Package: rCAX (via r-universe)

November 25, 2024

Type Package

Date 2023-11-28

Title Coordinated Assessments REST API R Client

Description This package is an R client for the StreamNet Coordinated Assessments HLI REST API.

Version 1.0.3

License GPL-3 + file LICENSE

URL https://github.com/nwfsc-cb/rCAX

BugReports https://github.com/nwfsc-cb/rCAX/issues

Encoding UTF-8

Language en-US

Imports crul (>= 1.3), jsonlite (>= 1.8), utils

RoxygenNote 7.2.1

Roxygen list(markdown = TRUE)

NeedsCompilation no

Author Eli Holmes [aut, cre], Mari Williams [aut], Katie Barnas [ctb]

Suggests knitr, rmarkdown, ggplot2

VignetteBuilder knitr

Depends R (>= 3.0.0)

Config/pak/sysreqs libssl-dev

Repository https://nmfs-opensci.r-universe.dev

RemoteUrl https://github.com/nwfsc-cb/rCAX

RemoteRef HEAD

RemoteSha cc3dc03f72318d3bcab714923ba1271ed4f09629

Contents

`rCAX-package` *rCAX*

Description

This package is an R client for the StreamNet Coordinated Assessments HLI REST API.

Authentication

The rCAX R package includes a pull key so that you can query the CAX database. Should you need to use your own key see the vignette on the API functions.

`caxesu` *ESU/DPS names*

Description

An internal dataset of the ESU/DPS names. Useful for filtering.

Examples

```
rCAX:::caxesu
```

caxtabs *CAX API table names and table_id*

Description

An internal dataset output from a call to "ca/tables" per REST API documentation. It is a data frame with a number of columns. The name and id columns are the ones used internally.

Examples

```
tab <- rCAX:::caxtabs
head(tab[,c("name", "id")])
```

cax_column_definitions *Column names and definitions*

Description

An internal dataset with column name and definitions. Produced from the CAX HLI Tabular Query Excel file definition tabs. It is a data frame with name and definition.

Examples

```
head(rCAX:::cax_column_definitions)
```

check_key *Get the key from the user system environment variable*

Description

Saved in `CAX_KEY` in the `.Renviron` file. Alternatively it can be passed in.

Usage

```
check_key(key)
```

Arguments

`key`	the key

References

This function is modeled off check_key() in `https://github.com/ropensci/rredlist`

internal_data *Internal data used by functions*

Description

rCAX has a number of internally stored data (in sysdata.rda) that are used by the functions. If the user wants to use these, preface the name with rCAX:::. For example, rCAX:::caxesu shows the list of ESU/DPS names.

Details

Internal data

- caxsuperpops: the SuperPopulations table
- caxtabs: the API table names and their table_ids. Used to look up table_id
- caxpops: the Populations table
- caxesu: the list of ESU/DPS names. Useful for filtering.
- nosa_xport_colnames, sar_xport_colnames, pni_xport_colnames, rpers_xport_colnames, presmolt_xport_colnames, juvout_xport_colnames: the column names that appear in the HLI Tabular Query output so that rCAX _xport functions produce the same columns in the same order.
- cax_column_definitions: a data frame with column name and definitions. Produced from the HLI Tabular Query Excel files

Examples

```
rCAX:::caxesu
```

rcax_base *Set the base API url*

Description

"https://api.streamnet.org/api/v1"

Usage

```
rcax_base()
```

Examples

```
rcax_base()
```

rcax_citation *Get the citation CAX version*

Description

Get the citation CAX version

Usage

```
rcax_citation()
```

Value

API citation as character string

Examples

```
rcax_citation()
```

rcax_datasets *Get the list of CAX datasets (tables) with ids*

Description

Get the list of CAX datasets (tables) with ids

Usage

```
rcax_datasets(
  cols = c("name", "id", "description"),
 GETargs = list(table_id = NULL, recordloc = "records", key = NULL, parse = TRUE),
  ...
)
```

Arguments

cols	the columns to return. default is name, description and id. use cols=NULL to get all the columns
GETargs	A list of arguments for the rcax_GET() call. These do not need to be specified by the user unless the user wants to change the default values. table_id This is the CAX table id. It is looked up using tablename and rCAX:::caxpops (an internal dataset). If table_id is passed in via GETargs, it will override the default table id lookup. recordloc This is the name of the record we want in the list returned from the GET call. key is the API key.
...	Curl options passed to HttpClient

References

API docs at https://www.streamnet.org/resources/exchange-tools/rest-api-documentation/

Examples

```
a <- rcax_datasets(cols=NULL)
colnames(a)

head(a[,1:2])
```

rcax_escdata *Get EscData table*

Description

Returns EscData table sorted by refid and countdate.

