Package 'inlpubs'

February 27, 2024

Title USGS INL Project Office Publications

Version 1.1.1

Description Contains bibliographic information for the U.S. Geological Survey (USGS) Idaho National Laboratory (INL) Project Office.

Depends R (>= 4.1)

Imports checkmate, stats, tm

Suggests connectapi, covr, graphics, htmltools, htmlwidgets, jsonlite, kableExtra, knitr, magick, markdown, pkgbuild, pkgdown, pkgload, reactable, renv, rmarkdown, rsconnect, RWeka, stringi, textutils, tinytest, utils, webshot2, wordcloud2

License CC0

URL https://rconnect.usgs.gov/INLPO/inlpubs-main/, https://code.usgs.gov/inl/inlpubs

BugReports https://code.usgs.gov/inl/inlpubs/-/issues

Copyright This software is in the public domain because it contains materials that originally came from the United States Geological Survey (USGS), an agency of the United States Department of Interior. For more information, see the official USGS copyright policy at https://www.usgs.gov/information-policies-and-instructions/copyrights-and-credits

Encoding UTF-8

SystemRequirements Text mining using n-grams requires Amazon Corretto (https://aws.amazon.com/corretto/).

LazyData true

LazyDataCompression xz

RoxygenNote 7.3.1

NeedsCompilation no

Author Jason C. Fisher [aut, cre] (<https://orcid.org/0000-0001-9032-8912>), Kerri C. Treinen [aut] (<https://orcid.org/0000-0003-0645-6810>), Allison R. Trcka [aut] (<https://orcid.org/0000-0001-8498-4737>)

Maintainer Jason C. Fisher <jfisher@usgs.gov>

Repository CRAN

Date/Publication 2024-02-27 17:00:02 UTC

R topics documented:

`authors` *Contributing Authors to INLPO Publications*

Description

Authors who have contributed to the publications by the U.S. Geological Survey (USGS), Idaho Water Science Center, Idaho National Laboratory Project Office (INLPO).

Usage

```
authors
```

Format

An object of class 'author' that inherits behavior from the 'data.frame' class and includes the following columns:

`author_id` Unique identifier for the author.

`name` Name of author, surname first and initials or given name.

`person` Information about the person like email address and ORCiD identifier.

`pub_id` Identifier(s) of the publication(s) the author has contributed to, referes to the primry key of the pubs data table.

`total_pub` Total number of publications.

`single_authored` Number of single-authored publications.

`multi_authored` Number of multi-authored publications.

`first_authored` Number of multi-authored publications where the researcher appears as first author.

`first_year` First year author published.

`last_year` Last year author published.

Source

Curated by INLPO staff.

Examples

```
# Subset Jason Fisher's information and display structure:
author <- authors["jfisher", ]
str(author, max.level = 3, width = 75, strict.width = "cut")

# Print author's given name:
author$person |> format(include = "given")
```

`make_wordcloud` *Create Word Cloud from Frequency Table of Words*

Description

Create a word cloud from a frequency table of words, and save to a PNG file. Requires R-packages **htmltools**, **htmlwidgets**, **magick**, **webshot2**, and **wordcloud2** are available. System dependencies include the the following: ImageMagick for displaying the PNG image, OptiPNG for PNG file compression, and Chrome- or a Chromium-based browser with support for the Chrome DevTools protocol. Use `find_chromate` function to find the path to the Chrome browser.

Usage

```
make_wordcloud(
  x,
  max_words = 200L,
  size = 1,
  shape = "circle",
  ellipticity = 0.65,
  ...,
  width = 910L,
  output = NULL,
  display = FALSE
)
```

Arguments

`x`	'data.frame'. A frequency table of words that includes "word" and "freq" in each column.
`max_words`	'integer' number. Maximum number of words to include in the word cloud.
`size`	'numeric' number. Font size, where the larger size indicates a bigger word.
`shape`	'character' string. Shape of the "cloud" to draw. Possible shapes include a "circle", "cardioid", "diamond", "triangle-forward", "triangle", "pentagon", and "star".

ellipticity	'numeric' number. Degree of "flatness" of the shape to draw, a value between 0 and 1.
...	Additional arguments to be passed to the wordcloud2 function.
width	'integer' number. Desired image width in pixels.
output	'character' string. Path to the output file, by default the word cloud is copied to a temporary file.
display	'logical' flag. Whether to display the saved PNG file in a graphics window. Requires access to the **magick** package.

