

Package ‘ennet’

February 3, 2021

Type Package

Title Utilities to Extract and Analyse Text Data from the Emergency Nutrition Network Forum

Version 0.2.1

Description The Emergency Nutrition Network or en-net forum is the go to online forum for field practitioners requiring prompt technical advice for operational challenges for which answers are not readily accessible in current guidelines. The questions and the corresponding answers raised within en-net can provide insight into what the key topics of discussion are within the nutrition sector. This package provides utility functions for the extraction, processing and analysis of text data from the online forum.

License GPL-3

Depends R (>= 2.10)

Imports magrittr, rvest, xml2, tibble, dplyr, stringr, lubridate, utils, tidyr

Suggests knitr, rmarkdown, testthat, spelling

Encoding UTF-8

Language en-GB

LazyData true

RoxygenNote 7.1.1

URL <https://katilingban.io/ennet/>,
<https://github.com/katilingban/ennet>

BugReports <https://github.com/katilingban/ennet/issues>

VignetteBuilder knitr

NeedsCompilation no

Author Ernest Guevarra [aut, cre] (<<https://orcid.org/0000-0002-4887-4415>>)

Maintainer Ernest Guevarra <ernest@guevarra.io>

Repository CRAN

Date/Publication 2021-02-03 16:20:02 UTC

R topics documented:

arrange_replies	2
arrange_views	3
count_authors	4
count_topics	5
count_topics_author	5
count_topics_day	6
count_topics_theme	7
create_db_topics_dailies	8
create_db_topics_daily	9
create_db_topics_hourlies	10
create_db_topics_interactions	11
create_db_topics_monthly	12
ennet-deprecated	12
ennet_dailies	13
ennet_hourlies	14
ennet_themes	14
ennet_topics	15
get_db_discussions	16
get_themes	17
get_themes_topics	17
get_theme_topics	18
get_topics_discussions	19
get_topic_discussions	19
update_topics	20
Index	21

arrange_replies	<i>Arrange topics based on number of replies</i>
-----------------	--

Description

Arrange topics based on number of replies

Usage

```
arrange_replies(
  topics,
  by_theme = TRUE,
  by_date = c("month_year", "year", "all")
)
```

Arguments

topics	A tibble of topics by theme from en-net forum produced through a call to <code>get_themes_topics()</code> .
by_theme	Logical. Should topics be grouped by theme? Default is <i>TRUE</i> .
by_date	Should topics be grouped by month of the year or just by year? Default is to group by month of the year.

Value

A tibble of topic views by theme and by specified date format arranged in descending order

Author(s)

Ernest Guevarra

Examples

```
library(magrittr)
ennet_topics %>% arrange_replies()
```

arrange_views	<i>Arrange topics based on number of views</i>
---------------	--

Description

Arrange topics based on number of views

Usage

```
arrange_views(
  topics,
  by_theme = TRUE,
  by_date = c("month_year", "year", "all")
)
```

Arguments

topics	A tibble of topics by theme from en-net forum produced through a call to <code>get_themes_topics()</code> .
by_theme	Logical. Should topics be grouped by theme? Default is <i>TRUE</i> .
by_date	Should topics be grouped by month of the year or just by year or overall? Default is to group by month of the year.

Value

A tibble of topic views by theme and by specified date format arranged in descending order

Author(s)

Ernest Guevarra

Examples

```
library(magrittr)
ennet_topics %>% arrange_views()
```

count_authors	<i>Count the number of topics by author, by theme and by date</i>
---------------	---

Description

Count the number of topics by author, by theme and by date

Usage

```
count_authors(  
  topics,  
  by_theme = TRUE,  
  by_date = c("month_year", "year", "all"),  
  .sort = TRUE  
)
```

Arguments

topics	A tibble of topics by theme from en-net forum produced through a call to get_themes_topics() .
by_theme	Logical. If <i>TRUE</i> (<i>default</i>), count by theme.
by_date	Should topics be counted by month of the year or just by year or total? Default is to count by month of the year.
.sort	Logical. Should output be sorted by count frequencies? Default is <i>TRUE</i> .

