

# 22 Python基礎

## os / shutil の利用

### osモジュール

- 以前、紹介している
  - 再度、どのようなメソッドが用意されていて、どのような処理が行えるのか確認してください。

**主な役割:** OSに依存する機能を提供し、ファイルやディレクトリの操作を行う

主なメソッド	説明
os.walk()	ディレクトリツリーを走査し、ファイルやディレクトリの一覧を取得する
os.system()	シェルコマンドを実行する
os.listdir()	指定したディレクトリ内のファイルとディレクトリの一覧を取得する
os.path.join()	パス名を結合する
os.environ	環境変数を取得・設定する
os.mkdir()	新しいディレクトリを作成する
os.rename()	ファイルやディレクトリの名前を変更する
os.remove()	ファイルを削除する

- 一部抜粋

### shutilモジュール

**主な役割:** 高水準のファイル操作を提供し、ファイルのコピーや削除などの複雑な処理を行う

主なメソッド	説明
<code>shutil.copyfile()</code>	ファイルの内容をコピーする
<code>shutil.copy()</code>	ファイルとパーミッションをコピーする
<code>shutil.copy2()</code>	ファイル、パーミッション、その他のメタデータをコピーする
<code>shutil.copytree()</code>	ディレクトリ階層全体をコピーする
<code>shutil.rmtree()</code>	ディレクトリ階層全体を削除する
<code>shutil.move()</code>	ファイルまたはディレクトリを移動（名前変更）する
<code>shutil.make_archive()</code>	圧縮アーカイブファイルを作成する
<code>shutil.disk_usage()</code>	ディスクの使用状況を取得する
<code>shutil.which()</code>	実行可能ファイルのパスを検索する
<code>shutil.unpack_archive()</code>	圧縮ファイルをアンパックする

- 一部抜粋

# Officeのファイルの仕組みを知ろう

---

- Microsoft Officeで作成したファイルはどのような構造になっているのだろうか？

- Wordを使用して、文書内に画像を2点貼り付けて保存してください。
  - 文章を書き込んでもOKです。
- 拡張子 `.docx` を `.zip` に強制的に書き換えてください。
- zipファイルとして展開してみます。  
以下の例では、`test1.docx`を元に作業したものです。

```
./test1
├── [Content_Types].xml
├── _rels
├── docProps
│   ├── app.xml
│   ├── core.xml
│   └── custom.xml
└── word
    ├── _rels
    │   └── document.xml.rels
    ├── document.xml
    ├── fontTable.xml
    ├── media
    │   ├── image1.jpeg
    │   └── image2.png
    ├── settings.xml
    ├── styles.xml
    ├── theme
    │   └── theme1.xml
```

- 貼り付けた画像は、`word/media` 内に保存されていることが分かります。

## 練習

---

- 以下のプログラムを作成してください。(ren1\_1.py)
  - 先の手作業で行ったものをコードに落とし込みます
    - 仮にfilename = 'test1.docx' とする
    - filename（拡張子 `docx`）をzipfile（拡張子 `zip`）にリネームする
    - zipfileを解凍する
    - 解凍したフォルダ内の `word/media` 内の画像ファイルを、別のフォルダに移動する

- zipfileを元のファイル名に戻す
- 展開したフォルダを削除する
- 例では、test1.docx→test1.docx\_mediaフォルダを生成した

```
./  
├─ test1.docx  
└─ test1.docx_media  
    ├─ image1.jpeg  
    └─ image2.png
```

- Excel , PowerPointも同様に、どのような構成になるかを確認してください
- 先のプログラムを改良してください。(ren1\_2.py)
  - Excelファイル (.xlsx)、PowerPoint (.pptx) も、同様に処理できるようにしてください
    - ファイル名の拡張子を用いて動作を切り分ける

```
./test2.xlsx_media/  
├─ image1.png  
└─ image2.jpeg
```

```
./test3.pptx_media/  
├─ image1.png  
├─ image2.jpeg  
└─ image3.jpeg
```

- コマンドライン引数を使用して、ファイル名を受け取り、処理できるようにしてください (ren1\_3.py)
  - 複数のファイルを指定できるようにする
  - ワイルドカードは適切に利用できるようにする
  - すでに、同名のフォルダ (～\_media) が存在する場合は、そのファイルの処理をSKIPする
  - 各種エラー処理を組み込む
    - どのようなエラーが発生しうるか検討すること
  - 引数にファイルが指定されない場合は、使い方を表示する
  - 極力関数化し、1つのMainプログラムだけで処理しないようにする
  - コメントは適切に設定する
- 適切に管理できるクラスを作成し、それを用いたプログラムを作成してください。(ren1\_4.py)

## 課題

---

- 上記、ren1\_4.pyを提出する