Package ‘MeSHSim’

Type Package

Author Jing Zhou, Yuxuan Shui

Maintainer Jing Zhou <12210240050@fudan.edu.cn>

Title MeSH (Medical Subject Headings) Semantic Similarity Measures

Version 1.7.0

Date 2015-01-12

Imports XML, RCurl

Depends R(>= 3.0.0)

Description Provide for measuring semantic similarity over MeSH headings and MEDLINE documents

License GPL-2

bioCViews Clustering, Software

NeedsCompilation no

R topics documented:

docInfo .......................... 2
docSim ................................ 2
headingSetSim .......................... 3
headingSim .......................... 4
MeshSimData ......................... 4
mheadingSim .......................... 5
mnodeSim .......................... 5
nodeInfo .......................... 6
nodeSim .......................... 7
termInfo .......................... 7

Index .......................... 9
docInfo  

*Get details of documents*

**Description**

Fetch information of a given article from PubMed.

**Usage**

```r
docInfo(pmid, verbose=FALSE, major=FALSE)
```

**Arguments**

- `pmid`: pmid of the desired article.
- `verbose`: whether the title and abstract of the article should be print out.
- `major`: whether only major terms should be returned.

**Value**

Document information of given PMID including title, abstract, MeSH headings

**Note**

Network connection is required for using this function.

**Examples**

```r
docInfo("1111111")
```

docSim  

*Similarity between articles*

**Description**

Calculate the similarity between two articles.

**Usage**

```r
docSim(pmid1, pmid2, method="SP", frame="node", major=FALSE, env=NULL)
```

**Arguments**

- `pmid1, pmid2`: pmids of two articles whose similarity is needed to be calculated.
- `method`: similarity measurement method, see Details for available methods.
- `frame`: available options are node and heading, decide whether using node-based or heading-based methods.
- `env`: the dataset to use.
- `major`: whether the calculation should only be based on major terms
headingSetSim

Details


Value

Semantic similarity between two MEDLINE documents, whose value is between 0 and 1.

Note

Network connection is required for using this function.

Examples

docSim("1111113","1111111")

headingSetSim headingSet1, headingSet2, method="SP", frame="node", env=NULL)

Arguments

headingSet1, headingSet2
two lists of headings

method similarity measurement method, see Details for available methods.

frame available options are node and heading, decide whether using node-based or heading-based methods.

env the dataset to use.

Details


Value

Semantic similarity between two MeSH heading sets, whose value is between 0 and 1.

Examples

headingSet1<-c("Lumbosacral Region", "Body Regions")
headingSet2<-c("Body Regions", "Abdomen", "Abdominal Cavity")
headingSetSim(headingSet1,headingSet2,"SP","node")
headingSim  

**Description**

Calculate similarity between two headings.

**Usage**

\[
\text{headingSim(heading1, heading2, method="SP", frame="node", env=NULL)}
\]

**Arguments**

- **heading1, heading2**: two headings or two lists of headings
- **method**: similarity measurement method, see Details for available methods.
- **frame**: available options are node and heading, decide whether using node-based or heading-based methods.
- **env**: the dataset to use.

**Details**


**Value**

Semantic similarity between two MeSH headings, whose value is between 0 and 1.

**Examples**

\[
\text{headingSim("Lumbosacral Region", "Body Regions")}
\]

---

**MeshSimData**  

**MeSH Dataset**

**Description**

These contents data of the whole MeSH tree, as well as information contents for every node and term.

This dataset will be auto loaded by the first invoked function of this package, if no other dataset is specified.
mheadingSim

Similarity between heading lists

Description

Calculate similarity matrix between two heading lists.

Usage

mheadingSim(headingList1, headingList2, method="SP", frame="node", env=NULL)

Arguments

headingList1, headingList2
two headings or two lists of headings
method
similarity measurement method, see Details for available methods.
frame
available options are node and heading, decide whether using node-based or heading-based methods.
env
the dataset to use.

Details


Value

Semantic similarity matrix between two MeSH heading lists.

Examples

headingList1<-c("Body Regions", "Abdomen", "Abdominal Cavity")
headingList2<-c("Lumbosacral Region", "Body Regions")
mheadingSim(headingList1,headingList2)

mnodeSim

Similarity between node lists

Description

Calculate similarity matrix between two MeSH node lists.

Usage

mnodeSim(nodeList1, nodeList2, method="SP", frame="node", env=NULL)
Arguments
nodeList1, nodeList2
two nodes or two lists of nodes
method
similarity measurement method, see Details for available methods.
frame
available options are node and heading, decide whether using node-based or
heading-based methods.
env
the dataset to use.

Details
method, LC: Leacock and Chodorow’s method, Li: Li’s method, Lord: Lord’s method, Resnik:
Resnik’s method, Lin: Lin’s method, JC: Jiang and Conrath’s method.

Value
Semantic similarity matrix between two MeSH node lists.

Examples
nodeList1<-c("B03.440.400.425.340.590", "B03.440.400.425.117.800.200")
nodeList2<-c("B03.440.400.425.340.590", "B03.440.400.425.117.800.200", "B03.440.400.425.127.100")
mnodeSim(nodeList1, nodeList2)

Description
This function returns a tree contains the given node.

Usage
nodeInfo(node, brief, env=NULL)

Arguments
node
a node name.
brief
brief model for nodeInfo
env
the dataset to use.

Value
Hierarchy information of node a

Examples
nodeInfo("B03.440.400.425.127")
nodeInfo("B03.440.400", brief=TRUE)
nodeSim

Similarity between nodes

Description

Calculate similarity between two MeSH nodes.

Usage

nodeSim(node1, node2, method="SP", frame="node", env=NULL)

Arguments

node1, node2
  two nodes or two lists of nodes
method
  similarity measurement method, see Details for available methods.
frame
  available options are node and heading, decide whether using node-based or
  heading-based methods.
env
  the dataset to use.

Details

method, LC: Leacock and Chodorow’s method, Li: Li’s method, Lord: Lord’s method, Resnik:
Resnik’s method, Lin: Lin’s method, JC: Jiang and Conrath’s method.

Value

Semantic similarity between two MeSH nodes, whose value is between 0 and 1.

Examples

nodeSim("B03.440.400.425.340.590", "B03.440.400.425.117.800.200")

termInfo

Details of MeSH terms

Description

This function returns a tree contains the given term.

Usage

termInfo(term, brief, env=NULL)

Arguments

term
  a term name.
brief
  whether to retrieve brief tree information of MeSH term
env
  the dataset to use.
Value

Hierarchy information of a given term

Examples

termInfo("Body Regions")
Index

«Topic MeSH
  docInfo, 2
  docSim, 2
  headingSetSim, 3
  headingSim, 4
  mheadingSim, 5

  docInfo, 2
  docSim, 2
  headingSetSim, 3
  headingSim, 4
  MeshSimData, 4
  mheadingSim, 5
  mnodeSim, 5
  nodeInfo, 6
  nodeSim, 7
  termInfo, 7