Question

Jack Abdo · Feb 2, 2016

IFind french and stemming

Hi,

I created with Studio a persistent class with the following field and index:

```
Property DescriptionDemande As %String(MAXLEN = "");
Index IDXBASDescriptionDemande On
(DescriptionDemande) As %iFind.Index.Basic(INDEXOPTION = 1, LANGUAGE = "fr", LOWER = 1);
```

INDEXOPTION is set to 1 for activating stemming. I'm indexing french documents. I have set lower to 1 because I want to do non case sensitive search.

I inserted a single french word "élément" in the field DescriptionDemande for testing purposes using this query: insert into mytable(DescriptionDemande) values(' élément')

the query SELECT * FROM mytable WHERE %ID %FIND searchindex(IDXBASDescriptionDemande,'élément',0) returns one row and that is fine.

But:

- 1. The query SELECT * FROM mytable WHERE %ID %FIND searchindex(IDXBASDescriptionDemande, 'éléments',0) returns nothing. If stemming was activated the ending 's' should be ignored and the query should return one row.
- 2. The query SELECT * FROM mytable WHERE %ID %FIND searchindex(IDXBASDescriptionDemande, 'Elément',0) returns nothing, despite that Elément and élément in french are the same word (the accent on the capital letters are not commonly used)

Why the stemming is not working? is there something missing in my index configuration.

What can I do tell iFind not to take accents into consideration, without modifying my own text of course.

Regards,

Jack Abdo.

#Beginner #Development Environment #Studio #Unstructured Data #iFind #Caché

Source URL: https://community.intersystems.com/post/ifind-french-and-stemming