Value

A tibble of topic counts by author, by theme and by specified date format

Author(s)

Ernest Guevarra

Examples

```
library(magrittr)
ennet_topics %>% count_authors()
```

count_topics	<i>Count the number of topics by theme and by date</i>
--------------	--

Description

Count the number of topics by theme and by date

Usage

```
count_topics(topics, by_date = c("month_year", "year", "all"), .sort = TRUE)
```

Arguments

topics	A tibble of topics by theme from en-net forum produced through a call to get_themes_topics() .
by_date	Should topics be grouped by month of the year or just by year or overall? Default is to group by month of the year.
.sort	Logical. Should output be sorted by count frequencies? Default is <i>TRUE</i> .

Value

A tibble of topic counts by theme and by specified date format

Author(s)

Ernest Guevarra

Examples

```
library(magrittr)
ennet_topics %>% count_topics(by_date = "month_year")
```

count_topics_author	<i>Count number of questions/topics posted on en-net by author</i>
---------------------	--

Description

Count number of questions/topics posted on en-net by author

Usage

```
count_topics_author(topics = get_themes_topics(), .sort = TRUE)

count_topics_author_time(
  topics = get_themes_topics(),
  by_time = c("day", "week", "month", "year"),
  .sort = TRUE
)
```

Arguments

topics	A tibble of topics by theme, by author, and by posting date from en-net forum produced through a call to get_themes_topics()
.sort	Logical. Should output be sorted by count frequencies? Default is <i>TRUE</i>
by_time	Should topics be counted by day, by week, by month or by year? Default is to count by day.

Value

A tibble of topic counts by specified grouping

Author(s)

Ernest Guevarra

Examples

```
## Get counts of topics by author
count_topics_author(topics = ennet_topics)

## Get counts of authors by author and by time
count_topics_author_time(topics = ennet_topics)
```

count_topics_day

Count number of questions/topics posted on en-net by time

Description

Count number of questions/topics posted on en-net by time

Usage

```
count_topics_day(topics = get_themes_topics(), .sort = FALSE)
count_topics_week(topics = get_themes_topics(), .sort = FALSE)
count_topics_month(topics = get_themes_topics(), .sort = FALSE)
count_topics_year(topics = get_themes_topics(), .sort = FALSE)
```

Arguments

topics	A tibble of topics by theme, by author, and by posting date from en-net forum produced through a call to get_themes_topics()
.sort	Logical. Should output be sorted by count frequencies? Default is <i>FALSE</i>

Value

A tibble of topic counts by specified time grouping

Author(s)

Ernest Guevarra

Examples

```
## Get counts of topics by day
count_topics_day(topics = ennet_topics)

## Get counts of topics by week
count_topics_week(topics = ennet_topics)

## Get counts of topics by month
count_topics_month(topics = ennet_topics)

## Get counts of topics by year
count_topics_year(topics = ennet_topics)
```

count_topics_theme	<i>Count number of questions/topics posted on en-net</i>
--------------------	--

Description

Count number of questions/topics posted on en-net

Usage

```
count_topics_theme(topics = get_themes_topics(), .sort = TRUE)

count_topics_theme_time(
  topics = get_themes_topics(),
  by_time = c("day", "week", "month", "year"),
  .sort = TRUE
)
```

Arguments

topics	A tibble of topics by theme, by author, and by posting date from en-net forum produced through a call to get_themes_topics()
.sort	Logical. Should output be sorted by count frequencies? Default is <i>TRUE</i>
by_time	Should topics be counted by day, by week, by month or by year? Default is to count by day.

Value

A tibble of topic counts by specified grouping

Author(s)

Ernest Guevarra

Examples

```
## Get counts of topics by theme
count_topics_theme(topics = ennet_topics)

## Get counts of topics by theme and by time
count_topics_theme_time(topics = ennet_topics)
```

```
create_db_topics_dailies
```

Create daily topics datasets for the ennet_db

Description

Create daily topics datasets for the ennet_db

Usage

```
create_db_topics_dailies(hourlies)
```


Arguments

hourlies A tibble of topics data usually produced by using the `create_db_topics_hourlies()` function

Value

A tibble of specified topics dataset created from data in the `ennet_db`

Author(s)

