例：



这个表单里面 “索引号” 和 “文号”都有空格，一般我们抽取索引号的内容这么写xpath

//table//td[contains(string(), "索 引 号")]

或者

//table//td[text() = "索 引 号"]

但是这样是提取不出来的，具体原因未知

解决方法

1） //td[normalize-space(text()) = '索 引 号']

2）//td[normalize-space(.) = '索 引 号']

3）//td[contains(text(), '索') and contains(text(), '引') and contains(text(), '号')]

4）//td[translate(text(), " &#13;&#10;&#09;&#xa;", "") = '索引号']

5）//td[matches(text(), '\s\*索\s+引\s+号\s\*')]

参考地址

<https://stackoverflow.com/questions/24362178/using-xpath-to-search-text-containing-whitespace?lq=1>