

Package ‘taxotools’

March 8, 2020

Type Package

Title Tools to Handle Taxonomic Lists

Version 0.0.43

Date 2020-03-08

Maintainer Vijay Barve <vijay.barve@gmail.com>

Description Some tools to work with taxonomic name lists.

License CC0

Imports taxize, wikitaxa, plyr, sqldf

Encoding UTF-8

LazyData true

RoxygenNote 7.0.2

BugReports <https://github.com/vijaybarve/taxotools/issues>

NeedsCompilation no

Author Vijay Barve [aut, cre] (<<https://orcid.org/0000-0002-4852-2567>>)

Repository CRAN

Date/Publication 2020-03-08 15:30:03 UTC

R topics documented:

cast_canonical	2
cast_cs_field	3
check_scientific	4
expand_name	4
get_itis_syn	5
guess_taxo_level	6
list_higher_taxo	6
list_itis_syn	8
list_wiki_syn	8
match_lists	9
melt_canonical	10
melt_cs_field	11
taxotools	12

cast_canonical	<i>Construct canonical names</i>
----------------	----------------------------------

Description

Construct canonical names using Genus, Species and Subspecies fields. At times due to spaces or NAs in the data fields, it makes it tricky to generate canonical names.

Usage

```
cast_canonical(dat, genus = "", species = "", subspecies = "")
```

Arguments

dat	data frame containing taxonomic list
genus	field name for Genus field
species	field name for Species field
subspecies	field name for Subspecies field

Value

a data frame containing Canonical names field added or repopulated using filed names for Genus, Species and Subspecies specified in parameters

See Also

Other Name functions: [check_scientific\(\)](#), [expand_name\(\)](#), [guess_taxo_level\(\)](#), [list_higher_taxo\(\)](#), [melt_canonical\(\)](#)

Examples

```
## Not run:  
cast_canonical(mylist,"genus","species","subspecies")  
  
## End(Not run)
```

cast_cs_field	<i>Build a character (comma) separated List within field</i>
---------------	--

Description

Builds a character (comma) separated list within a field given a data frame with primary field repeating values and secondary field with values to be character separated in the same field (secondary)

Usage

```
cast_cs_field(data, pri, sec, duplicate = FALSE, sepchar = ",")
```

Arguments

data	data frame containing primary and secondary data columns
pri	Primary field name (repeating values)
sec	Secondary field (values would be added to same record, comma separated)
duplicate	If true, duplicate entries are allowed in secondary field
sepchar	Character separator between the data items. Default is comma

Value

a data frame with two fields Primary and secondary (comma separated list)

See Also

Other List functions: [melt_cs_field\(\)](#)

Examples

```
## Not run:  
scnames <- c("Abrothrix longipilis", "Abrothrix jelskii")  
SynList <- list_itis_syn(scnames)  
cast_cs_field(SynList,"Name","Syn")  
  
## End(Not run)
```

check_scientific	<i>Parse and resolve a scientific name string</i>
------------------	---

Description

Parse the name using GNR and GBIF parse API to make sure the name is scientific name

Usage

```
check_scientific(name)
```

Arguments

name	scientific name string to be checked
------	--------------------------------------

Value

Resolved canonical name (NULL if not matched)

See Also

Other Name functions: [cast_canonical\(\)](#), [expand_name\(\)](#), [guess_taxo_level\(\)](#), [list_higher_taxo\(\)](#), [melt_canonical\(\)](#)

Examples

```
check_scientific("Akodon longipilis (Waterhouse, 1837)")
check_scientific("Mus longipilis Waterhouse, 1837")
check_scientific("Akodon hershkovitzi Patterson, Gallardo, and Freas, 1984")
```

expand_name	<i>Expands Scientific name</i>
-------------	--------------------------------

Description

At times the genus is specified with first character and '.' rather than repeating genus names every time. These are either synonyms or species of the same genus listed one below another. To convert these names to canonical names, we need to expand the genus name (typically) using previous entry in the list.

