

Package: RKEA (via r-universe)

November 1, 2024

Version 0.0-6

Title R/KEA Interface

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 RKEAjars (>= 5.0-1), rJava (>= 0.6-3), tm

SystemRequirements Java (>= 5.0)

License GPL-2

Author Ingo Feinerer [aut], Kurt Hornik [aut, cre]

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

NeedsCompilation no

Date/Publication 2015-04-03 17:27:56

Repository <https://kurthornik.r-universe.dev>

RemoteUrl <https://github.com/cran/RKEA>

RemoteRef HEAD

RemoteSha 3979474b6f49f6b455dcbb27fc3455f075f25515

Contents

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

corpus	A list of character vectors containing the text documents, e.g., a Corpus object as provided by package tm .
keywords	A list of character vectors containing the keywords for each document in corpus.
model	A character giving the path where the created model should be stored.
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., details on vocabularies and supported formats for these).

When controlled vocabularies are used (by default: no), the `voc` argument should give the file path to the respective files without their extensions. When `vocformat` is "skos", the extension must be ".rdf"; when "text", there must be files with extensions ".en", ".rel" and ".use".

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

* **file**

createModel, [2](#)

extractKeywords, [3](#)

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

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