Ernest Guevarra

Examples

```
themes <- ennet_themes$themes
x <- ennet_hourlies[ennet_hourlies$Theme == themes[3], ]
create_db_topics_dailies(hourlies = x)
```

create_db_topics_daily

Create daily topics dataset for the ennet_db

Description

Create daily topics dataset for the `ennet_db`

Usage

```
create_db_topics_daily(
  repo = "katilingban/ennet_db",
  branch = "main",
  .date = Sys.Date() - 1,
  fn = NULL
)
```

Arguments

repo A character value of the GitHub user and repository name combination identifying the GitHub location for `ennet_db`. Default is `katilingban/ennet_db`.

branch A character value for the branch name from which to retrieve data. Default is `main`.

.date A character value or vector of date/dates for which to create a topics dataset for the `ennet_db`

fn A character value or vector of filenames for hourly topics dataset found in `ennet_db`

Value

A tibble of daily topics dataset created from data in the ennet_db

Author(s)

Ernest Guevarra

Examples

```
##  
fn <- c("ennet_topics_2021-01-17_00:54:48.csv")  
  
create_db_topics_daily(.date = "2021-01-17", fn = fn)
```

create_db_topics_hourlies

Create hourly topics datasets for the ennet_db

Description

Create hourly topics datasets for the ennet_db

Usage

```
create_db_topics_hourlies(  
  repo = "katilingban/ennet_db",  
  branch = "main",  
  .date = Sys.Date()  
)
```

Arguments

repo	A character value of the GitHub user and repository name combination identifying the GitHub location for ennet_db. Default is katilingban/ennet_db.
branch	A character value for the branch name from which to retrieve data. Default is main.
.date	A character value or vector of date/dates for which to create a topics dataset for the ennet_db

Value

A tibble of specified topics dataset created from data in the ennet_db

Author(s)

Ernest Guevarra

Examples

```
create_db_topics_hourlies(.date = "2020-12-31")
```

```
create_db_topics_interactions
```

Create various topics interactions datasets for the ennet_db

Description

Create various topics interactions datasets for the ennet_db

Usage

```
create_db_topics_interactions(  
  dailies,  
  id = c("daily", "weekly", "monthly", "yearly")  
)
```

Arguments

dailies	A tibble of topics data usually produced by using the <code>create_db_topics_dailies()</code> function
id	A character value for data identifier. Possible choices are <i>daily</i> , <i>weekly</i> , <i>monthly</i> , or <i>yearly</i> .

Value

A tibble of specified topics dataset created from data in the ennet_db

Author(s)

Ernest Guevarra

Examples

```
themes <- ennet_themes$themes  
x <- ennet_dailies[ennet_dailies$Theme == themes[3], ]  
create_db_topics_interactions(dailies = x, id = "yearly")
```

```
create_db_topics_monthly
```

Create monthly topics dataset for the ennet_db

Description

Create monthly topics dataset for the ennet_db

Usage

```
create_db_topics_monthly(  
  repo = "katilingban/ennet_db",  
  branch = "main",  
  .date = Sys.Date() - 1  
)
```

Arguments

repo	A character value of the GitHub user and repository name combination identifying the GitHub location for ennet_db. Default is katilingban/ennet_db.
branch	A character value for the branch name from which to retrieve data. Default is main.
.date	A character value or vector of date/dates for which to create a topics dataset for the ennet_db

Value

A tibble of monthly topics dataset created from data in the ennet_db

Author(s)

Ernest Guevarra

Examples

```
create_db_topics_monthly(.date = "2021-01-03")
```

ennet-deprecated

Deprecated functions in ennet

Description

These functions still work but will be removed (defunct) in the next version of ennet.