Usage

```
expand_name(fullname, shortname)
```

Arguments

fullname full scientific name
shortname scientific name with short form genus name to expand the Genus

Value

scientific name with Genus expanded using reference name provided as parameter

See Also

Other Name functions: [cast_canonical\(\)](#), [check_scientific\(\)](#), [guess_taxo_level\(\)](#), [list_higher_taxo\(\)](#), [melt_canonical\(\)](#)

Examples

```
expand_name("Addax gibbosa", "A. mytilopes")  
expand_name("Oryx addax", "O. nasomaculatus")
```

get_itis_syn *Get ITIS Synonyms for a Scientific Name*

Description

Fetch Synonyms using ITIS web service

Usage

```
get_itis_syn(scname)
```

Arguments

scname Scientific Name

Value

a list containing synonyms

See Also

Other ITIS functions: [list_itis_syn\(\)](#)

Examples

```
## Not run:  
get_itis_syn("Abrothrix longipilis")  
get_itis_syn("Abditomys latidens")  
  
## End(Not run)
```

guess_taxo_level *Guess the taxonomic level of Scientific Name*

Description

Guesses the taxonomic level i.e. Genus, Species or Subspecies based on number of words

Usage

```
guess_taxo_level(name)
```

Arguments

name scientific name string to be checked

See Also

Other Name functions: [cast_canonical\(\)](#), [check_scientific\(\)](#), [expand_name\(\)](#), [list_higher_taxo\(\)](#), [melt_canonical\(\)](#)

Examples

```
guess_taxo_level("Akodon longipilis")
guess_taxo_level("Akodon")
guess_taxo_level("Abrocoma cinerea shistacea")
```

list_higher_taxo *Get higher taxonomy data for list of names*

Description

Retrieve higher taxonomy information (like Family and Order) for each record from the "Encyclopedia of Life" web API.

Usage

```
list_higher_taxo(
  indf,
  canonical,
  genus = FALSE,
  verbose = FALSE,
  progress = TRUE
)
```

Arguments

indf	input data frame containing taxonomic list
canonical	field name containing scientific names
genus	If TRUE, use only genus level data to get taxonomy
verbose	If TRUE, displays each name string for which the higher taxonomy is sought
progress	If TRUE prints progress bar and messages on the console.

Details

This function makes use of certain functions in the [taxize](#) package. It scans and retrieves the taxonomic hierarchy for each scientific name (or just genus name) in the data set. When new data are retrieved, they are stored in a local sqlite database, `taxo.db`, for faster further access.

Value

data frame with added / updated columns

- "Kingdom" Kingdom of the Scientific name
- "Phylum" Phylum of the Scientific name
- "Order_" Order of the Scientific name
- "Family" Family of the Scientific name
- "Genus" Genus of the Scientific name

and also saves a local copy of taxonomy downloaded for future use in 'taxo.db' sqlite file

See Also

Other Name functions: [cast_canonical\(\)](#), [check_scientific\(\)](#), [expand_name\(\)](#), [guess_taxo_level\(\)](#), [melt_canonical\(\)](#)

Examples

```
## Not run:
mylist <- data.frame("canonical" = c("Abrothrix longipilis",
                                   "Mus longipilis",
                                   "Abrothrix jelskii",
                                   "Cardinalis cardinalis",
                                   "Danaus plexippus"),
                    stringsAsFactors = F)

my_taxo_list <- list_higher_taxo(mylist, "canonical")

## End(Not run)
```

list_itis_syn *Get ITIS Synonyms for list of names*

Description

Fetch Synonyms from ITIS

Usage

```
list_itis_syn(namelist)
```

Arguments

namelist list of scientific names

Value

a data frame containing names (passed) and synonyms

See Also

Other ITIS functions: [get_itis_syn\(\)](#)

Examples

```
## Not run:  
list_itis_syn("Abrothrix longipilis")  
#list_itis_syn(c("Abditomys latidens", "Abeomelomys sevia", "Abrothrix jelskii" ))  
  
## End(Not run)
```

list_wiki_syn *Get Wikipedia Synonyms for list of names*

Description

Fetch Synonyms from Wikipedia and clean them for use

Usage

```
list_wiki_syn(namelist)
```

Arguments

namelist list of scientific names

Value

a data frame containing names, synonyms and Canonical synonyms matched with GBIF backbone taxonomy

- Name : Scientific name
- WikiName : Wikipedia page name
- OrigSyn : Original synonym returned by Wikipedia
- Syn : Synonym in canonical form, matched with GBIF

Examples

```
list_wiki_syn("Abrothrix illutea")
#list_wiki_syn(c("Abditomys latidens", "Abeomelomys sevia", "Abrocoma schistacea"))
```

match_lists	<i>match two taxonomic lists</i>
-------------	----------------------------------

Description

match two taxonomic lists using canonical names

Usage

```
match_lists(master, checklist, masterfld, checklistfld)
```

Arguments

master	master taxonomic list
checklist	match taxonomic list
masterfld	field name for canonical name in master list
checklistfld	field name for canonical name in match list

