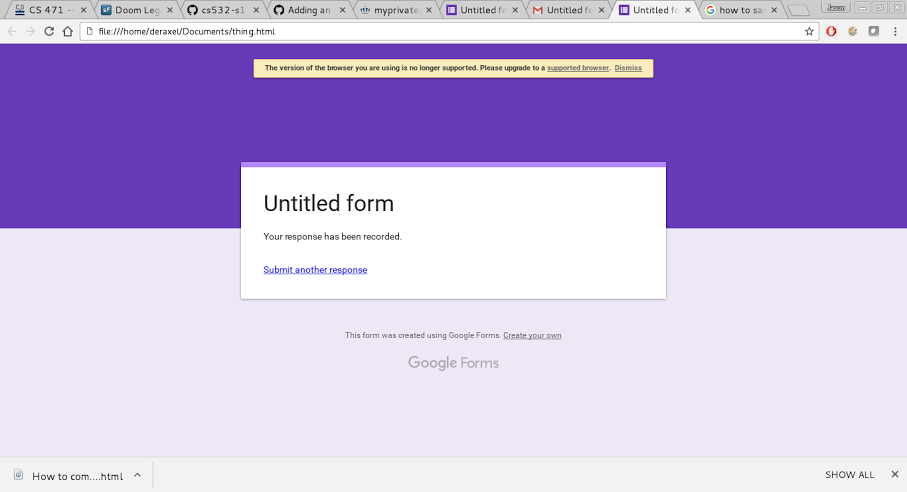
Jason Howse

CS 432

Thursday 5.45 pm – 8:25 pm

Jhows001@odu.edu

1. curl -d "entry.1873578523=some of the things&submit=Submit" "<https://docs.google.com/forms/d/e/1FAIpQLSe7ckfuSj5gc7tfEkYkSZnwMRML9IZ3iideULeu-B9I8E6oXg/formResponse>"



2. see dotsave.py

import urllib2

import re

def getlinks(website):

openWebsite = urllib2.urlopen(website)

domain=re.findall('((http|ftp).+[.].?..[/])',website)

domainName=domain[0][0]

domainName=domainName[:-1]

html=openWebsite.read()

allPdflinks=re.findall('="[^"]+[.]pdf"',html)

alllinks=re.findall('"((http|ftp)s?://.\*?)"',html)

i=0

i2=0

i2=len(allPdflinks)

while (i<i2):

allPdflinks[i]=allPdflinks[i][2:]

allPdflinks[i]=allPdflinks[i][:-1]

if(not(allPdflinks[i].startswith('http') or allPdflinks[i].startswith('ftp'))):

allPdflinks[i]=domainName+allPdflinks[i]

i+=1

i=0

while (i<i2):

try:

print allPdflinks[i]+" is "+urllib2.urlopen(allPdflinks[i]).headers["content-length"]+" bytes"

except:

i=i+0

i+=1

=============================== RESTART: Shell ===============================

>>> import dotsave

>>> dotsave.getlinks("https://www.odu.edu/about/visitors/campus-map")

https://www.odu.edu/content/dam/odu/images/maps/campus/campus-map.pdf is 38809925 bytes

>>> dotsave.getlinks("http://www.cs.odu.edu/~mln/teaching/cs532-s17/test/pdfs.html")

http://www.cs.odu.edu/~mln/pubs/ht-2015/hypertext-2015-temporal-violations.pdf is 2184076 bytes

http://www.cs.odu.edu/~mln/pubs/tpdl-2015/tpdl-2015-annotations.pdf is 622981 bytes

http://www.cs.odu.edu/~mln/pubs/tpdl-2015/tpdl-2015-off-topic.pdf is 4308768 bytes

http://www.cs.odu.edu/~mln/pubs/tpdl-2015/tpdl-2015-stories.pdf is 1274604 bytes

http://www.cs.odu.edu/~mln/pubs/tpdl-2015/tpdl-2015-profiling.pdf is 639001 bytes

http://www.cs.odu.edu/~mln/pubs/jcdl-2014/jcdl-2014-brunelle-damage.pdf is 2205546 bytes

http://www.cs.odu.edu/~mln/pubs/jcdl-2015/jcdl-2015-mink.pdf is 1254605 bytes

http://www.cs.odu.edu/~mln/pubs/jcdl-2015/jcdl-2015-arabic-sites.pdf is 709420 bytes

http://www.cs.odu.edu/~mln/pubs/jcdl-2015/jcdl-2015-dictionary.pdf is 2350603 bytes

>>> dotsave.getlinks("https://www.odu.edu/academics/academic-records/transcripts")

https://www.odu.edu/content/dam/odu/offices/university-registrar1/docs/transcript-request-form.pdf is 97729 bytes

https://www.odu.edu/content/dam/odu/offices/university-registrar1/docs/transcript-request-form.pdf is 97729 bytes

https://www.odu.edu/content/dam/odu/offices/university-registrar1/docs/transcript-request-form.pdf is 97729 bytes

3.

IN: DH

SCC: ABCG

OUT: MO

Tendrils: IKLP

Tubes: N

Disconnected: EFJ