Details

Function	Notes
<code>count_topics()</code>	Please use <code>count_topics_theme()</code> instead
<code>count_authors()</code>	Please use <code>count_topics_author()</code> instead

ennet_dailies	<i>Daily extracts of topics dataset from en-net online forum</i>
---------------	--

Description

Daily extracts of topics dataset from en-net online forum

Usage

`ennet_dailies`

Format

A tibble with 90 rows and 7 columns:

Variable	Description
Theme	Thematic areas in the en-net forum
Topic	Short description of the topic/question being discussed/raised
Author	Name of person who raised the topic/question
Posted	Date topic/question was posted on en-net forum
Link	URL of the topic/question being discussed/raised
Interaction	Type of interaction. Either <i>Views</i> or <i>Replies</i>
Extraction	Date and time when data was extracted
n	Number or count
Extraction Date	Date when data was extracted

Source

<https://www.en-net.org>

Examples

`ennet_dailies`

ennet_hourlies *Hourly extracts of topics dataset from en-net online forum*

Description

Hourly extracts of topics dataset from en-net online forum

Usage

ennet_hourlies

Format

A tibble with 643844 rows and 8 columns:

Variable	Description
Theme	Thematic areas in the en-net forum
Topic	Short description of the topic/question being discussed/raised
Author	Name of person who raised the topic/question
Posted	Date topic/question was posted on en-net forum
Link	URL of the topic/question being discussed/raised
Interaction	Type of interaction. Either <i>Views</i> or <i>Replies</i>
Extraction	Date and time when data was extracted
n	Number or count

Source

<https://www.en-net.org>

Examples

ennet_hourlies

ennet_themes *Themes from en-net forum retrieved on 17 January 2021*

Description

Themes from en-net forum retrieved on 17 January 2021

Usage

ennet_themes

Format

A tibble with 18 rows and 2 columns:

Variable	Description
themes	Thematic areas in the en-net forum
links	URL of the thematic area

Source

<https://www.en-net.org>

Examples

ennet_themes

ennet_topics	<i>Topics from en-net forum retrieved on 17 January 2021</i>
--------------	--

Description

Topics from en-net forum retrieved on 17 January 2021

Usage

ennet_topics

Format

A tibble with 3045 rows and 7 columns:

Variable	Description
Theme	Thematic areas in the en-net forum
Topic	Short description of the topic/question being discussed/raised
Views	Number of views of the topic/question being discussed/raised
Author	Name of person who raised the topic/question
Posted	Date topic/question was posted on en-net forum
Link	URL of the topic/question being discussed/raised
Replies	Number of replies to topic/question being discussed/raised

Details

Please note that this dataset is made available in the package primarily as a guide for the user and as testing data for the code. Users are advised not to use this dataset for actual analysis or reporting and instead make a call to `get_theme_topics()` or to `get_themes_topics()`. This is because some

topics and some author names have been converted to NA as they contain non-ASCII characters which are not allowed as a text encoding format for packaged data.

Source

<https://www.en-net.org>

Examples

```
ennet_topics
```

```
get_db_discussions      Retrieve data from ennet_db GitHub repository
```

Description

Retrieve data from ennet_db GitHub repository

Usage

```
get_db_discussions(repo = "katilingban/ennet_db", branch = "main")

get_db_topics(
  repo = "katilingban/ennet_db",
  branch = "main",
  id = c("daily", "weekly", "monthly", "yearly")
)
```

Arguments

repo	A character value for the GitHub user and repository name combination identifying the GitHub location for ennet_db. Default is katilingban/ennet_db.
branch	A character value for the branch name from which to retrieve data. Default is main.
id	A character value for data identifier. Possible choices are <i>daily</i> , <i>weekly</i> , <i>monthly</i> , or <i>yearly</i> .

Value

A tibble of the specified dataset

Author(s)

Ernest Guevarra

Examples

```
## Retrieve discussions dataset
get_db_discussions()

## Retrieve en-net topics yearly interactions dataset
get_db_topics(id = "yearly")
```

get_themes	<i>Get list of thematic areas from en-net.org</i>
------------	---

Description

Get list of thematic areas from en-net.org

Usage

```
get_themes(base = "https://www.en-net.org")
```

Arguments

base Base URL of the en-net site. Set to <https://www.en-net.org/>

Value

A tibble containing the thematic areas from **en-net forum** and the corresponding URLs for each thematic area

Examples

```
get_themes()
```

get_themes_topics	<i>Get topics from multiple themes</i>
-------------------	--

Description

Get topics from multiple themes

Usage

```
get_themes_topics(themes = get_themes())
```

Arguments

themes A tibble containing thematic areas and URLs for thematic area pages

Value

A tibble of all topics across all thematic areas with their respective URLs

Examples

```
themes <- ennet_themes
get_themes_topics(themes = themes[4, ])
```

get_theme_topics *Get theme topics*

Description

Get theme topics

Usage

```
get_theme_topics(link)
```

Arguments

link URL of a specific thematic area

Value

A tibble of all topics for the specified thematic area.

