

Package: europarl (via r-universe)

March 1, 2025

Type Package

Title Scrap Data from Europarlament's Website

Version 0.1.0

Author Szymon Gorka

Maintainer The package maintainer <yourself@somewhere.net>

Description Scrap data from europarlament's website.

License GPL

URL <https://ropengov.github.io/europarl/>,
<https://github.com/rOpenGov/europarl>

BugReports <https://github.com/rOpenGov/europarl/issues>

Imports DBI, dplyr, lubridate, magrittr, readr, RMySQL, RSelenium,
rvest, stringr, tidyverse, XML, xml2

Encoding UTF-8

LazyData true

RoxygenNote 7.1.2

X-schema.org-isPartOf <http://ropengov.org/>

X-schema.org-keywords ropengov

Config/pak/sysreqs libfontconfig1-dev libfreetype6-dev libfribidi-dev
make libharfbuzz-dev libicu-dev libjpeg-dev libmysqlclient-dev
libpng-dev libtiff-dev libxml2-dev libssl-dev libx11-dev
zlib1g-dev

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

RemoteUrl <https://github.com/rOpenGov/europarl>

RemoteRef HEAD

RemoteSha f6bda62a433f8c5595078f3b93df877d706cac6f

Contents

create_database	2
deputies_get_history_of_services	2
get_all_deputies	3
get_eurogroup	3
get_history	3
get_languages	3
get_more_info	4
get_nationalparty	4
get_statements	4
read_config	5
safe_html	6
statements_core	7
statements_get_all_statements	7
statements_get_list_of_statements	7
statements_get_statement	7
statements_get_time	8
statements_update_statements	8
strings_identical	8

Index	9
--------------	----------

create_database	<i>Creating database</i>
-----------------	--------------------------

Description

Function create_database creates a dabase with

Usage

create_database(dbname, user, password, host)

deputies_get_history_of_services
<i>Get national parties, eu groups, postions</i>

Description

Get national parties, eu groups, postions

Usage

deputies_get_history_of_services(home_page, deputy_id)

`get_all_deputies` *Get data about all deputies*

Description

Get data about all deputies

Usage

```
get_all_deputies(term = 1)
```

`get_eurogroup` *Get data about all deputies*

Description

Get data about all deputies

Usage

```
get_eurogroup(deputy_id, date = Sys.Date())
```

`get_history` *Get parties, eu groups, postions*

Description

Get parties, eu groups, postions

Usage

```
get_history(home_page, deputy_id)
```

`get_languages` *Get all languages in europarl*

Description

Get all languages in europarl

Usage

```
get_languages()
```

get_more_info	<i>Get nationality, date of birth, place of birth and/or date of death</i>
---------------	--

Description

Get nationality, date of birth, place of birth and/or date of death

Usage

```
get_more_info(home_page)
```

get_nationalparty	<i>Get data about all deputies</i>
-------------------	------------------------------------

Description

Get data about all deputies

Usage

```
get_nationalparty(deputy_id, date = Sys.Date())
```

get_statements	<i>Get all statements for P8</i>
----------------	----------------------------------

Description

Get all statements for P8

Usage

```
get_statements(deputy_id, browser)
```

read_config	<i>Loads data from configuration file</i>
-------------	---

Description

File should contains:

- dbname = "dbname"
- host = "host"
- username = "username"
- password = "password"

Usage

```
read_config(  
  file = system.file("config/db_config.txt", package = "europarl"),  
  delim = " "  
)
```

Arguments

delim	a delim parametr in read_delim
name	file name or path to file

Value

A tibble with dbname, host, username and password for database connction.

Examples

```
## Not run:  
read_config()  
  
read_config(file = "path/name.txt", delim = " ")  
  
## End(Not run)
```

`safe_html`*Safe html scrapping*

Description

Function `safe_html` tries to download the URL several times.

Usage

```
safe_html(page, time = 60, attempts = 10)
```

Arguments

<code>page</code>	requested URL
<code>time</code>	sleep interval after each failure
<code>attempts</code>	max number of tries (if there is a problem with connection)

Details

Function `safe_html` performs 10 (by default) attempts to download the URL and waits 60sec (by default) after each failure

Value

character vector

Author(s)

Przemyslaw Biecek

Examples

```
## Not run:  
page <- paste0('http://www.sejm.gov.pl/Sejm7.nsf/',  
              'wypowiedz.xsp?posiedzenie=15&dzien=1&wyp=008')  
safe_html(page)  
## End(Not run)
```

statements_core *Get statements text and details.*

Description

Function statements_core downloads content and details(language, time) of the statement.

Usage

```
statements_core(url)
```

statements_get_all_statements
Get all statements for P8

Description

Get all statements for P8

Usage

```
statements_get_all_statements(deputy_id, browser, term_of_office = 8)
```

statements_get_list_of_statements
Return links, titles, dates of stamenets of given deputie.

Description

Return links, titles, dates of sttamenets of given deputie.

Usage

```
statements_get_list_of_statements(deputy_id, browser, term_of_office = 8)
```

statements_get_statement
Get statements text and details.

Description

Function statements_core downloads content and details(language, time) of the statement.

Usage

```
statements_get_statement(url)
```

statements_get_time *Get time of statements*

Description

Get time of statements

Usage

```
statements_get_time(url)
```

Arguments

url A url of statements

Value

Returns duration, start time and end time of statement

statements_update_statements
 Update stamenets in data abse

Description

Update stamenets in data abse

Usage

```
statements_update_statements(deputy_id, browser, term_of_office = 8, db)
```

strings_identical *Remove white signs and comapre two strings*

Description

Remove white signs and comapre two strings

Usage

```
strings_identical(x, y)
```

Value

Returns TRUE or FALSE

Index

`create_database`, 2

`deputies_get_history_of_services`, 2

`get_all_deputies`, 3

`get_eurogroup`, 3

`get_history`, 3

`get_languages`, 3

`get_more_info`, 4

`get_nationalparty`, 4

`get_statements`, 4

`read_config`, 5

`safe_html`, 6

`statements_core`, 7

`statements_get_all_statements`, 7

`statements_get_list_of_statements`, 7

`statements_get_statement`, 7

`statements_get_time`, 8

`statements_update_statements`, 8

`strings_identical`, 8