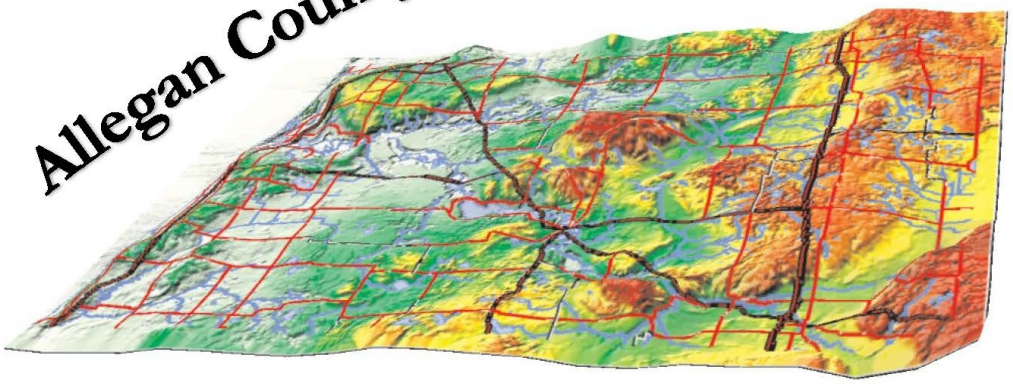
Allegan County GIS Services



Batch PDF Optimization

WWW.ALLEGANCOUNTY.ORG/GIS

JUNE 21, 2019

Contents

0.0.1 PDF Optimizer

Purpose and Summary

Purpose

Optimization of any number of pdf documents

Summary

A Python script creates a list of .pdf docs in a folder. The list is used to write a .txt document in which every line is a DOS command to optimize each of the .pdf documents and save them to another location. The .txt must be saved as a .bat. When executed the batch process calls ghost script for the optimization.

Requirements

Software

- ghostscript
- python 2.7 and a Python IDE
- A text editor

About ghostscript

ghostscript is used for the optimization. ghostscript is an interpreter for the PostScript language and for PDF [**?**].

Licensing

ghostscript is available opensource under AGPL conditions. more information can be found here.

Download

ghostscript can be downloloaded here.

note:

The output of this script is bdoc.txt, Save as a .bat to execute the optimize.

Python(2.7)

This script was developed in python 2.7

The Python Script

The output is a batch file

```
#-------------------------------------------------------------------------------
# Name:        Optimize
#
# Purpose:      Batch optimize pdfs
#
# Notes:        This script creates a list of pdf files in a source folder and
#                   then creates a .txt that can be used as a .bat file to optimize
#                   all of the pdfs in the source folder to a new location.
# Author:      BMay
#
# Created:     06/20/2019
#-------------------------------------------------------------------------------
################################################################################
#   Imports and Relative path folder setup
################################################################################
import os, sys
project = os.path.dirname(os.path.dirname(__file__))
processing = os.path.join(project, 'processing')
build = os.path.join(project,'build')

################################################################################
#   String vars for each line of the .bat file
################################################################################
inString1 = "gswin32 -sDEVICE=pdfwrite -dCompatibilityLevel=1.4 -dPDFSETTINGS=/ebook#
 -dNOPAUSE -dQUIET -dBATCH -sOutputFile=H:\\2019ParcelAtlas\\optimized\\"
inString2 = " H:\\2019ParcelAtlas\\20190619\\"
usString = '_' # Underscore string to add to file names

################################################################################
#   Source pdfs path
################################################################################
sourcepdf = os.path.join(project, '20190619x') # folder with pdfs to be optimized
```

```
#############################################################################
#   new .txt
#############################################################################
batchdoc = os.path.join(processing,"bDoc.txt") # new .txt that can be used as a .bat

#############################################################################
# Main
#############################################################################
if __name__ == "__main__":
    list1 = os.listdir(sourcepdf) # assemble list of all files in sourcepdf
    l = open(batchdoc,'w') # open .txt doc to write lines
    for i in list1: # iterate list of files
        #newi = i[0:] # allows slicing on file name if chars need to be removed
        #print newi
        #t = inString1 + usString + newi + inString2 + i + "\n"
        t = inString1 + usString + i + inString2 + i + "\n" # assemble each string
        print t
        l.write(t) # write each string
    l.close()
```

WINDOWS BATCH FILE

A line from the batch file looks like:

```
gswin32 -sDEVICE=pdfwrite -dCompatibilityLevel=1.4 -dPDFSETTINGS=/ebook -dNOPAUSE
        -dQUIET -dBATCH -sOutputFile=J:\Project\2018ParcelAtlas\build\optimized 02-
        001-001-00.pdf J:\Projects\2018ParcelAtlas\build\2018071 \_02-001-001-00.pdf
```

To execute the batch file: change the extension of the scripts output from .txt to .bat. Double click the .bat to execute.