

Package: iucnr (via r-universe)

October 12, 2026

Title Match Species Names to 'IUCN' Red List Data

Version 0.0.3

Description Facilitates matching species names with the International Union for Conservation of Nature ('IUCN') Red List of Threatened Species data. Streamlines determining conservation statuses and vernacular names through exact and fuzzy taxonomic matching. Designed to support biodiversity research and conservation planning relying on local data from the 'iucnrdata' package.

License MIT + file LICENSE

Config/testthat/edition 3

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 8.0.0

URL <https://paulesantos.github.io/iucnr/>,
<https://github.com/PaulESantos/iucnr>

BugReports <https://github.com/PaulESantos/iucnr/issues>

Depends R (>= 4.1.0)

LazyData true

LazyDataCompression xz

Imports assertthat, cli, data.table, dplyr, fuzzyjoin, lifecycle,
stringr, tibble

Suggests knitr, rmarkdown, rlang, testthat (>= 3.0.0), iucnrdata

VignetteBuilder knitr

Additional_repositories <https://paulesantos.r-universe.dev>

Config/roxygen2/version 8.0.0

Config/pak/sysreqs libicu-dev

Repository <https://paulesantos.r-universe.dev>

Date/Publication 2026-08-04 05:25:13 UTC

RemoteUrl <https://github.com/PaulESantos/iucnr>
RemoteRef HEAD
RemoteSha 7739e9eed280b2dc8fb4a68e1e2b7eee8c45535a

Contents

filter_iucn_genus	2
genus_exact_match	3
genus_fuzzy_match	4
get_iucn_by_category	5
get_iucn_by_taxonomy	6
get_iucn_common_name	7
get_iucn_status	9
iucn_data_version	10
iucn_genus	10
iucn_match	11
species_direct_match	12
species_fuzzy_match	13
Index	14

filter_iucn_genus	<i>Filter and Match Genera from a Species List</i>
-------------------	--

Description

[Experimental]

This function processes a list of species names to match their genera with a database. It first attempts to match genera exactly and then applies fuzzy matching for unmatched cases. The function provides detailed messages about the matching process, including the number of exact matches, fuzzy matches, and unmatched inputs.

Usage

```
filter_iucn_genus(splist, verbose = getOption("iucnr.verbose", FALSE))
```

Arguments

- | | |
|---------|---|
| splist | A character vector containing the species names to be processed. Each name should include the genus and species (e.g., "Panthera uncia"). |
| verbose | Logical value indicating whether to print informative messages. Defaults to <code>getOption("iucnr.verbose", FALSE)</code> . |

Details

The function follows these steps:

1. Preprocess the species list using the helper function `format_df()`.
2. Perform exact genus matching using `genus_exact_match()`.
3. If there are unmatched genera, apply fuzzy matching using `genus_fuzzy_match()`.
4. Combine results from both exact and fuzzy matching into a single tibble.
5. Identify unmatched genera and provide a summary of matching statistics via informative messages.

Value

A tibble with the following columns:

- `submitted_name`: The original species name submitted.
- `or_genus_name`: The genus extracted from the submitted name.
- `matched_genus`: The genus name that was matched (either exactly or via fuzzy matching).
- `direct_match_genus`: A logical value indicating whether the genus was matched exactly.
- `fuzzy_match_genus`: A logical value indicating whether the genus was matched using fuzzy matching.

If any genera remain unmatched, they will have NA in the `matched_genus` column, and both `direct_match_genus` and `fuzzy_match_genus` will be FALSE or NA.

Examples

```
filter_iucn_genus(c("Panthera uncia", "Cedrela odorata"))
```

genus_exact_match	<i>Exact Genus Match for IUCN Data</i>
-------------------	--

Description

[Experimental]

Matches the genus names from a given tibble to a reference IUCN genus dataset and returns a tibble with the matching status.

Usage

```
genus_exact_match(df)
```

Arguments

`df` A tibble, result of internal function `format_df`.

Details

The function:

- Matches the genus names in `or_genus_name` to the IUCN genus dataset (`iucnr::iucn_genus`).
- Combines matched and unmatched rows, adding a boolean indicator column (`direct_match_genus`) to show match status.