Value

File path to the word cloud plot in PNG format.

Author(s)

J.C. Fisher, U.S. Geological Survey, Idaho Water Science Center

Examples

```
## Not run:
  file <- wordcloud2::demoFreq |>
    make_wordcloud(size = 1.5, display = interactive())

  unlink(file)

## End(Not run)
```

---

mine_text *Mine Text Components in the INLPO Publications*

Description

Performs a word frequency text analysis of Idaho National Laboratory Project Office (INLPO) publications.

Usage

```
mine_text(
  pubs,
  components = c("title", "abstract"),
  ngmin = 1L,
  ngmax = ngmin,
  lowfreq = 1L
)
```

Arguments

pubs	'pub' class. Bibliographic information, see pubs dataset for details.
components	character vector. One or more text components to analyze. Choices include the "title", "abstract", "annotation", and "bibentry" of the document.
ngmin, ngmax	integer number. Splits strings into *n-grams* with given minimal and maximal numbers of grams. An n-gram is an ordered sequence of n words taken from the body of a text. Requires the **RWeka** package is available and that the environment variable JAVA_HOME points to where the Java software is located. Recommended for single text compoents only.
lowfreq	integer number. Lower frequency bound. Words that occur less than this bound are excluded from the returned frequency table.

Details

HTML entities are decoded when the **textutils** package is available.

Value

A word frequency table giving the number of times each word occurs in a publication's text component(s). A table column represents a single publication that is identified using its bibentry-key. And each row provides frequency counts for a particular word (also known as a 'term').

Author(s)

J.C. Fisher, U.S. Geological Survey, Idaho Water Science Center

See Also

make_wordcloud function to create a word cloud.

Examples

```
m <- head(pubs, 3) |> mine_text()
head(m)

## Not run:
  d <- data.frame(word = rownames(m), freq = rowSums(m))
  file <- make_wordcloud(d, display = interactive())
  unlink(file)

## End(Not run)
```

pubs *Publications of the INLPO*

Description

Bibliographic information for reports, articles, maps, and theses related to scientific monitoring and research conducted by the U.S. Geological Survey (USGS), Idaho Water Science Center, Idaho National Laboratory Project Office (INLPO).

Usage

```
pubs
```

Format

An object of class 'pub' that inherits behavior from the 'data.frame' class and includes the following columns:

`pub_id` Unique identifier for the publication.

`institution` Name of the institution that published and/or sponsored the report.

`type` Type of publication.

`text_ref` Text reference (also known as the in-text citation) that excludes the year of publication.

`year` Year of publication.

`author_id` Identifier(s) of the author(s), referes to the primry key of the `authors` data table.

`title` Title of publication.

`bibentry` Bibliographic entry of class `bibentry`.

`abstract` Abstract of publication.

`annotation` Annotation of publication.

`annotation_src` Identifier for the annotation source publication (Knobel and others, 2005; Bartholomay, 2022).

Source

Many of these publications are available through the USGS Publications Warehouse.

References

Bartholomay, R.C., 2022, Historical development of the U.S. Geological Survey hydrological monitoring and investigative programs at the Idaho National Laboratory, Idaho, 2002-2020: U.S. Geological Survey Open-File Report 2022-1027 (DOE/ID-22256), 54 p., doi:10.3133/ofr20221027.

Knobel, L.L., Bartholomay, R.C., and Rousseau, J.P., 2005, Historical development of the U.S. Geological Survey hydrologic monitoring and investigative programs at the Idaho National Engineering and Environmental Laboratory, Idaho, 1949 to 2001: U.S. Geological Survey Open-File Report 2005–1223 (DOE/ID–22195), 93 p., doi:10.3133/ofr20051223.

Examples

```
# Subset Fisher and others (2012) and display structure:
id <- "FisherOthers2012"
pub <- pubs[id, ]
str(pub, max.level = 3, width = 75, strict.width = "cut")

# Print suggested citation:
attr(unclass(pub$bibentry[[1]])[[1]], which = "textVersion")

# Print authors full name:
format(pub$bibentry[[1]]$author, include = c("given", "family"))

# Print abstract:
pub$abstract
```

Index