Package 'SDGdetector'

May 24, 2023

Title Detect SDGs and Targets in Text

Version 2.7.3

Description Identify 17 Sustainable Development Goals and associated 169 targets in text.

URL https://github.com/Yingjie4Science/SDGdetector

BugReports https://github.com/Yingjie4Science/SDGdetector/issues

Imports dplyr, magrittr, stringr, ggplot2, tidyr, grDevices, rnaturalearth, utils, scales, magick

Depends R (>= 3.5.0)

License GPL (>= 3)

Encoding UTF-8

RoxygenNote 7.2.3

LazyData true

NeedsCompilation no

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

Author Yingjie Li [aut, cre] (<https://orcid.org/0000-0002-8401-0649>), Veronica Frans [aut] (<https://orcid.org/0000-0002-5634-3956>), Yongze Song [aut] (<https://orcid.org/0000-0003-3420-9622>), Meng Cai [aut] (<https://orcid.org/0000-0002-8318-572X>), Yuqian Zhang [aut] (<https://orcid.org/0000-0001-7576-2526>), Jianguo Liu [aut] (<https://orcid.org/0000-0001-6344-0087>)

Maintainer Yingjie Li <yingjieli.edu@gmail.com>

Repository CRAN

Date/Publication 2023-05-23 23:52:04 UTC

R topics documented:

add_sdg_pattern 2
codelist_panel 3

country_region_names . . . 3
detect_region . . . 4
func_AND_vector . . . 4
func_OR_vector . . . 5
list_of_un_goals_targets . . . 6
lookaround_nearby_n . . . 6
plot_sdg_bar . . . 7
plot_sdg_map . . . 8
SDGdetector . . . 8
sdgstat . . . 9
sdg_color . . . 10
sdg_icon . . . 10
sdg_icons . . . 11
SDG_keys . . . 11
shp . . . 12
summarize_sdg . . . 12

Index **14**

`add_sdg_pattern` *Users Can Add Customized Patterns for Each SDG or Target*

Description

Users Can Add Customized Patterns for Each SDG or Target

Usage

```
add_sdg_pattern(sdg_id, x, operator = "AND", quiet = FALSE)
```

Arguments

`sdg_id`	SDG Goal's ID or Target's ID, in the format of 'SDGx_y', e.g., SDG1_1, SDG2_general
`x`	A vector of strings
`operator`	'AND', 'OR' to combine a vector of keywords for identifying SDG Goals or Targets.
`quiet`	Logical. Suppress info message

Value

A regerx string

Examples

```
terms_new <- c("improve", "farmer", "income")
add_sdg_pattern(sdg_id = 'SDG1_2', x = terms_new, operator = 'AND')
```

codelist_panel *List of Names and ISO Code for Countries*

Description

List of Names and ISO Code for Countries

Usage

```
codelist_panel
```

Format

`codelist_panel`**:**
A data frame with 28941 rows and 55 columns:

country.name.en Country name in English

iso2c, iso3c 2 & 3 letter ISO country codes

year Year ...

Source

https://en.wikipedia.org/wiki/List_of_countries_and_territories_by_land_and_maritime_borders

country_region_names *Datasets of country and region names.*

Description

Datasets of country and region names.

Usage

```
country_region_names
```

Format

`country_region_names`: A data frame with 644 rows and 3 variables

Author(s)

Yingjie Li <yingjieli.edu@gmail.com>

detect_region *Detect country or region names in text for further mapping*

Description

Detect country or region names in text for further mapping.

Usage

```
detect_region(x, col)
```

Arguments

x	Data frame or a string
col	Column name for text to be assessed

Value

Returns the tool text outputs.

Examples

```
x <- c("This paper explores the method and results from an independent
evidence based assessment of Australia's progress towards the SDGs",
"Last year alone, the United States experienced 14 separate billion-dollar
 disasters related to climate change")
col <- data.frame(x)
regions <- detect_region(x, col)
```

func_AND_vector **Last update on***: 3/31/2022*

Description

New changes:

Usage

```
func_AND_vector(v)
```

Arguments

v	a vector of characters

Details

Compare to the earlier version, we made the following changes

1. Instead of combining multiple term lists by `OR` for one particular target, it is more intuitive and accurate to add each alternative term list to the search term table or database directly.
2. Added `Look around` function to more accurately match SDG targets.

