Package 'sorvi'

October 14, 2022

Type Package

Title Functions for Finnish Open Data

Version 0.8.20

Date 2022-05-27

MailingList rOpenGov <ropengov-forum@googlegroups.com>

Description Misc support functions for rOpenGov and open data downloads.

License BSD_2_clause + file LICENSE

VignetteBuilder knitr

BugReports <https://github.com/ropengov/sorvi/issues>

URL <https://github.com/ropengov/sorvi>,

<https://CRAN.R-project.org/package=sorvi>,

<https://ropengov.github.io/sorvi/>

Depends R (>= 3.3.0)

Imports dlstats, dplyr, ggplot2, gh, tidyr, purrr, rlang, utils,
rvest, xml2, lubridate, checkmate

Suggests gridExtra, RColorBrewer, knitr, rmarkdown, Cairo

Encoding UTF-8

RoxygenNote 7.2.0

NeedsCompilation no

Author Leo Lahti [aut, cre] (<<https://orcid.org/0000-0001-5537-637X>>),
Juuso Parkkinen [aut],
Joona Lehtomaki [aut],
Pyry Kantanen [ctb] (<<https://orcid.org/0000-0003-2853-2765>>)

Maintainer Leo Lahti <leo.lahti@iki.fi>

Repository CRAN

Date/Publication 2022-05-30 10:00:02 UTC

R topics documented:

---

sorvi-package *Algorithmic Tools for Open Data in Finland*

Description

The sorvi package hosts various functions that are mainly helpful in rOpenGov package mainte-
nance, package authoring and drawing graphs for presentations. Additionally it has some functions
that do not (yet) have their own package but are useful in some contexts.

Details

Package:	sorvi
Type:	Package
Version:	0.8.18
Date:	2010-2022
License:	BSD 2-clause
LazyLoad:	yes

Author(s)

Leo Lahti, Juuso Parkkinen, Jussi Paananen, Joona Lehtomaki, Einari Happonen, Juuso Haapanen,
and Pyry Kantanen <louhos@googlegroups.com>

References

See citation("sorvi") https://github.com/rOpenGov/sorvi

Examples

```
library(sorvi)
```

---

cran_downloads | *Get CRAN download statistics*

Description

Produces a tibble or a visualization of package download statistics.

Usage

```
cran_downloads(
  pkgs = "all",
  output = "tibble",
  sum = "by_month",
  plot.scale = 11,
  use.cache = TRUE
)
```

Arguments

pkgs	Package name(s). Default is "all", which prints statistics for all rOpenGov packages. You can also input 1 or more package names as a vector.
output	"tibble" (default) or "plot". With sum "by_month" and "by_year" "plot" outputs a line chart, with "total" it outputs a bar chart.
sum	"by_month" (default), "by_year" or "total"
plot.scale	integer, default is 11. Smaller numbers decrease the size of plot elements, larger numbers make them larger.
use.cache	Cache downloaded statistics. Default is TRUE

Details

This function is intended for easy retrieval and visualization of rOpenGov package download statistics from CRAN. It is an evolution of an R script by antagomir. As such it retains some features that were present in the original R script and were deemed useful for rOpenGov's internal use. This function may or may not be useful in other instances.

Value

tibble or a ggplot2 line chart or a bar chart

Author(s)

Leo Lahti, Pyry Kantanen <pyry.kantanen@gmail.com>

Examples

```
## Not run:
df <- cran_downloads(pkgs = "eurostat", sum = "total", use.cache = FALSE)
kable(df)

## Compare two packages
p1 <- cran_downloads(pkgs = "eurostat", sum = "by_year", output = "plot")
p2 <- cran_downloads(pkgs = "osmar", sum = "by_year", output = "plot")
gridExtra::grid.arrange(p1, p2, nrow = 2)

## End(Not run)
```

---

get_ifpi_charts   *Get IFPI Finland music consumption statistics*

Description

Download chart position data from ifpi.fi

Usage

```
get_ifpi_charts(channel = "radio", year = NA, week = NA)
```

Arguments

channel	Options: "radio", "albumit", "singlet", "fyysiset-albumit"
year	year as numeric. Default is NA, returning charts from current year. Charts are available from 2014 onwards.
week	week as numeric. Default is NA, returning most last possible charts. Week cannot be the current week. Please note that number of weeks differ between years. For simplicity's sake valid weeks are set to be between 1 and 53. Use e.g. 'lubridate::isoweek' to check how many weeks a given year has.

Details

Web scraping function that is inspired by Sauravkaushik8 Kaushik's blog post "Beginner's Guide on Web Scraping in R" on analyticsvidhya.com. Downloads chart data from Musiikkituottajat - IFPI Finland ry website. Please note that this function works only with IFPI Finland website!

The output has the following columns:

- rank: Rank on chart
- artist: Artist name
- song_title: Song title
- rank_last_week: Rank on chart on the previous week. RE if the song has re-entered the chart
- chart_woc: Weeks on chart
- week: Week number of observation
- year: Year of observation

Value

tibble

Author(s)

Pyry Kantanen <pyry.kantanen@gmail.com>

See Also

Original tutorial in https://www.analyticsvidhya.com/blog/2017/03/beginners-guide-on-web-scraping-in-r-us

---

gh_issue_stats *GitHub issues statistics*

Description

Get statistics about GitHub issues from GitHub API.

Usage

```
gh_issue_stats(
  owner = "ropengov",
  repo = "geofi",
  issue.type = NA,
  time.from = NA,
  time.to = NA
)
```

Arguments

owner	Repository owner / organization. Default is "ropengov"
repo	Repository name. Default is "geofi"
issue.type	Type of issues printed: "issue", "PR" or NA printing all (default).
time.from	Start date in ISO 8601 format: YYYY-MM-DDTHH:MM:SSZ. Default is "2010-09-01T00:00:00Z".
time.to	End date in ISO 8601 format: YYYY-MM-DDTHH:MM:SSZ. Default is Sys.time().

Details

This function is intended for easy information retrieval about rOpenGov package issues and pull requests. More specifically, this function returns a tibble containing information on issue id, title, status (open or closed), number of comments, who opened it, when it was created, what was the openers status (rOpenGov organization member, package contributor or a regular user who opened e.g. a bug issue) and what is the type of the issue.

GitHub Issues API handles Pull Requests and Issues similarly and therefore this function returns both types by default. Different types of issues can be filtered by using the issue.type parameter.

Kudos for this function go to Jennifer Bryan. The changes made here are mostly related to adding additional fields (opener_type, issue_type) to the output tibble and writing a function around these original contributions. The scope of this function is to mainly help rOpenGov team analyze the type of user feedback we get via GitHub issues and therefore the scope of this function is very limited.

Value

tibble

Author(s)

Original scripts by Jennifer Bryan (jennybc), function by Pyry Kantanen <pyry.kantanen@gmail.com>

See Also

GitHub Issues API documentation: <https://docs.github.com/en/rest/reference/issues>

Original "analyze GitHub stuff with R" repository: <https://github.com/jennybc/analyze-github-stuff-with-r>

---

load_sorvi_data   *Supporting Data*

Description

Load custom data sets.

Usage

```
load_sorvi_data(data.id, verbose = TRUE)
```

Arguments

data.id   data ID to download (see details)

verbose   verbose

Details

The following data sets are available:

  • translation_provincesTranslation of Finnish province (maakunta) names (Finnish, English).

Value

Data set. The format depends on the data.

Author(s)

Leo Lahti <leo.lahti@iki.fi>

References

See citation("sorvi")

Examples

```
translations <- load_sorvi_data("translation_provinces")
```

Index