Value

a list with data frames containing matched records, records only in master and checklist and statistics about the records including Jaccard index

Examples

```
## Not run:
master <- data.frame("canonical" = c("Abrothrix longipilis",
                                     "Acodon hirtus",
                                     "Akodon longipilis apta",
                                     "Akodon longipilis castaneus",
                                     "Chroeomys jelskii",
                                     "Acodon jelskii pyrrhotis"),
                    stringsAsFactors = F)
checklist <- data.frame("canonical" = c("Abrothrix longipilis",
                                       "Akodon longipilis apta",
                                       "Akodon longipilis castaneus",
                                       "Abrothrix jelskii",
                                       "Acodon jelskii pyrrhotis"),
                       stringsAsFactors = F)
match_lists(master, checklist, "canonical", "canonical")

## End(Not run)
```

melt_canonical

Deconstruct canonical names

Description

Deconstruct canonical names into Genus, Species and Subspecies fields.

Usage

```
melt_canonical(dat, canonical = "", genus = "", species = "", subspecies = "")
```

Arguments

dat	data frame containing taxonomic list
canonical	field name for canonical names
genus	field name for Genus
species	field name for Species
subspecies	field name for Subspecies

Value

a data frame containing Genus, Species and Subspecies fields added or repopulated using data in canonical name field.

See Also

Other Name functions: [cast_canonical\(\)](#), [check_scientific\(\)](#), [expand_name\(\)](#), [guess_taxo_level\(\)](#), [list_higher_taxo\(\)](#)

Examples

```
## Not run:
mylist <- data.frame("canonical" = c("Abrothrix longipilis",
                                     "Acodon hirtus",
                                     "Akodon longipilis apta",
                                     "AKODON LONGIPILIS CASTANEUS",
                                     "Chroeomys jelskii",
                                     "Acodon jelskii pyrrhotis"),
                    stringsAsFactors = F)
melt_canonical(mylist,"canonical","genus","species","subspecies")

## End(Not run)
```

melt_cs_field	<i>Generate a list melting character (comma) separated field values into multiple records</i>
---------------	---

Description

Builds a list, melting character (comma) separated field values given a data frame with a field with repeating values

Usage

```
melt_cs_field(data, melt, sepchar = ",")
```

Arguments

data	data frame containing a data columns with character(comma) separated values
melt	Field name with character(comma) separated values
sepchar	Character separator between the data items. Default is comma

Value

a data frame with separate records for each value in field specified

See Also

Other List functions: [cast_cs_field\(\)](#)

Examples

```
## Not run:
scnames <- c("Abrothrix longipilis", "Abrothrix jelskii")
syn_list <- list_itis_syn(scnames)
cs_syn_list <- cast_cs_field(syn_list , "Name", "Syn")
syn_list_new <- melt_cs_field(cs_syn_list, "Syn")

## End(Not run)
```

taxotools

taxotools: Tools to handle taxonomic data

Description

Some tools to work with lists taxonomic names.

List functions

- [cast_cs_field](#)
- [melt_cs_field](#)
- [match_lists](#)

Name functions

- [expand_name](#)
- [check_scientific](#)
- [guess_taxo_level](#)
- [cast_canonical](#)
- [melt_canonical](#)
- [list_higher_taxo](#)

ITIS functions

- [get_itis_syn](#)
- [list_itis_syn](#)

Wiki functions

- [list_wiki_syn](#)

Citation

- Barve, V., (2019). taxotools: Tools to handle taxonomic data (R package V 0.0.39). Retrieved from <https://cran.r-project.org/web/packages/taxotools/index.html>

Index

cast_canonical, [2](#), [4-7](#), [10](#), [12](#)
cast_cs_field, [3](#), [11](#), [12](#)
check_scientific, [2](#), [4](#), [5-7](#), [10](#), [12](#)

expand_name, [2](#), [4](#), [4](#), [6](#), [7](#), [10](#), [12](#)

get_itis_syn, [5](#), [8](#), [12](#)
guess_taxo_level, [2](#), [4](#), [5](#), [6](#), [7](#), [10](#), [12](#)

list_higher_taxo, [2](#), [4-6](#), [6](#), [10](#), [12](#)
list_itis_syn, [5](#), [8](#), [12](#)
list_wiki_syn, [8](#), [12](#)

match_lists, [9](#), [12](#)
melt_canonical, [2](#), [4-7](#), [10](#), [12](#)
melt_cs_field, [3](#), [11](#), [12](#)

taxize, [7](#)
taxotools, [12](#)