

Debian で始める Python Programming

taiseiyo

Mail Address: taiseiyo11@gmail.com

2022 年 2 月 19 日

Python とは？

Python(パイソン) はオブジェクト指向のプログラミング言語であり、手軽に始める事ができ実用的な本格的なプログラムの開発を行う事が可能

言語ランキング (TIOBE Software 調べ) : 1 位

<https://www.tiobe.com/tiobe-index/>

最新の安定バージョン : 3.10.2

本日のソース一式

```
git clone https://github.com/taiseiyo/pman.git
# debian branch の doc 以下に資料等は置いてます
```

Python 導入：apt install を使う (Debian)

メリット

- ▶ インストールが簡単

```
sudo apt install python3 python3-pip
```

デメリット

- ▶ バージョンが古い
 - ▶ バージョン 3.10 以降は文法チェッカーが優秀 (<https://www.python.org/dev/peps/pep-0626/>)
 - ▶ buster(debian 10.x 系) → python 3.7
 - ▶ bullseye(debian 11.x 系) → python 3.9

```
cat /etc/os-release or cat /etc/debian_version
```

Python 導入：source からビルド

メリット

- ▶ 最新バージョンが手に入る

デメリット

- ▶ 手間がかかる (シェルでまとめておく と 便利)

```
wget https://www.python.org/ftp/python/3.10.2/Python-3.10.2.tar.xz
tar Jxvf Python-3.10.2.tar.xz
# debian なら tar よりunar がオススメ
cd Python-3.10.2
./configure
sudo make # ディレクトリの python が使用可能になる
# PATH が通っている所の Python を置き換える場合
# sudo make altinstall
```

pman コマンド

発想

- ▶ Debian なら man コマンドがあるが、同じ事を Python でも行いたい
- ▶ いつでもモジュールのマニュアルが欲しい

手法

1. Python 公式サイトから標準ライブラリのデータ (HTML) を取得
2. HTML から文字列を取得

実践：必要なモジュール

- ▶ beautifulsoup4
 - ▶ HTML ファイルや XML ファイルからデータを抽出するための Python ライブラリ
- ▶ lxml
 - ▶ <https://pypi.org/project/lxml/>
 - ▶ Python で xml や html を扱うためのライブラリ。Web サイトの html を解析して情報を抽出するプログラムを作成する場合などに多く用いる

pman 作成の環境構築 (/pman/prog/ の想定)

```
python -m venv test
cd test
source bin/activate
pip install -r ../requirements.txt
# 作業終了後 deactivate
```

実践：プログラム作成

HTML 取得

- ▶ `urllib.request.urlopen()`

HTML 解析

- ▶ `BeautifulSoup()`

最後に

GitHub: