Package 'peruflorads43'

<div align="center">August 21, 2023</div>

Title Reviewed Official Classification of Endangered Wild Flora
Species in Peru

Version 0.1.1

Description Provide users with a convenient way to access and analyze information on endangered plant species in Peru based on `Decreto Supremo N 043-2006-AG - Aprueban categorizacion de especies amenazadas de flora silvestre`<https://sinia.minam.gob.pe/normas/aprueban-categorizacion-especies-amenazadas-flora-silvestre>.

Encoding UTF-8

URL https://github.com/PaulESantos/peruflorads43,
https://paulesantos.github.io/peruflorads43/

BugReports https://github.com/PaulESantos/peruflorads43/issues

LazyData true

LazyDataCompression xz

License MIT + file LICENSE

Suggests knitr, rmarkdown, testthat (>= 3.0.0)

RoxygenNote 7.2.3

Depends R (>= 3.5.0),

Config/testthat/edition 3

Maintainer Paul E. Santos Andrade <paulefrens@gmail.com>

NeedsCompilation no

Author Paul E. Santos Andrade [aut, cre]
(<https://orcid.org/0000-0002-6635-0375>)

Repository CRAN

Date/Publication 2023-08-21 13:20:09 UTC

R topics documented:

---

category_ds043_2006	*Get DS043-2006-AG: Aprueban Categorizacion de Especies Ame-nazadas de Flora Silvestre. 13-07-2006 data*

Description

This function takes a species list and tries to match a name in the "DS043-2006-AG: Aprueban Categorizacion de Especies Amenazadas de Flora Silvestre", subseting information for each species. If the name_submitted is a valid name, it will be the duplicated in accepted_name column, else the accepted_name column will display the closest name given the maximum distance defined in max_distance

Usage

```
category_ds043_2006(splist, max_distance = 0.2)
```

Arguments

splist	A character vector specifying the input taxon, each element including genus and specific epithet and, potentially, infraspecific rank, infraspecific name and author name. Only valid characters are allowed (see base::validEnc).
max_distance	match when comparing the submitted name with the closest name matches in the species listed in the "DS043-2006-AG: Aprueban Categorizacion de Especies Amenazadas de Flora Silvestre". The distance used is a generalized Levenshtein distance that indicates the total number of insertions, deletions, and substitutions allowed to match the two names. It can be expressed as an integer or as the fraction of the binomial name. For example, a name with length 10, and a max_distance = 0.1, allow only one change (insertion, deletion, or substitution). A max_distance = 2, allows two changes.

Value

A table with the accepted name and catalog data of the species.

Examples

```
splist <- c("Cleistocactus clavispinus",
            "Welfia alfredi",
            "Matucana haynei")
category_ds043_2006(splist)
```

---

ds43_2006_sps_class *List of species name in tab_ds43_2006 separeted by category*

Description

The 'ds43_2006_sps_class' includes all species separeted by genus, epithet, author, subspecies, variety, and id (position in the tab_ds43_2006).

Usage

```
ds43_2006_sps_class
```

Format

A data.frame.

Examples

```
data(ds43_2006_sps_class)
```

---

ds_043_2006_ag *Species names list from DS043-2006-AG Aprueban Categorizacion de Especies Amenazadas de Flora Silvestre*

Description

Species names list from DS043-2006-AG Aprueban Categorizacion de Especies Amenazadas de Flora Silvestre

Usage

```
ds_043_2006_ag
```

Format

A tibble with the following columns:

categoria A character vector.

accepted_name A character vector. The list of the accepted plant taxa names according to the Taxonomic Name Resolution Service - TNRS.

accepted_family A character vector. The corresponding family name of the accepted_name.

