Package 'reptiledb.data'

July 5, 2025

Title Reptile Database Data

Version 0.0.0.1

Description Provides easy access to 'The Reptile Database', a comprehensive catalogue of all living reptile species and their classification. This package includes taxonomic data for over 10,000 reptile species, approximately 2,800 of which are subspecies, covering all extant reptiles. The dataset features taxonomic names, synonyms, distribution data, type specimens, and literature references, making it ready for research and analysis. Data is sourced from 'The Reptile Database' <http://www.reptile-database.org/>.

License MIT + file LICENSE

Depends R (>= 4.1.0)

Suggests testthat (>= 3.0.0), knitr, rmarkdown

Config/testthat/edition 3

Encoding UTF-8

RoxygenNote 7.3.2

LazyData true

URL https://github.com/PaulESantos/reptiledb.data

BugReports https://github.com/PaulESantos/reptiledb.data/issues

Imports httr, rvest, stringr, tibble

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

NeedsCompilation no

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

Repository CRAN

Date/Publication 2025-07-05 15:40:02 UTC

Contents

`check_data_update` *Check if reptile database data needs updating based on date comparison*

Description

This function checks if the local reptile database data is up-to-date by comparing the date extracted from the local dataset name with the date from the latest available file on The Reptile Database website.

Usage

```
check_data_update(silent = FALSE, check_connection = TRUE)
```

Arguments

`silent`	Logical. If TRUE, suppresses messages and only returns results. Default is FALSE.
`check_connection`	Logical. If TRUE, checks internet connection before attempting to access online data. Default is TRUE.

Value

A list containing the following elements:

update_needed Logical. TRUE if an update is needed, FALSE otherwise

local_info List. Information about the local dataset

remote_info List. Information about the remote dataset

message Character. Status message describing the comparison result

recommendation Character. Recommendation for user action

local_date Character. Date of local data in YYYY-MM-DD format

remote_date Character. Date of remote data in YYYY-MM-DD format (if available)

remote_filename Character. Filename of the remote file (if available)

days_difference Numeric. Number of days difference between local and remote data (if both dates available)

If an error occurs or internet connection is not available, only the message element will contain relevant error information.

Examples

```
# Silent check (no messages) - requires internet connection
update_status <- check_data_update(silent = TRUE)

# Verbose check with connection verification
update_status <- check_data_update(silent = FALSE, check_connection = TRUE)

# Check without internet connection verification
update_status <- check_data_update(check_connection = FALSE)
```

`extract_date_from_name` *Extract date from dataset name or filename*

Description

Extract date from dataset name or filename

Usage

```
extract_date_from_name(name, type = "local")
```

Arguments

`name`	Dataset name or filename
`type`	Type of name ("local" or "remote")

Value

A Date object representing the extracted date, or NULL if extraction fails. For local datasets, expects pattern "reptiledb_MMYYYY" (e.g., reptiledb_012025). For remote files, expects pattern "reptile_checklist_YYYY_MM.xlsx".

`get_latest_reptile_download` *Get Latest Reptile Database Download Link*

Description

This function retrieves the most recent download link for reptile database files from the Reptile Database website. It searches for files from the current year first, and if none are found, searches for files from the previous year.

Usage

```
get_latest_reptile_download(
  base_url = "http://www.reptile-database.org/data/",
  current_year = as.numeric(format(Sys.Date(), "%Y")),
  file_types = c("xls", "xlsx", "zip"),
  return_info = FALSE
)
```

Arguments

`base_url`	Character string. The base URL of the reptile database data page. Default is "http://www.reptile-database.org/data/".
`current_year`	Numeric. The current year to search for files. Default is the current system year.
`file_types`	Character vector. File extensions to search for. Default is c("xls", "xlsx", "zip").
`return_info`	Logical. If TRUE, returns a list with detailed information about the found file. If FALSE, returns only the URL. Default is FALSE.

Details

The function performs web scraping on the specified URL to find download links. It prioritizes files from the current year, but will fall back to the previous year if no current year files are available.

The function requires the following packages: rvest, dplyr, and stringr. These packages must be installed before using this function.

Value

If `return_info = FALSE`, returns a character string with the URL of the most recent file, or NULL if no suitable file is found. If `return_info = TRUE`, returns a list containing:

url Character. The complete URL of the file

filename Character. The name of the file

file_type Character. The file extension

extraction_date Date. The date when the link was extracted

source_page Character. The source webpage URL

Returns NULL if no suitable file is found or if an error occurs during web scraping.

See Also

http://www.reptile-database.org/ for more information about the Reptile Database.

Examples

```
# Get just the URL - requires internet connection
url <- get_latest_reptile_download()

# Get detailed information
info <- get_latest_reptile_download(return_info = TRUE)
```

```
# Search for specific file types
zip_url <- get_latest_reptile_download(file_types = "zip")

# Search for files from a specific year
url_2024 <- get_latest_reptile_download(current_year = 2024)
```

`reptiledb_012025` *Reptile Checklist with Subspecies Information*

Description

A comprehensive dataset extracted from The Reptile Database containing taxonomic and nomenclatural information for reptile species and their subspecies. This tibble includes detailed columns related to authorship, type species, and taxonomic changes.

Usage

```
reptiledb_012025
```

Format

A tibble with 14,474 rows and 16 columns:

order Taxonomic order of the reptile (e.g., `"Sauria"`).

family Taxonomic family (e.g., `"Scincidae"`).

genus Genus name.

epithet Species epithet (second part of the species name).

species Full species name (genus + epithet).

species_author Primary author(s) of the species name.

species_name_year Year the species was described.

subspecies_name Epithet of the subspecies (if any).

subspecie_author_info Full author citation of the subspecies.

subspecies_name_author Author(s) of the subspecies name.

subspecies_year Year the subspecies was described.

type_species Name of the type species, if available.

change Text description of any taxonomic or nomenclatural change.

rdb_sp_id Unique identifier assigned by The Reptile Database.

nomenclature_change Logical flag indicating if a nomenclatural change has occurred (`TRUE` / `FALSE`).

nomenclature_change_species Logical flag indicating if the nomenclatural change affects the species level (`TRUE` / `FALSE`).

Details

This dataset is part of the `reptiledb.data` package and provides structured access to reptile taxonomy data, enabling users to filter, analyze, or visualize species and subspecies information across multiple reptile families and genera.

Source

http://www.reptile-database.org/

Index