The EscData table_id is set automatically using a saved data frame from a rcax_datasets() call. The table is saved in R/sysdata.rda.

rcax_escdata() will download 1000 records.

Usage

```
rcax_escdata(
  tablename = "EscData",
  flist = NULL,
  qlist = NULL,
  cols = NULL,
  sortcols = c("countdate", "refid"),
  type = c("data.frame", "colnames"),
 GETargs = list(table_id = NULL, recordloc = "records", key = NULL, parse = TRUE),
  ...
)
```

Arguments

tablename	The name of the table in the CAX API. See vignette("CAX_Tables", package = "rCAX")
flist	A filter for the query. See details and rcax_filter()
qlist	Additional query parameters. See rcax_table_query() for details.
cols	column names to return. Use cols=NULL if you want all columns. Names are case insensitive.
sortcols	The columns to sort on for the returned table. The order of sortcols indicates the order of sorting. Start with the smallest group and work outwards, e.g. c("spawningyear", "popid", "MPG"). Names are case insensitive.

`type` whether to return the table ("data.frame") or colnames with definitions ("colnames"). Default is to return the table as a data frame.

`GETargs` A list of arguments for the `rcax_GET()` call. These do not need to be specified by the user unless the user wants to change the default values. `table_id` This is the CAX table id. It is looked up using `tablename` and `rCAX:::caxpops` (an internal dataset). If `table_id` is passed in via `GETargs`, it will override the default table id lookup. `recordloc` This is the name of the record we want in the list returned from the GET call. `key` is the API key.

`...` Curl options passed to `HttpClient`

Value

data frame if `type="data.frame"` (default) and the colnames if `type="colnames"`

Examples

```
# First 5 columns of first 5 results
rcax_escdata(cols = NULL, qlist=list(limit=5))[,1:5]

# to print the first column names and definitions
head(rcax_escdata(type="colnames"))
```

`rcax_filter` *Create filter to add to GET call*

Description

The table returned by an API query can be filtered using the filter query parameter. The value of the parameter is JSON. Examples are

- `filter=[{"field":"popid","value":"7","type":"string"}]`
- `filter=[{"field":"popid","value":[7,8],"type":"list"}]`
- `filter=[{"field":"esu_dps","value":"Salmon, Chinook (Snake River spring/summer-run ESU)","type":'`

Usage

```
rcax_filter(x)
```

Arguments

`x` list of property (column) names, values in this format `list(colname=value)`. If the property can take multiple values, this is passed in as `list(colname=c(value1, value2))`

Details

The property names in x are the column names in a table being returned by rcax_table_query(). Note the values are not case sensitive. So a value of Spring will return the data with Spring and spring.

crul properly formats the filter given the JSON. For reference, here is a url example. Note that in the filter double quotes are required. A single quote will not filter.

```
httr::GET('https://api.streamnet.org/api/v1/ca.json?table_id=4EF09E86-2AA8-4C98-A983-A272C2C2C7E3&X
```

Value

The JSON for the filter query parameter

Examples

```
rcax_filter(list(commonpopname="GRCAT"))
rcax_filter(list(popid=c(7,8)))
```

rcax_GET *Make a GET call to API*

Description

Make a GET call to API

Usage

```
rcax_GET(path, key = NULL, query = NULL, ...)
```

Arguments

path	what to add after the base api path
key	A CAX API key. See details.
query	a list of query parameters with their values, e.g. list(param=value)
...	Curl options passed to HttpClient

Details

API Key: rCAX includes a "read-only" API key. Thus for normal read-only use, you do not need an API key. If you need to use your own key, see the instructions in vignette("api_get", package = "rCAX").

User Agent String: rcax_ua() sets the user agent string that is passed to the CAX REST API. If the rCAX package code is reused (for another package), make sure to change the user agent string to indicate that the queries are from another package.