Use `AND` to Concatenate a Vector of Terms

Value

A character

Examples

```
words <- c('apple', 'bean', 'food')
func_AND_vector(v= words)
```

`func_OR_vector` *Use* `OR` *to Concatenate a Vector of Terms*

Description

Use `OR` to Concatenate a Vector of Terms

Usage

```
func_OR_vector(v)
```

Arguments

v	a vector of characters

Value

A character

Examples

```
words <- c('apple', 'bean', 'food')
func_OR_vector(v= words)
```

list_of_un_goals_targets *The Names, ID, and Descriptions of all the 17 SDGs and 169 Targets*

Description

The Names, ID, and Descriptions of all the 17 SDGs and 169 Targets

Usage

```
list_of_un_goals_targets
```

Format

`list_of_un_goals_targets`**:**
A data frame with 169 rows and 3 columns:

GoalID The ID of each SDG

GoalName The name of each SDG

target_id_un The name of each Target

target_desc_un The description for each Target

Source

https://unstats.un.org/sdgs/indicators/indicators-list/

lookaround_nearby_n *Look Around*

Description

Look around to match pattern in a sentence

Usage

```
lookaround_nearby_n(word_ls1, word_ls2, n, exclude = "", third_AND_string = "")
```

Arguments

word_ls1	is a string, which includes a list of words connected by "|" that indicates 'OR'
word_ls2	is a string, which includes a list of words connected by "|" that indicates 'OR'
n	is a number, indicates the number of words to look around
exclude	is a vector, including a list of words to be excluded from match
third_AND_string	similar to word_ls1 or word_ls2, it is a string that includes a list of words connected by "|" that indicates 'OR'

Value

A regex string

Examples

```
con1 <- c('apple', 'bean', 'food')
con2 <- c('big', 'delicious')
lookaround_nearby_n(word_ls1 = con1, word_ls2 = con2, n = 2, exclude = "", third_AND_string = "")
```

`plot_sdg_bar` *SDG bar plot*

Description

SDG bar plot

Usage

```
plot_sdg_bar(data, sdg = "sdg", value = "value", quiet = FALSE)
```

Arguments

`data`	Data frame as the input
`sdg`	Vector with SDG code to be visualized.
`value`	The value, e.g., number of SDGs, to be show in the thematic map
`quiet`	Logical. Suppress info message

Value

Returns the tool text outputs.

Examples

```
data("sdgstat")
plot_sdg_bar(sdgstat, sdg = "SDG", value = "Value")
```

plot_sdg_map *SDG Map Plot*

Description

SDG map plot

Usage

```
plot_sdg_map(data, sdg = sdg, value = value,
                    country = country, by_sdg = TRUE)
```

Arguments

data	Data frame as the input
sdg	Vector with SDG code to be visualized.
value	The value, e.g., number of SDGs, to be show in the thematic map
country	Country that are associated with the SDGs.
by_sdg	If mapping by SDG, TRUE or FALSE.

Value

Returns the tool text outputs.

Examples

```
data("sdgstat")
plot_sdg_map(sdgstat,
  sdg = "SDG", value = "Value",
  country = "Country", by_sdg = FALSE
)
```

SDGdetector *Identify SDGs in text*

Description

Identify 17 Sustainable Development Goals and associated 169 targets in text.

Usage

```
SDGdetector(x, col, quiet = FALSE)
```

Arguments

x	Data frame or a string
col	Column name for text to be assessed
quiet	Logical. Suppress info message

Details

In 2015, leaders worldwide adopted 17 Sustainable Development Goals (SDGs) with 169 targets to be achieved by 2030 (https://sdgs.un.org). The framework of SDGs serves as a blueprint for shared prosperity for both people and the earth. `SDGdetector` identifies both direct and indirect expressions of SDGs and associated targets in chunks of text. It takes a data frame with a specified column of text to process as inputs and outputs a data frame with original columns plus matched SDGs and targets.

Value

Data frame with the same columns as the `df` plus one extra column named "sdgs", which list the occurrence (or hits) of SDG goals or targets detected from each sentence in rows. Users can further use our function `summarize_sdg()` to clean the result for visulization.

