

Dateien mit Python herunterladen

Um mit Python Dateien aus dem Web herunterzuladen wird das:

Requests

Modul benötigt.

Requests muss manuell installiert werden:

```
Py -m pip install requests
```

```
import requests
```

```
#Lade die komplette Datei in den Zwischenspeicher:  
res = requests.get('http://christopher-schwerdt.de/sherlock.txt')
```

```
#Zeige erste 256 Zeichen von res:  
print(res.text[:256])
```

Datei aus dem Web abspeichern

```
import requests
```

```
#Lade die komplette Datei in den Zwischenspeicher:  
res = requests.get('http://christopher-schwerdt.de/sherlock.txt')  
#Prüfe, ob alles okay. Ansonsten werfe Exeption  
res.raise_for_status()
```

```
#Öffne neue Datei im Filesystem als "Binary Writer"  
playFile = open('Sherlock.txt', 'wb')  
# Schreibe Daten vom res Objekt in die Datei:  
# iteriere durch die Daten vom res Objekt(Einhunderttausend bytes je chunk).  
for chunk in res.iter_content(100000):  
    playFile.write(chunk)
```

```
#Dannach schließe den Writer  
playFile.close()
```

Zusammenfassung

Um eine Datei aus dem Web mittels Python herunterzuladen, muss folgendermaßen vorgegangen werden:

1. `Request.get()` mit der herunterzuladenden Datei aufrufen.
2. `Open()` mit "wb" erstellt eine neue, leere Datei im "write binary" Modus.
3. Mittels einer Schleife durch das res Objekt iterieren mittels der `iter_content()` Methode.
4. Die `Write()` Methode für jeden "chunk" (jede Iteration) aufrufen, um den Inhalt in die Datei zu schreiben.
5. Am Ende `close()` aufrufen, um die Datei sauber (ab)zuschließen.

Abschlussaufgabe

Schreibe ein Programm (in Python), welches automatisiert Dateien aus dem Internet herunterlädt.

Hinweise und Anforderungen:

1. Es muss per Kommandozeile aufgerufen werden können.
2. Der erste Parameter entspricht dem Ziel Dateiformat (z.B. jpg, txt, ...)
3. Der zweite Parameter entspricht der URL.

Beispielaufruf:

```
py .\3_Downloader.py jpg https://fleurop.cstatic.io/media/c3/8d/cd/1679328524/fleurop-  
blog-blumenarten.jpg
```



Download.jpg

Lösung:

```
import requests, sys

if len(sys.argv) > 1:
    format = sys.argv[1]
    pathToContent= sys.argv[2]

#Lade die komplette Datei in den Zwischenspeicher:
res = requests.get(pathToContent)
#Prüfe, ob alles okay. Ansonsten werfe Exeption
res.raise_for_status()

#Öffne neue Datei im Filesystem als "Binary Writer"
playFile = open('Download.' +format, 'wb')
# Schreibe Daten vom res Objekt in die Datei:
# iteriere durch die Daten vom res Objekt(Einhunderttausend bytes je chunk).
for chunk in res.iter_content(100000):
    playFile.write(chunk)

#Dannach schließe den Writer
playFile.close()
```