

Package ‘RKEA’

February 14, 2012

Version 0.0-3

Date 2010-01-22

Title R/KEA interface

Author Ingo Feinerer, Kurt Hornik

Maintainer Kurt Hornik <Kurt.Hornik@R-project.org>

Description An R interface to KEA (Version 5.0). KEA (for Keyphrase Extraction Algorithm) allows for extracting keyphrases from text documents. It can be either used for free indexing or for indexing with a controlled vocabulary. For more information see <http://www.nzdl.org/Kea/>.

Imports rJava (>= 0.6-3), tm

SystemRequirements Java (>= 5.0)

License GPL-2

Repository CRAN

Date/Publication 2012-01-24 16:13:45

R topics documented:

createModel 2
extractKeywords 3

Index 4

`createModel`*Create a KEA Model*

Description

Create a keyphrase extraction model.

Usage

```
createModel(corpus, keywords, model, voc = "none", vocformat = "")
```

Arguments

<code>corpus</code>	A list of character vectors containing the text documents, e.g., a Corpus object as provided by package <code>tm</code> .
<code>keywords</code>	A list of character vectors containing the keywords for each document in corpus.
<code>model</code>	A character giving the path where the created model should be stored.
<code>voc</code>	A character pointing to a controlled vocabulary.
<code>vocformat</code>	A character giving the format of voc.

Details

A tutorial on keyword extraction is located at <http://www.nzdl.org/Kea/Download/Kea-5.0-Readme.txt>. There you can find details on the internals of KEA, including various parameter settings (e.g., valid arguments for `voc` and `vocformat`).

Value

Invisibly returns `model`, i.e., the path to the created KEA model.

Author(s)

Ingo Feinerer

References

<http://www.nzdl.org/Kea/>

See Also

[extractKeywords](#)

extractKeywords	<i>Extract Keywords</i>
-----------------	-------------------------

Description

Extract keywords from text documents.

Usage

```
extractKeywords(corpus, model, voc = "none", vocformat = "")
```

Arguments

corpus	A list of character vectors containing the text documents, e.g., a Corpus object as provided by package <code>tm</code> , used for keyword extraction.
model	A character giving the path to a KEA model.
voc	A character pointing to a controlled vocabulary.
vocformat	A character giving the format of voc.

Details

A tutorial on keyword extraction is located at <http://www.nzdl.org/Kea/Download/Kea-5.0-Readme.txt>. There you can find details on the internals of KEA, including various parameter settings (e.g., valid arguments for voc and vocformat).

Value

A list of character vectors corresponding to the keywords in corpus.

Author(s)

Ingo Feinerer

References

<http://www.nzdl.org/Kea/>

See Also

[createModel](#)

Index

*Topic **file**

createModel, [2](#)

extractKeywords, [3](#)

createModel, [2](#), [3](#)

extractKeywords, [2](#), [3](#)