Examples

```
themes <- ennet_themes
get_theme_topics(link = themes$links[4])
```

`get_topics_discussions`*Get the discussion and details of discussion for a set of topics*

Description

Get the discussion and details of discussion for a set of topics

Usage

```
get_topics_discussions(links)
```

Arguments

`links` A tibble of topics containing URL of topic discussion. This is provided using a call to `get_theme_topics()` or `get_themes_topics()`

Value

A tibble containing the topic discussions for selected topic/s

Examples

```
links <- get_theme_topics(link = (ennet_themes$links)[4])
get_topics_discussions(link = links[1:3, ])
```

`get_topic_discussions` *Get the discussion and other details for a particular topic*

Description

Get the discussion and other details for a particular topic

Usage

```
get_topic_discussions(link)
```

Arguments

`link` URL for topic discussion

Value

A tibble containing the topic question

Examples

```
links <- get_theme_topics(link = (ennet_themes$links)[4])
get_topic_discussions(link = links$Link[1])
```

update_topics	<i>Update en-net topics</i>
---------------	-----------------------------

Description

Update en-net topics

Usage

```
update_topics(freq = 0)
```

Arguments

freq	A numeric value for time in seconds for frequency of retrieval of updated en-net topics. Defaults to 0 for a single retrieval (no repeats)
------	--

Value

A tibble containing the topics across all thematic areas in **en-net forum**

Examples

```
if(interactive()) update_topics()
```

Index

* datasets

- ennet_dailies, [13](#)
 - ennet_hourlies, [14](#)
 - ennet_themes, [14](#)
 - ennet_topics, [15](#)
- [arrange_replies](#), [2](#)
- [arrange_views](#), [3](#)
- [count_authors](#), [4](#)
- [count_authors\(\)](#), [13](#)
- [count_topics](#), [5](#)
- [count_topics\(\)](#), [13](#)
- [count_topics_author](#), [5](#)
- [count_topics_author\(\)](#), [13](#)
- [count_topics_author_time](#)
([count_topics_author](#)), [5](#)
- [count_topics_day](#), [6](#)
- [count_topics_month](#) ([count_topics_day](#)), [6](#)
- [count_topics_theme](#), [7](#)
- [count_topics_theme\(\)](#), [13](#)
- [count_topics_theme_time](#)
([count_topics_theme](#)), [7](#)
- [count_topics_week](#) ([count_topics_day](#)), [6](#)
- [count_topics_year](#) ([count_topics_day](#)), [6](#)
- [create_db_topics_dailies](#), [8](#)
- [create_db_topics_dailies\(\)](#), [11](#)
- [create_db_topics_daily](#), [9](#)
- [create_db_topics_hourlies](#), [10](#)
- [create_db_topics_hourlies\(\)](#), [9](#)
- [create_db_topics_interactions](#), [11](#)
- [create_db_topics_monthly](#), [12](#)
- [ennet-deprecated](#), [12](#)
- [ennet_dailies](#), [13](#)
- [ennet_hourlies](#), [14](#)
- [ennet_themes](#), [14](#)
- [ennet_topics](#), [15](#)
- [get_db_discussions](#), [16](#)
- [get_db_topics](#) ([get_db_discussions](#)), [16](#)
- [get_theme_topics](#), [18](#)
- [get_theme_topics\(\)](#), [15](#), [19](#)
- [get_themes](#), [17](#)
- [get_themes_topics](#), [17](#)
- [get_themes_topics\(\)](#), [3–8](#), [15](#), [19](#)
- [get_topic_discussions](#), [19](#)
- [get_topics_discussions](#), [19](#)
- [update_topics](#), [20](#)