The reference dataset `iucnr::iucn_genus` is an internal dataset included in the `iucnr` package.

Value

A tibble with the following columns:

submitted_name Original submitted name from the input tibble.

or_genus_name Original genus name from the input tibble.

matched_genus Matched genus name from the IUCN reference dataset, or NA if no match was found.

direct_match_genus Logical column indicating whether the genus was successfully matched (TRUE) or not (FALSE).

Examples

```
df <- iucnr::format_df(c("Panthera uncia", "Cedrela odorata"))
genus_exact_match(df)
```

genus_fuzzy_match	<i>Fuzzy Match Genus Name</i>
-------------------	-------------------------------

Description

[Experimental]

Tries to fuzzy match the genus name from the input dataset to the IUCN genus list (`iucnr::iucn_genus`). Uses `fuzzyjoin::stringdist_left_join()` for approximate string matching, allowing for minor differences in the genus names.

Usage

```
genus_fuzzy_match(df)
```

Arguments

df A tibble resulting of applying function `genus_exact_match`.

Details

The function handles cases where the genus name in the input data differs slightly (e.g., due to typos). It assigns the closest match based on string distance (with a maximum distance of 1) and identifies unmatched entries. For genera with multiple fuzzy matches, the function issues a warning, saves the ambiguous matches to a CSV file, and proceeds by selecting the first match. The user is advised to manually curate these ambiguous cases and re-run the function if necessary.

Value

Returns a tibble with additional columns with information about fuzzy matching:

- `matched_genus`: Contains the matched genus name (if a match was found) or NA otherwise.
- `fuzzy_match_genus`: Logical column indicating whether the genus was successfully matched (TRUE) or not (FALSE).
- `fuzzy_genus_dist`: The string distance between the original and matched genus names (only for matched rows).

Examples

```
# Example input tibble

df <- tibble::tibble(
  submitted_name = c("Panthera uncia", "Lynx lynx", "Ara militaris",
                    "Zonotrichia capensis", "Cyathula coriacea", "Impatiens confusa"),
  or_genus_name = c("Panthera", "Lynx", "Ara", "Zonotrichia", "Cyathula", "Impatiens")
)

# Exact match
node_1_processed <- genus_exact_match(df)

# Fuzzy matching

node_1_FALSE <- node_1_processed |>
  dplyr::filter(direct_match_genus == FALSE)

node_2_processed <- genus_fuzzy_match(node_1_FALSE)

print(node_2_processed)
```

get_iucn_by_category *Retrieve IUCN Species Data by Red List Category*

Description

[Experimental]

Filters IUCN species data from the packaged dataset by Red List conservation status categories such as Critically Endangered ("CR"), Endangered ("EN"), Vulnerable ("VU"), etc.

Usage

```
get_iucn_by_category(category = c("CR", "EN", "VU"), format = "tbl")
```

Arguments

- | | |
|----------|--|
| category | Character vector specifying the Red List conservation status categories to filter by. Accepts short codes (e.g., "CR", "EN", "VU", "NT", "LC", "DD", "EX", "EW") or full category names (e.g., "Critically Endangered", "Vulnerable"). Default is c("CR", "EN", "VU") (all threatened categories). |
| format | A character string specifying the output format: <ul style="list-style-type: none"> • "tbl": Returns a tibble with species, taxonomy, and Red List status (default). • "vect": Returns a character vector of scientific names. |

Value

Depending on format: a tibble or a character vector of species scientific names.

Examples

```
# Get all Critically Endangered and Endangered species
cr_en_species <- get_iucn_by_category(category = c("CR", "EN"), format = "tbl")
head(cr_en_species)
```

get_iucn_by_taxonomy *Retrieve IUCN Species Data by Taxonomic Hierarchy*

Description**[Experimental]**

Filters IUCN species and conservation status data by higher taxonomic ranks such as kingdom, phylum, class, order, family, or genus.

