
baidubaike Documentation

Release 1.3.0

yakiang

January 29, 2016

1	Page	3
2	Baike.search	5
3	Exception	7
3.1	PageError	7
3.2	DisambiguationError	7
3.3	VerifyError	7
3.4	Other	7
4	Indices and tables	9

There're two main functionalities:

To create a page:

```
>>> from baidubaike import Page
>>> page = Page('google')
```

get the basic information about the page as a dict:

```
>>> info = page.get_info()
>>> print info['title'], info['url']

>>> print info.get('last_modify_time')
>>> print info.get('creator')
>>> print info.get('page_view')
# these keys might not be included
```

get the content of the page as a string:

```
>>> page.get_content()
```

get urls referring to other pages as an OrderedDict:

```
>>> links = page.get_inurls()
>>> for word in links:
...     print word, links[word]
```

get a list of tags of the page:

```
>>> page.get_tags()
```

get a list of reference links:

```
>>> ref = page.get_references()
>>> for r in ref:
...     print r['title']
...     print r['url']
```

besides, you can also create a Page with an existing url referring to it:

```
>>> page = Page('http://baike.baidu.com/view/105.htm')
```

you can set the encoding, by default it is *utf-8*:

```
>>> page = Page('google', encoding='gbk')
```

Baike.search

By default it queries 10 results at page 1:

```
>>> from baidubaike import Search
>>> search = Search('google')
```

to get more results at other page number:

```
>>> search = Search('google', results_n=20, page_n=2)
# get 20 results at page 2
```

to get the results:

```
>>> results = search.get_results()

>>> for r in results:
...     print r['title']
...     print r['url']
...     print r['discription']
```

Exception

3.1 PageError

Raised when no such a page exists.

3.2 DisambiguationError

Raised when a word causes two or more results.

While throwing errors it will return some items that you may refer to.

3.3 VerifyError

Raised when a page needs a verifying-code.

This may be caused when you query too frequently.

3.4 Other

Raised by `requests`

Indices and tables

- `genindex`
- `modindex`
- `search`