Examples

```
my_col <- c("our goal is to end poverty globally", "this product
contributes to slowing down climate change")
my_text <- data.frame(my_col)
SDGdetector(my_text, my_col)
```

---

sdgstat *Datasets of SDG statistics.*

Description

Datasets of SDG statistics.

Usage

```
sdgstat
```

Format

`sdgstat`: A data frame with 62 rows and 4 variables

Author(s)

Yingjie Li <yingjieli.edu@gmail.com>

sdg_color *Color scheme for the 17 SDGs*

Description

Color scheme for the 17 SDGs

Usage

```
sdg_color(x)
```

Arguments

x	A number, which indicates the SDG ID

Value

HTML color code of a specified SDG

Examples

```
sdg_color(1)
sdg_color(x = 1:17)
```

sdg_icon *Icons for SDGs*

Description

The sdg_icon function provides the specific icon for each SDG

Usage

```
sdg_icon(x, res = 200)
```

Arguments

x	Numeric code for each SDG, ranging from 1 to 17
res	Resolution of SDG icon. Default: res = 200 indicates resizing proportionally to 200px

Examples

```
sdg_icon(x = 17, res = 300)
```

sdg_icons *List SDG Icons*

Description

List SDG Icons

Usage

```
sdg_icons
```

Format

sdg_icons: External pointer of class "magick-image"

Source

https://www.un.org/sustainabledevelopment/wp-content/uploads/2019/01/SDG_Guidelines_AUG_2019_Final.pdf

SDG_keys *SDG_keys*

Description

Database of SDG search terms

Datasets of SDG keys.

Usage

```
data(SDG_keys)

SDG_keys
```

Format

An object of class data.frame with 557 rows and 3 columns.

SDG_keys: A data frame with 557 rows and 3 variables

Details

The search terms are developed at the "Target" level (SDG Goal/Target/Indicator) to extract SDG-related statements. These SDG search terms can be "direct mention", such as "SDG 1", or "indirect mention", which means a statement aligns with the description of certain SDGs or targets. For example, "Our company has embraced CO2 emissions mitigation as a priority within our sustainability strategy") is an indirect mention of "SDG 13.a" ("Implement the commitment... in the context of meaningful mitigation actions and ...").

Author(s)

Yingjie Li <yingjieli.edu@gmail.com>

Examples

```
data(SDG_keys)
```

---

shp *Datasets of shapefiles.*

Description

Datasets of shapefiles..

Usage

```
shp
```

Format

shp: A data frame with 241 rows and 6 variables

Author(s)

Yingjie Li <yingjieli.edu@gmail.com>

---

summarize_sdg *Summarize results from SDGdetector at either the Goal level or Target level.*

Description

Summarize results from SDGdetector at either the Goal level or Target level.

Usage

```
summarize_sdg(data, sum_by = "target", quiet = FALSE)
```

Arguments

data	Data frame or a string
sum_by	The group level to be chosen for data summary. Default parameter is "target", and can also set at "goal" level.
quiet	Logical. Suppress info message

Value

Data frame with at least one column named "SDG" or "Target", and one column `Freq` that represent the total hits.

Examples

```
library(SDGdetector)
df <- data.frame(col = c(
    'our goal is to end poverty globally',
    'this product contributes to slowing down climate change'))
data <- SDGdetector(x = df, col = col)
summarize_sdg(data, sum_by = 'target', quiet = FALSE)
```

Index

∗ **datasets**
codelist_panel, 3
list_of_un_goals_targets, 6
sdg_icons, 11
SDG_keys, 11
∗ **dataset**
country_region_names, 3
SDG_keys, 11
sdgstat, 9
shp, 12

add_sdg_pattern, 2

codelist_panel, 3
country_region_names, 3

detect_region, 4

func_AND_vector, 4
func_OR_vector, 5

list_of_un_goals_targets, 6
lookaround_nearby_n, 6

plot_sdg_bar, 7
plot_sdg_map, 8

sdg_color, 10
sdg_icon, 10
sdg_icons, 11
SDG_keys, 11
SDGdetector, 8
sdgstat, 9
shp, 12
summarize_sdg, 12