

EShw04

1. 简述查询时分析器的查找顺序?

用于搜索特定字段的分析器由以下内容顺序确定：

在查询中指定分析器。

mapping参数指定的search_analyzer

mapping参数中指定的analyzer

在索引设置指定的default_search分析器

在索引设置指定的default分析器

standard分析器。

2. 自定义分析器， 分析文本： "I'm a :) person, and you?", 实现如下功能:

- (1)、去除英文停用词
- (2)、转换小写
- (3)、去除html标签
- (4、将:)转换为*hadppy*

```

GET my_index

PUT my_index
{
  "settings": {
    "analysis": {
      "analyzer": {
        "my_custom_analyzer": {
          "type": "custom",
          "tokenizer": "standard",
          "char_filter": [
            "html_strip", "my_char_filter"
          ],
          "filter": [
            "lowercase",
            "asciifolding"
          ]
        }
      },
      "char_filter": {
        "my_char_filter": {
          "type": "mapping",
          "mappings": [
            ":) => _happy_",
            ":( => _sad_"
          ]
        }
      }
    }
  }
}

POST my_index/_analyze
{
  "analyzer": "my_custom_analyzer",
  "text": "I'm a :) person, and <b>you</b>?"
}

```

```

1 {
2   "tokens": [
3     {
4       "token": "i'm",
5       "start_offset": 0,
6       "end_offset": 3,
7       "type": "<ALPHANUM>",
8       "position": 0
9     },
10    {
11      "token": "a",
12      "start_offset": 4,
13      "end_offset": 5,
14      "type": "<ALPHANUM>",
15      "position": 1
16    },
17    {
18      "token": "_happy_",
19      "start_offset": 6,
20      "end_offset": 8,
21      "type": "<ALPHANUM>",
22      "position": 2
23    },
24    {
25      "token": "person",
26      "start_offset": 9,
27      "end_offset": 15,
28      "type": "<ALPHANUM>",
29      "position": 3
30    },
31    {
32      "token": "and",
33      "start_offset": 17,
34      "end_offset": 20,
35      "type": "<ALPHANUM>",
36      "position": 4
37    },
38    {
39      "token": "you",
40      "start_offset": 25,
41      "end_offset": 32,
42      "type": "<ALPHANUM>",
43      "position": 5
44    }
45  ]
46 }

```

In []: