

ХОЧУ ПОМОЧЬ

# Функции escape() и unescape() модуля html в Python



besteq.ru

РЕКЛАМА

## Этикетировочная машина BESTEQ-HALS. Аналог ОБ-КЭТ-С2

Машина для нанесения бумажной круговой этикетки на консервную тару. До 500 бан/мин

Узнать цену

### Бесплатное занятие английским в Яндекс Практиуме

Полноценное занятие с преподавателем, а не презентация курсов

Устный тест на уровень языка

Практика английского

Узнать больше

Функции escape() и unescape() модуля html в Python

## Символы &, < и > в безопасные последовательности HTML

В этом разделе будут рассмотрены только две функции [html.escape](#) и [html.unescape](#) используемые в Python.

В этом разделе подмодуль [html.entities](#) будет рассмотрен только часто используемый словарь

### html.escape(s):

Функция html.escape(s) переводит символы &, < и > в строке s в безопасные последовательности HTML.

Если необходимо отобразить на HTML странице текст, который может содержать вышеуказанные символы - внедрение вредоносного JavaScript-кода.

Если необязательный флаг quote имеет значение True, то также переводятся символы (") и ('). Это поможет включить в HTML-атрибут значение, разделенное кавычками, например как в <a id="...">.

```
>>> import html
>>> xxs = """<script type = 'text/javascript'>
... alert('This XXS!!!');
... </script>"""
>>> rez = html.escape(xxs)
>>> print(rez)
# &lt;script type = &#x27;text/javascript&#x27;&gt;
# alert(&#x27;This XXS!!!&#x27;);
# &lt;/script&gt;
```

### html.unescape(s):

Функция html.unescape(s) преобразует все именованные и числовые ссылки на символы (например &gt;, &#62;, &#x3e;) в строке s в соответствующие символы Юникода. Другими словами, эта функция обратная [html.escape\(\)](#).

Функция html.unescape() использует правила, определенные стандартом HTML 5 для допустимых и недопустимых ссылок на символы, а также список именованных ссылок на символы HTML 5.

```
>>> import html
>>> htm = "&lt;script type = &#x27;text/javascript&#x27;&gt;
... alert(&#x27;This XXS!!!&#x27;);
... &lt;/script&gt;"""
>>> rez = html.unescape(htm)
>>> print(rez)
# <script type = 'text/javascript'>
# alert('This XXS!!!');
# </script>
```

### html.entities.html5:

Атрибут html5 подмодуля html.entities представляет собой [словарь](#), который сопоставляет HTML5 именованные ссылки на символы с их эквивалентными символами Unicode. Например html5['gt;'] == '>'.

Вверх

Обратите внимание, что конечная точка с запятой ';' включена в имя (например, 'gt;'), но некоторые имена принимаются стандартом даже без точки с запятой: в этом случае имя присутствует с символом ';' и без него.

>>> from html.entities import html5

РЕКЛАМА · 18+ ⓘ



practicum.yandex.ru

**Бесплатное занятие английским в Яндекс**

[DOCS-Python.ru™](#), 2023 г.

Полноценное занятие с преподавателем, а не презентация курсов

Устный тест на уровень языка >

Практика английского >

Узнать больше

(Внимание! При копировании материала ссылка на источник обязательна)

[@docs\\_python\\_ru](#)

Вверх

https://docs-python.ru/standart-library/funktsii-escape-unescape-modulja-html-python/2/2