

Regular Expression Search

Contents of this Page

- [Regular Expression Implementation Details](#)
 - [Algorithm:](#)
 - [Example of use:](#)
 - [Associated JUnits:](#)

Regular Expression Implementation Details

The regular expression search searches against lower cased text. Additionally, this searches against the entire string as a single token, rather than the tokenized string.

Algorithm:

The Regular Expression search has the following characteristics:

- This search searches only lower cased text.
- It searches on the untokenized lower cased property value.
- Analyzers are not applied to the expression. However, the expression is lower cased (this is an explicit step done outside of Lucene by LexEVS code).

Example of use:

The following examples are based on the Automobiles coding scheme.

Example 1:

Search string: automobi.*

Lucene query: untokenizedLCPropertyValue:automobi.*

Result: 1 result

- entity code: A0001
- entity description: Automobile

Example 2:

Search string: .*utomobile

Lucene query: untokenizedLCPropertyValue:.*utomobile

Result: 1 result

- entity code: A0001
- entity description: Automobile

Associated JUnits:

Junits can be found here: <https://github.com/lexevs/lexevs/blob/master/lbTest/src/test/java/org/LexGrid/LexBIG/Impl/function/query/lucene/searchAlgorithms/TestRegExp.java>