

ХОЧУ ПОМОЧЬ

Модуль

# Модуль python-magic в Python, идентификация типа файла



Распродажа на Авито.

avito.ru РЕКЛАМА - 16+

Больше информации на сайте рекламодателя

Подробнее

[Сторонние пакеты и модули Python3.](#) / Модуль python-magic в Python, идентификация типа файла

## Идентификация типа файла по содержанию

[Модуль python-magic](#) представляет собой интерфейс Python для библиотеки идентификации типов файлов libmagic. Библиотека libmagic идентифицирует типы файлов, проверяя их заголовки в соответствии с предопределенным списком типов файлов. Эта функциональность предоставляется командной строке командным файлом Unix.

Модуль будет полезен при идентификации типа файла переданного по сети интернет (например, который был загружен на веб-сервер) перед его использованием, чтобы избежать выполнения вредоносного кода (например, передан файл .php с расширением .pdf).

## Установка модуля python-magic в виртуальное окружение.

Модуль python-magic размещен на PyPI, поэтому установка относительно проста.

```
# создаем виртуальное окружение, если нет
$ python3 -m venv .venv --prompt VirtualEnv
# активируем виртуальное окружение
$ source .venv/bin/activate
# обновляем `pip`
(VirtualEnv):~$ python3 -m pip install -U pip
# ставим модуль `python-magic`
(VirtualEnv):~$ python3 -m pip install -U python-magic
```

Установка в Windows:

```
> pip install python-magic-bin
```

## Примеры использования модуля tomli\_w.

```
>>> import magic
>>> magic.from_file("testdata/test.pdf")
# 'PDF document, version 1.2'
```

Рекомендуется использовать по крайней мере первые 2048 байт, так как меньшее количество может привести к неправильной идентификации:

```
>>> magic.from_buffer(open("testdata/test.pdf", "rb").read(2048))
# 'PDF document, version 1.2'

# извлечение MIME типа
>>> magic.from_file("testdata/test.pdf", mime=True)
# 'application/pdf'
```

Существует также класс magic.Magic(), который обеспечивает более прямое управление, включая переопределение файла базы данных модуля python-magic и включение определения кодировки символов. Этот класс не рекомендуется для общего использования. В частности, это небезопасно для совместного использования несколькими потоками, и при попытке сделать это произойдет сбой.

Вверх

```
>>> fl = magic.Magic(uncompress=True)
>>> fl.from_file('testdata/test.gz')
# 'ASCII text (gzip compressed data, was "test", last modified: Sat Jun 28 21:32:52 2008, from Unix)'
```

Можно также можно комбинировать аргументы:

```
>>> f = magic.Magic(mime=True, uncompress=True)
>>> f.from_file('testdata/test.gz')
# 'text/plain'
```