Usage

```
get_iucn_by_taxonomy(
  kingdom = NULL,
  phylum = NULL,
  class = NULL,
  order = NULL,
  family = NULL,
  genus = NULL,
  status = NULL,
  format = "tbl"
)
```

Arguments

kingdom	Optional character string or vector specifying kingdom name(s) (e.g., "PLANTAE", "ANIMALIA").
phylum	Optional character string or vector specifying phylum name(s).
class	Optional character string or vector specifying class name(s) (e.g., "MAMMALIA", "AMPHIBIA").
order	Optional character string or vector specifying order name(s) (e.g., "PRIMATES", "CARNIVORA").
family	Optional character string or vector specifying family name(s) (e.g., "FELIDAE", "ORCHIDACEAE").
genus	Optional character string or vector specifying genus name(s) (e.g., "Panthera", "Puya").
status	Optional character vector to filter by Red List conservation status category. Accepts short codes (e.g., "CR", "EN", "VU", "NT", "LC", "DD", "EX", "EW") or full names (e.g., "Critically Endangered").
format	A character string specifying the output format: "tbl": Returns a tibble with taxonomic hierarchy and Red List status (default). "vect": Returns a character vector of scientific names.

Value

Depending on format: a tibble or a character vector of species scientific names.

Examples

```
# Search Endangered or Critically Endangered species in Felidae family
felidae_threatened <- get_iucn_by_taxonomy(
  family = "Felidae",
  status = c("CR", "EN"),
  format = "tbl"
)
head(felidae_threatened)
```

get_iucn_common_name *Retrieve Vernacular (Common) Names for Species*

Description**[Experimental]**

This function retrieves vernacular (common) names for species using IUCN data. It filters vernacular names based on the taxonID and merges them with species metadata. The output can be customized as either a vector of common names or a tibble containing detailed information.

Usage

```
get_iucn_common_name(
  splist,
  format = "vect",
  verbose = getOption("iucnr.verbose", FALSE)
)
```

Arguments

- | | |
|---------|---|
| splist | A character vector containing the scientific names of species for which vernacular names are to be retrieved. |
| format | A character string specifying the output format. Options are: <ul style="list-style-type: none"> • "vect": Returns a vector of common names. • "tbl": Returns a tibble with species details and their vernacular names. Default is "tbl". |
| verbose | Logical value indicating whether to print informative messages. Defaults to getOption("iucnr.verbose", FALSE). |

Value

Depending on the format argument:

- If "vect", a character vector of common names.
- If "tbl", a tibble with the following columns:
 - submitted_name** The scientific name of the species as submitted.
 - matched_species** The matched species name from IUCN data.
 - common_name** The vernacular (common) name. If no common name is found, it returns "___".

Examples

```
species <- c("Panthera uncia", "Cedrela odorata")
result <- get_iucn_common_name(splist = species, format = "tbl")
print(result)

common_names <- get_iucn_common_name(splist = species, format = "vect")
print(common_names)
```

get_iucn_status	<i>Retrieve Conservation Status of Species</i>
-----------------	--

Description

[Experimental]

This function retrieves the conservation status of species from the `iucn_data` dataset. It supports multiple species names and uses `data.table` for efficient performance. The output format can be customized to either a vector of conservation statuses or a tibble.

Usage

```
get_iucn_status(  
  splist,  
  format = "vect",  
  verbose = getOption("iucnr.verbose", FALSE)  
)
```

Arguments

<code>splist</code>	A character vector containing species names to search for in the <code>iucn_data</code> dataset.
<code>format</code>	A character string specifying the output format. Options are: "vect": Returns a vector of conservation statuses. "tbl": Returns a tibble with species names and conservation statuses. Default is "vect".
<code>verbose</code>	Logical value indicating whether to print informative messages. Defaults to <code>getOption("iucnr.verbose", FALSE)</code> .

Value

A vector or tibble, depending on the specified format, with species names and their corresponding conservation statuses. If no match is found for a species, it will return "no match found".

Examples

```
species <- c("Panthera uncia", "Cedrela odorata")  
result <- get_iucn_status(splist = species, format = "tbl")  
print(result)
```

iucn_data_version	<i>Identify Version of Packaged IUCN Data Snapshot</i>
-------------------	--

Description

[Experimental]

Returns information about the version of the IUCN Red List data snapshot packaged and available in the offline environment (iucnrdata).