References

API docs at https://www.streamnet.org/resources/exchange-tools/rest-api-documentation/

This function is modeled off `rl_GET()` in https://github.com/ropensci/rredlist

Examples

```
## Not run:
rcax_GET("ca/tables")

## End(Not run)
```

rcax_hli *Return fish HLI tables from Coordinated Assessments data eXchange*

Description

Returns the HLI table with some additional metadata, such as NMFS_PopID and update dates, in the default `tabletype="xport` case. This is the table that one can download from the CAP Fish HLIs Tabular Query. The HLI is specified using its short code:

- NOSA: Natural Origin Spawner Abundance
- SAR: Smolt to Adult Ratios
- PNI: Proportionate Natural Influence of supplementation hatcheries
- RperS: Recruits per Spawner
- JuvOut: Juvenile Outmigrants
- PreSmolt: Presmolt Abundance

The query will download 1000 records by default; you can change this by passing in `qlist=list(limit=1100)` for example. To make a filtered query e.g. for one ESU or popid) pass in the `flist` argument. Field names and values are case insensitive, e.g. `POPID`, `popid` and `PopID` return the same result. If your filter value is not in the table, the returned table will be empty. Examples:

- `flist=list(popid=7)` return values for popid 7.
- `flist=list(nmfs_popid=6)` return values for NMFS_PopID 6.
- `flist=list(esu_dps="Salmon, Chinook (Snake River spring/summer-run ESU)")` return values for one ESU. Use `rCAX:::caxesu` to see a list of ESU_DPS names.

Usage

```
rcax_hli(
  hli = c("NOSA", "SAR", "PNI", "PreSmolt", "JuvOut", "RperS"),
  flist = NULL,
  qlist = NULL,
  cols = NULL,
  sortcols = NULL,
  type = c("data.frame", "colnames"),
  tabletype = c("xport", "base"),
 GETargs = list(table_id = NULL, recordloc = "records", key = NULL, parse = TRUE),
  ...
)
```

Arguments

hli	The HLI short name: NOSA, SAR, PNI, PreSmolt, JuvOut, RperS
flist	A filter for the query. See details and rcax_filter()
qlist	Additional query parameters. See rcax_table_query() for details.
cols	column names to return. Use cols=NULL if you want all columns. Names are case insensitive.
sortcols	The columns to sort on for the returned table. The order of sortcols indicates the order of sorting. Start with the smallest group and work outwards, e.g. c("spawningyear", "popid", "MPG"). Names are case insensitive.
type	whether to return the table ("data.frame") or colnames with definitions ("colnames"). Default is to return the table as a data frame.
tabletype	"xport" (default) or "base". XPort has additional useful metadata.
GETargs	A list of arguments for the rcax_GET() call. These do not need to be specified by the user unless the user wants to change the default values. table_id This is the CAX table id. It is looked up using tablename and rCAX:::caxpops (an internal dataset). If table_id is passed in via GETargs, it will override the default table id lookup. recordloc This is the name of the record we want in the list returned from the GET call. key is the API key.
...	Curl options passed to HttpClient

Details

The XPort tables have a few columns that do not appear in the Excel files that one can download from StreamNet: "species", "publish", "num", "hli", "agency", "esudps", "hli_id". "esudps" is a slight variant on "esu_dps" and is removed. "num" is a sort flag and is removed. "hli_id" is an internal reference id and is removed. "publish" will always be Yes but is left in. The other columns are left in. The colnames are re-sorted into the order found in the downloaded Excel files with the extra columns added to the end.

The table columns are sorted into the order that appears in CAP Fish HLIs Tabular Query. If you want to see the original columns in the original order, use cols=1:50, say, to see the first 50 columns in the original order.

If tabletype="base" is used, the base tables without the added metadata are returned. The data are the same just missing the extra metadata.

If you want to see only the column names use type="colnames.

