Package 'TNRS'

January 11, 2023

Title Taxonomic Name Resolution Service

Version 0.3.3

Description Provides access to the Taxonomic Name Resolution Service <`https://github.com/ojalaquellueva/tnrsapi`> through R. The user supplies plant taxonomic names and the package returns resolved taxonomic names along with information on decisions. Optionally, the package can also be used to parse taxonomic names.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

Imports jsonlite, httr

Suggests knitr, rmarkdown, testthat, vcr (>= 0.6.0), devtools

VignetteBuilder knitr

RoxygenNote 7.2.2

Depends R (>= 3.5.0)

NeedsCompilation no

Author Brian Maitner [aut, cre] (<`https://orcid.org/0000-0002-2118-9880`>),
Brad Boyle [aut],
Paul Efren [ctb]

Maintainer Brian Maitner <`bmaitner@gmail.com`>

Repository CRAN

Date/Publication 2023-01-11 13:10:02 UTC

R topics documented:

`TNRS` *Resolve plant taxonomic names*

Description

Resolve plant taxonomic names.

Usage

```
TNRS(
  taxonomic_names,
  sources = c("tropicos", "wcvp"),
  classification = "tropicos",
  mode = "resolve",
  matches = "best",
  accuracy = NULL,
  skip_internet_check = FALSE,
  ...
)
```

Arguments

`taxonomic_names`	Data.frame containing two columns: 1) Row number, 2) Taxonomic names to be resolved (or parsed). Alternatively, a character vector of names can be supplied.
`sources`	Character vector. Taxonomic sources to use. Default is c("tropicos","wcvp"). Options include "tropicos", "usda", "wfo", and "wcvp".
`classification`	Character. Family classification to use. Currently options include "tropicos" (the default), which is equivalent to APGIII, and "wfo".
`mode`	Character. Options are "resolve" and "parse". Default option is "resolve"
`matches`	Character. Should all matches be returned ("all") or only the best match ("best", the default)?
`accuracy`	numeric. If specified, only matches with a score greater than or equal to the supplied accuracy level will be returned. If left NULL, the default threshold will be used.
`skip_internet_check`	Should the check for internet connectivity be skipped? Default is FALSE.
`...`	Additional parameters passed to internal functions

Value

Dataframe containing TNRS results.

Note

usda = United States Department of Agriculture, wfo = World Flora Online, wcvp = World Checklist of Vascular Plants.

For queries of more than 5000 names, the function will automatically divide the query into batches of 5000 names and then run the batches one after the other. Thus, for very large queries this may take some time. When this is the case, a progress bar will be displayed.

IMPORTANT: Note that parallelization of queries is automatically handled by the API, and so there is no need to further parallelize in R (in fact, doing so may actually slow things down!).

Examples

```
## Not run:
# Take a subset of the testfile to speed up runtime
tnrs_testfile <- tnrs_testfile[1:20, ]

results <- TNRS(taxonomic_names = tnrs_testfile)

# Inspect the results
head(results, 10)

## End(Not run)
```

`TNRS_base` *Resolve a small batch of plant taxonomic names*

Description

Resolve a small batch of plant taxonomic names

Usage

```
TNRS_base(
  taxonomic_names,
  sources = c("tropicos", "wcvp"),
  classification = "tropicos",
  mode = "resolve",
  matches = "best",
  accuracy = NULL,
  skip_internet_check = FALSE,
  ...
)
```

Arguments

`taxonomic_names`	Data.frame containing two columns: 1) Row number, 2) Taxonomic names to be resolved (or parsed). Alternatively, a character vector of names can be supplied.
`sources`	Character. Taxonomic sources to use. Default is c("tropicos","wcvp"). Options include "tropicos", "usda", "wfo", and "wcvp".
`classification`	Character. Family classification to use. Currently options include "tropicos" (the default), which is equivalent to APGIII, and "wfo".
`mode`	Character. Options are "resolve" and "parse". Default option is "resolve"
`matches`	Character. Should all matches be returned ("all") or only the best match ("best", the default)?
`accuracy`	numeric. If specified, only matches with a score greater than or equal to the supplied accuracy level will be returned.
`skip_internet_check`	Should the check for internet connectivity be skipped? Default is FALSE.
`...`	Additional parameters passed to internal functions

Value

Dataframe containing TNRS results.

Note

This function is primarily used as an internal function of TNRS and can only handle relatively small batches of names.

usda = United States Department of Agriculture, wfo = World Flora Online, wcvp = World Checklist of Vascular Plants.

`TNRS_citations` *Get citation information*

Description

Returns information needed to cite the TNRS

Usage

```
TNRS_citations(skip_internet_check = FALSE, ...)
```

Arguments

`skip_internet_check`	Should the check for internet connectivity be skipped? Default is FALSE.
`...`	Additional parameters passed to internal functions

Value

Dataframe containing bibtex-formatted citation information

Note

This function provides citation information in bibtex format that can be used with reference manager software (e.g. Paperpile, Zotero). Please do remember to cite both the sources and the TNRS, as the TNRS couldn't exist without these sources!

Examples

```
{
  citation_info <- TNRS_citations()
}
```

`TNRS_metadata` *Get TNRS metadata*

Description

Returns metadata on TNRS including version and citation information

Usage

```
TNRS_metadata(bibtex_file = NULL, skip_internet_check = FALSE)
```

Arguments

`bibtex_file`	Optional output file for writing bibtex citations.
`skip_internet_check`	Should the check for internet connectivity be skipped? Default is FALSE.

Value

List containing: (1) bibtex-formatted citation information, (2) information about TNRS data sources, and (3) TNRS version information.

Note

This function provides citation information in bibtex format that can be used with reference manager software (e.g. Paperpile, Zotero). Please remember to cite both the sources and the TNRS, as the TNRS couldn't exist without these sources!

This function is a wrapper that returns the output of the functions TNRS_citations, TNRS_sources, and TNRS_version.

Examples

```
{
  metadata <- TNRS_metadata()
}
```

`TNRS_sources` *Get information on sources used by the TNRS*

Description

Return metadata about the current TNRS sources

Usage

```
TNRS_sources(skip_internet_check = FALSE, ...)
```

Arguments

`skip_internet_check`	Should the check for internet connectivity be skipped? Default is FALSE.
`...`	Additional parameters passed to internal functions

Value

Dataframe containing information about the sources used in the current TNRS version.

Examples

```
{
  sources <- TNRS_sources()
}
```

`tnrs_testfile` *100 scientific names.*

Description

A dataset containing scientific names for 100 taxa. Names vary in accuracy and correctness.

Usage

```
tnrs_testfile
```

Format

A data frame with 100 rows and 2 variables:

ID Unique integer identifying each row

taxon Scientific name, possibly containing errors ...

Source

`https://github.com/ojalaquellueva/TNRSapi`

`TNRS_version` *Get metadata on current TNRS version*

Description

Return metadata about the current TNRS version

Usage

```
TNRS_version(skip_internet_check = FALSE, ...)
```

Arguments

`skip_internet_check`	Should the check for internet connectivity be skipped? Default is FALSE.
`...`	Additional parameters passed to internal functions

Value

Dataframe containing current TNRS version number, build date, and code version.

Examples

```
{
  TNRS_version_metadata <- TNRS_version()
}
```

Index