

Package ‘edeR’

August 29, 2016

Type Package

Title Email Data Extraction Using R

Version 1.0.0

Date 2014-02-17

Author Jaynal Abedin

Maintainer Jaynal Abedin <joystatru@gmail.com>

Description This package will allow to connect with email server through Internet Message Access Protocol (IMAP) and extract header information e.g. from, to, cc, subject, date and time. User will supply their email address and password along with other options. Initially this package is developed only for Gmail. To run the functions from this package user have to have IMAP enabled Gmail account.

Depends rJava, rjson, rJython

SystemRequirements Java Runtime Environment

License GPL-2

NeedsCompilation no

Repository CRAN

Date/Publication 2014-02-17 12:12:37

R topics documented:

edeR	2
extractBetween	3
extractKeyword	4
extractKeywordB	5
extractMbox	6

Index	8
--------------	----------

Description

This package allows user to connect with Gmail server through IMAP and extract email header information from, to, cc, subject, and datetime. User can extract email data from any specific folder or label using the argument `folder`. To access "Drafts" and "Sent Mail" we need to input the argument as "[Gmail]/Drafts" and "[Gmail]/Sent Mail". Moreover user can search email with specific keyword in subject line or in email body message.

Details

Package:	edeR
Type:	Package
Depends:	rJava, rjson, rJython
System dependency:	Java runtime environment
Version:	1.0.0
Date:	2014-02-17
License:	GPL-2

Author(s)

Jaynal Abedin

Maintainer: Jaynal Abedin <joystatru@gmail.com>

Examples

```
## Not run:
# Extract maximum 5 emails from the specified folder.
# User needs to provide his/her valid Gmail address and password.
dat1 <- extractMbox(username="username@gmail.com",
  password="password",
  folder="folder",
  nmail=5)

# To extract maximum of 5 emails from inbox between June 6, 2013 to June 6, 2013
dat2 <- extractBetween(username="username@gmail.com",
  password="password",
  folder="inbox",
  startDate="06-Jun-2013",
  endDate="06-Jun-2013",
  nmail=5)
```

```
# To extract maximum 5 emails with the word "keyword" within subject line
dat3 <-extractKeyword(username="username@gmail.com",
                      password="password",
                      kw="keyword",
                      nmail=5)
# To extract maximum 5 emails with the word "keyword" within body message
dat4 <-extractKeywordB(username="username@gmail.com",
                       password="password",
                       kw="keyword",
                       nmail=5)

## End(Not run)
```

extractBetween	<i>Extracts email within given date range</i>
----------------	---

Description

This function extracts email data from a specified folder with given date range

Usage

```
extractBetween(username, password, folder, startDate, endDate, nmail = -1)
```

Arguments

username	A character string of email address for example, "xxx@gmail.com"
password	A character string of password of the email address for example "xyz123"
folder	A character string of folder or label name from where we want to extract email data
startDate	A character starting of date (e.g "06-Jun-2013") of interest of the date range. The format of date should be same as shown in the example, any other format will give error
endDate	A character string of end date (e.g. "06-Jun-2013")of interest of the date range
nmail	An integer for the number of emails to extract, default is all emails (nmail=-1)

Details

This function is able to extract emails from any specified folder within a given date range. To extract "Sent Mail" or "Drafts" we have to specify the argument differently. To extract "Sent Mail", the argument is folder="[Gmail]/Sent Mail"

Value

n_message	Number of available email in the specified folder
data	The extracted data frame with the field unique id (uid), sender (from), receiver either direct or cc'd (to), subject of mail (subj), time stamp (datetime)

Note

To extract "Sent Mail", the argument is folder="[Gmail]/Sent Mail"

Author(s)

Jaynal Abedin

See Also

[extractMbox](#)

Examples

```
## Not run:
# To extract maximum of 5 emails from inbox between June 6, 2013 to June 6, 2013
dat <-extractBetween(username="username@gmail.com",
                    password="password",
                    folder="INBOX",
                    startDate="06-Jun-2013",
                    endDate="06-Jun-2013",
                    nmail=5)

## End(Not run)
```

extractKeyword	<i>Extracts email with specified keyword in the subject line</i>
----------------	--

Description

This function extracts email data with specific keyword in the subject line

Usage

```
extractKeyword(username, password, kw, nmail = -1)
```

Arguments

username	A character string of email address for example, "xxx@gmail.com"
password	A character string of password of the email address for example "xyz123"
kw	A character string of specific keyword that will be searched within subject line
nmail	An integer for the number of emails to extract, default is all emails (nmail=-1)

Details

Extracts emails with exact keyword kw in the subject line

Value

n_message	Number of available email in the specified folder
data	The extracted data frame with the field unique id (uid), sender (from), receiver either direct or cc'd (to), subject of mail (subj), time stamp (datetime)

Author(s)

Jaynal Abedin

See Also

[extractKeywordB](#)

Examples

```
## Not run:
# To extract maximum 5 emails with the word "keyword" within subject line
dat <-extractKeyword(username="username@gmail.com",
                    password="password",
                    kw="keyword",
                    nmail=5)

## End(Not run)
```

extractKeywordB	<i>Extracts email with specified keyword in the email body text</i>
-----------------	---

Description

This function extracts email data with specific keyword in the body message

Usage

```
extractKeywordB(username, password, kw, nmail = -1)
```

Arguments

username	A character string of email address for example, "xxx@gmail.com"
password	A character string of password of the email address for example "xyz123"
kw	A character string of specific keyword that will be searched within subject line
nmail	An integer for the number of emails to extract, default is all emails (nmail=-1)

Details

Extracts emails matching the exact keyword kw in the body message

Value

n_message	Number of available email in the specified folder
data	The extracted data frame with the field unique id (uid), sender (from), receiver either direct or cc'd (to), subject of mail (subj), time stamp (datetime)

Author(s)

Jaynal Abedin

See Also

[extractKeyword](#)

Examples

```
## Not run:
# To extract maximum 5 emails with the word "keyword" within body message
dat <-extractKeywordB(username="username@gmail.com",
                      password="password",
                      kw="keyword",
                      nmail=5)

## End(Not run)
```

extractMbox	<i>Extracts emails from specified folder</i>
-------------	--

Description

This function extracts email data from a specified folder

Usage

```
extractMbox(username, password, folder, nmail = -1)
```

Arguments

username	A character string of email address for example, "xxx@gmail.com"
password	A character string of password of the email address for example "xyz123"
folder	A character string of specific folder or label name from where the email data will be extracted
nmail	An integer for the number of emails to extract, default is all emails (nmail=-1)

Details

This function allow users to extract email data from specified folder. By default this function will extract all available emails from the specified folder but user can restrict number of emails to extract using nmail.

Value

n_message	Number of available email in the specified folder
data	The extracted data frame with the field unique id (uid), sender (from), receiver either direct or cc'd (to), subject of mail (subj), time stamp (datetime)

Note

To extract "Sent Mail", the argument is `folder="[Gmail]/Sent Mail"`

Author(s)

Jaynal Abedin

See Also

[extractBetween](#)

Examples

```
## Not run:  
# Extract maximum 5 emails from the specified folder  
dat<-extractMbox(username="username@gmail.com",  
                 password="password",  
                 folder="folder",  
                 nmail=5)  
  
## End(Not run)
```

Index

*Topic **package**

edeR, [2](#)

edeR, [2](#)

extractBetween, [3](#), [7](#)

extractKeyword, [4](#), [6](#)

extractKeywordB, [5](#), [5](#)

extractMbox, [4](#), [6](#)