "OldLogo".
- The replacementLogo.png that comes with BH_Rebrand.py is adapted the aspect ratio (width/height) of the "OldLogo" in the BHOS template:
  - Width:1541 pixel Height: 727pixel

In other documents, which are based on a different template, the “OldLogo” may have a different width/height and resulting aspect ratio.

In this case it may be required to create a specific, replacement logo, with white padding added at the left/right or top/bottom so that it matches the aspect ratio with respect to width/height in pixels of the “OldLogo”.

3.7 OldLogoPath =

Filepath & filename of the logo in the XML-structure, which shall be replaced.

{filetype} will be automatically replaced with the foldername acc. to the file type (Word, Excel, Powerpoint.

Note: This setting will only be observed if ReplaceHeaderImage=False. See 3.4.

The logo’s filename depends on the template that was used to create the MS Office file.

The path and size inside the MS Office file can be obtained by “unzipping” the file, e.g. by opening with “7-Zip”.

In many cases the logo will be in the path:

```
OldLogoPath = {filetype}/media/image1.png
```

If the logo with a matching path is found, it will only be replaced, if its size in bytes matches the value in “OldLogoSize =”.

Else, there will be note on the log-file, e.g.:

- “Found alternative image in word/media/image1.png with size 8362 bytes”

3.8 OldLogoSize =

Filesize in bytes of the old logo in the XML-structure

Note: This setting will only be observed if ReplaceHeaderImage=False. See 3.4

The logo’s filesize depends on the template that was used to create the MS Office files
E.g: For Word documents based on the former BHOS “BHGE” Template:

```
OldLogoSize = 82068
```

If a logo with the matching size is found, it will be replaced, independent of its file-name or path.

Note: This may result in undesired replacements, if there should be other images than the logo in the current document with the same size in bytes like “OldLogoSize”.

This option should only be use on a set of files where it is known that they were derived from the same template, so that the filesize of the logo is known.

3.9 String replacements

Strings in the XML structure, which shall be replaced.
Usage: `Old string -> New String`

The replacement supports the usage of Regular Expressions.
E.g.:

```
B(<[\S\s]*?>)?H(<[\S\s]*?>)?G(<[\S\s]*?>)?E -> Baker Hughes
```

This expression replaces all occurrences of *BHGE*.

It turned out in the XML structure that there may be tags between the letters, so that a simple replacement of “BHGE” will not capture all occurrences.

3.9.1 Company name

Use distinct replacements for occurrences of the company name.

E.g.:

```
BAKER HUGHES, A GE COMPANY; LLC -> BAKER HUGHES COMPANY
Baker Hughes, a GE Company; LLC -> Baker Hughes Company
Baker Hughes, a GE company, LLC -> Baker Hughes Company
Baker Hughes, a GE company -> Baker Hughes Company
```

Note that in the XML structure there may also be tags inside the company name, so that a simple replacement of the plain string will not capture all occurrences.

3.9.2 Font substitution

The GE font *GE Inspira Sans* shall be replaced with the *Arial* font.

```
="GE Inspira Sans -> ="Arial
```

3.9.3 Color replacement

Color values are stored as hexadecimal.values in XML.

The current values in the document can be obtained from the formatting menu, e.g. in Word (Color -> more colors -> custom)

- BHOS Header Gradient line left blue to green
  `#00B5E2 -> #018374`
  `="00B5E2 -> ="018374`

- BHOS Header Gradient line center blue to green, Heading color, header text color
  `#005EB8 -> #019F84`
  `="005EB8 -> ="018374`

- BHOS Header Gradient line right blue to green
  `#10069F -> #02BC94`
  `="10069F -> ="02BC94`

- dark blue to green
  `="333399 -> ="018374`

4 Execution

4.1 Run bh_rebrand.py

7. Open command prompt window
8. Execute activation script once before starting bh_rebrand.py
   `Scripts\activate.bat`
   This is only required once within the current command prompt window
9. Execute command bh_rebrand.py [configuration file]
   E.g.: `d:\rebrand\bh_rebrand.py d:\rebrand\default.cfg`

4.2 Check log-file

A log file is created per each run as a CSV-file in the folder “LogFolder =”.
Name schema: BH_Rebrand_YYYY-MM-DD_hhmmss.csv

- E.g. `BH_Rebrand_2021-11-23_164244.csv`

The log file contains the columns: Inputfile, Outputfile, Status, Notes

Check the columns status and notes for information regarding the execution.