References

DS043-2006-AG: Aprueban Categorizacion de Especies Amenazadas de Flora Silvestre. 13-07-2006

Examples

```
data(ds_043_2006_ag)
str(ds_043_2006_ag)
```

---

search_ds043	*Search species name present in the DS043-2006-AG: Aprueban Categorizacion de Especies Amenazadas de Flora Silvestre. 13-07-2006*

Description

This function takes a species list and tries to match a name in theCategorizacion de Especies Amenazadas de Flora Silvestre, checking if the name is listed in tha dataset.

Usage

```
search_ds043(splist, max_distance = 0.1)
```

Arguments

splist   A character vector specifying the input taxon, each element including genus and specific epithet and, potentially, infraspecific rank, infraspecific name and author name. Only valid characters are allowed (see `base::validEnc`).

max_distance match when comparing the submitted name with the closest name matches in the species listed in the "Categorizacion de Especies Amenazadas de Flora Silvestre". The distance used is a generalized Levenshtein distance that indicates the total number of insertions, deletions, and substitutions allowed to match the two names. It can be expressed as an integer or as the fraction of the binomial name. For example, a name with length 10, and a max_distance = 0.1, allow only one change (insertion, deletion, or substitution). A max_distance = 2, allows two changes.

Value

A character vector that can have three different output values. The first value, "Present," indicates whether the species name is fully matched with the names listed in the 'Categorizacion de Especies Amenazadas de Flora Silvestre'. The second value, "P_updated_name," provides a fuzzy matching of species names. If a species name is not listed in the catalogue, the third value returned will be an empty string.

Examples

```
  # Search for multiple species vector
splist <- c("Cleistocactus clavispinus",
            "Welfia alfredi",
            "Matucana haynei")

  search_ds043(splist)

  # Search for multiple species data.frame
  # base
  df_splist <- data.frame(splist = splist)
  df_splist$peutimber <- search_ds043(df_splist$splist)
```

---

tab_ds43_2006 *List of Plant Species Name accordingly with the DS043-200-AG.*

Description

The 'tab_ds43_2006' contains records belonging to all the species DS043-200-AG.

Usage

```
tab_ds43_2006
```

Format

A tibble with the following columns:

id_cat The fixed species id of the input taxon.

input_genus A character vector. The input genus of the corresponding species name.

input_epitheton A character vector. The specific epithet of the corresponding species name.

rank A character vector. The taxonomic rank: "species","subspecies", "variety", of the corresponding species name.

input_subspecies_epitheton A character vector. If the indicated rank is below species, the subspecies epithet input of the corresponding species name.

taxonomic_status A character vector. description if a taxon is classified as 'accepted', 'synonym', 'no opinion'. According to the Taxonomic Name Resolution Service - TNRS.

accepted_name A character vector. The list of the accepted plant taxa names according to the Taxonomic Name Resolution Service - TNRS.

accepted_family A character vector. The corresponding family name of the accepted_name.

accepted_name_author A character vector. The corresponding author name of the accepted_name, staying empty if the taxonomic_status is "Synonym" or "No opinion".

References

DS043-2006-AG: Aprueban Categorizacion de Especies Amenazadas de Flora Silvestre. 13-07-2006

Examples

```
data(tab_ds43_2006)
str(tab_ds43_2006)
```

---

tab_ds43_2006_position

List of the number positions of the first 3 letters of the species name in the tab_ds43_2006

Description

The 'tab_ds43_2006_position' reports the position (in term of number of rows) of the first three letters (triphthong) for the plant names stored in the variable 'accepted_name' of the table 'tab_ds43_2006'. This indexing system speeds up of the search on the largest list using the package.

Usage

```
tab_ds43_2006_position
```

Format

A data frame with 305 observations on the following 3 variables.

position A character vector. It is the position of the first 3 letters of the species name in the tab_ds43_2006.

triphthong A character vector. First 3 letters of the species name in the tab_ds43_2006.

genus A character vector. Corresponding Genus name.

Examples

```
data(tab_ds43_2006_position)
str(tab_ds43_2006_position)
```

Index