Usage

```
iucn_data_version(details = FALSE)
```

Arguments

details	A logical value. If FALSE (default), returns a character string with the IUCN Red List data version (e.g., "Version 2026-1"). If TRUE, returns a tibble with detailed version metadata including data version, package version, freshness status, and data source URL.
---------	--

Value

Depending on details: a character string or a tibble containing version metadata.

Examples

```
# Get version string
iucn_data_version()

# Get detailed version metadata
iucn_data_version(details = TRUE)
```

iucn_genus	<i>IUCN Genus Database</i>
------------	----------------------------

Description

The iucn_genus dataset is part of the iucnr package and contains information on the genera of various species listed in the **IUCN Red List of Threatened Species**. Each genus is associated with one or more internal identifiers corresponding to its listings within the IUCN database.

Usage

```
iucn_genus
```

Format

A tibble with 22,813 rows and 2 columns:

genus_name Character. The name of the genus.

internal_ids Character. A string of internal IUCN identifiers for each genus, separated by " - " when multiple identifiers are present.

Details

This dataset provides a comprehensive reference for genus-level classifications in the IUCN database, facilitating taxonomic analyses and cross-referencing with other species data. The `internal_ids` column can include multiple identifiers, representing various listings for the same genus within the IUCN records.

Source

IUCN Red List of Threatened Species

iucn_match	<i>Perform IUCN species matching at genus and species levels</i>
------------	--

Description**[Experimental]**

This function matches a list of species against the IUCN database using genus filtering, exact species matching, and fuzzy species matching. The output includes matched species along with their IUCN Red List category and criteria.

Usage

```
iucn_match(splist, verbose = getOption("iucnr.verbose", FALSE))
```

Arguments

<code>splist</code>	A dataframe or character vector containing species names to be matched.
<code>verbose</code>	Logical value indicating whether to print informative messages. Defaults to <code>getOption("iucnr.verbose", FALSE)</code> .

Value

A dataframe with the following columns:

- `id`: Row number identifier.
- `submitted_name`: Original species names from the input.
- `matched_species`: Matched species name from the IUCN database (or --- if not matched).
- `redlist_category`: Red List category of the matched species (or --- if not available).
- `internal_taxon_id`: IUCN taxon identifier (or NA).

Examples

```
iucn_match(c("Panthera uncia", "Cedrela odorata"))
```

species_direct_match	<i>Perform direct species matching against IUCN data</i>
----------------------	--

Description

This function matches species names (submitted_name) from the input dataframe directly to the IUCN species list. It identifies matches and non-matches based on exact string comparison, and returns a dataframe with match status and relevant columns.

Usage

```
species_direct_match(df)
```

Arguments

df	A dataframe containing species names in the column submitted_name and their matched genera in matched_genus.
----	--

Value

A dataframe that includes:

- submitted_name: Original species names from the input.
- matched_genus: The genus matched from the input.
- matched_species: The species name matched directly in the IUCN dataset.
- direct_match_species: Boolean indicating whether the species was directly matched (TRUE) or not (FALSE).

Examples

```
df <- filter_iucn_genus("Panthera uncia")  
species_direct_match(df)
```

species_fuzzy_match	<i>Perform fuzzy species matching against IUCN data</i>
---------------------	---

Description

This function matches a dataframe of species names (submitted_name) to the IUCN database using fuzzy string matching. It returns a dataframe with matched and unmatched entries, including additional details on the matching process.

Usage

```
species_fuzzy_match(df)
```

Arguments

df	A dataframe containing species names in the column submitted_name and their matched genera in matched_genus.
----	--

Value

A dataframe with the original input, matched species names, fuzzy string distances, and match status.

Examples

```
df <- filter_iucn_genus("Panthera uncia")
species_fuzzy_match(df)
```

Index

* **datasets**

iucn_genus, [10](#)

filter_iucn_genus, [2](#)

genus_exact_match, [3](#)

genus_fuzzy_match, [4](#)

get_iucn_by_category, [5](#)

get_iucn_by_taxonomy, [6](#)

get_iucn_common_name, [7](#)

get_iucn_status, [9](#)

iucn_data_version, [10](#)

iucn_genus, [10](#)

iucn_match, [11](#)

species_direct_match, [12](#)

species_fuzzy_match, [13](#)