Value

data frame if type="data.frame" (default) and the colnames if type="colnames"

See Also

rcax_table_query(), rcax_filter(), rcax_table_name()

Examples

```
# return NMFS_PopID of second record
# Note the part after $ is case sensitive
id <- rcax_hli("NOSA", qlist=list(limit=2))$nmfs_popid[2]

a <- rcax_hli("NOSA",
         flist=list(nmfs_popid=id),
         cols=c("nmfs_popid", "spawningyear", "tsaej", "nosaej")
         )
head(a)

# First 3 columns
rcax_hli("NOSA", qlist=list(limit=3))[,1:3]

# to print the first column names and definitions
head(rcax_hli("NOSA", type="colnames"))
```

rcax_parse *Parse results from API call*

Description

Parse results from API call

Usage

```
rcax_parse(x, parse)
```

Arguments

x	what needs to be parsed
parse	logical true/false

References

This function is modeled off rl_parse() in https://github.com/ropensci/rredlist

See Also

rcax_GET() for examples

rcax_superpops *Get SuperPopulations table*

Description

Returns SuperPopulations table

The SuperPopulations table_id is set automatically using a saved data frame from a `rcax_datasets()` call. The table is saved in `R/sysdata.rda`.

`rcax_superpops()` will download 1000 records. Pass in `flist` to filter.

Usage

```
rcax_superpops(
  tablename = "SuperPopulations",
  flist = NULL,
  qlist = NULL,
  cols = NULL,
  sortcols = c("popid"),
  type = c("data.frame", "colnames"),
 GETargs = list(table_id = NULL, recordloc = "records", key = NULL, parse = TRUE),
  ...
)
```

Arguments

`tablename`	The name of the table in the CAX API. See `vignette("CAX_Tables", package = "rCAX")`
`flist`	A filter for the query. See details and `rcax_filter()`
`qlist`	Additional query parameters. See `rcax_table_query()` for details.
`cols`	column names to return. Use cols=NULL if you want all columns. Names are case insensitive.
`sortcols`	The columns to sort on for the returned table. The order of sortcols indicates the order of sorting. Start with the smallest group and work outwards, e.g. `c("spawningyear", "popid", "MPG")`. Names are case insensitive.
`type`	whether to return the table ("data.frame") or colnames with definitions ("colnames"). Default is to return the table as a data frame.
`GETargs`	A list of arguments for the `rcax_GET()` call. These do not need to be specified by the user unless the user wants to change the default values. `table_id` This is the CAX table id. It is looked up using `tablename` and `rCAX:::caxpops` (an internal dataset). If `table_id` is passed in via `GETargs`, it will override the default table id lookup. `recordloc` This is the name of the record we want in the list returned from the GET call. `key` is the API key.
`...`	Curl options passed to `HttpClient`

Value

data frame if `type="data.frame"` (default) and the colnames if `type="colnames"`

See Also

`rcax_table_query()`, `rcax_filter()`

Examples

```
# print the first columns and definitions
a <- rcax_superpops(type="colnames")
paste(a$name, a$definition, sep=": ")[1:5]

# get one record
rcax_superpops(qlist=list(limit=1))
```

`rcax_table_cols` *Returns the colnames to use for tables*

Description

Returns the columns in the same order as the downloaded Excel files have in https://www.streamnet.org/data/hli. Adds on a few extra columns. The colnames are saved as internal data in `R/sysdata.rda` and defined in `inst/docs/create_sysdata.R`

Usage

```
rcax_table_cols(hli, type = c("xport", "base"))
```

Arguments

`hli`	The HLI short name: NOSA, SAR, PNI, RperS, JuvOut, or PreSmolt
`type`	XPort table or base table

Details

- NOSA: Natural origin spawner abundance
- SAR: SAR
- PNI: PNI
- RperS: Recruits per spawner
- JuvOut: Short for JuvenileOutmigrants.
- PreSmolt: Short for PresmoltAbundance.

Value

A vector of column names

Examples

```
# Show the first 5 colnames
rcax_table_cols("NOSA")[1:5]
```

`rcax_table_name` *Return table name given the HLI short name*

Description

- NOSA: Natural origin spawner abundance
- SAR: SAR
- PNI: PNI
- RperS: Recruits per spawner
- JuvOut: Short for JuvenileOutmigrants.
- PreSmolt: Short for PresmoltAbundance.

Usage

```
rcax_table_name(hli, type = c("xport", "base"))
```

Arguments

`hli`	The HLI short name: NOSA, SAR, PNI, RperS, JuvOut, or PreSmolt
`type`	XPort table or base table

Examples

```
rcax_table_name("NOSA", type="xport")
```

`rcax_table_query` *Return tables by table name from Coordinated Assessments data eXchange*

Description

This is the base function for queries for most of the CAX tables. Default queries will download 1000 records. You will want to make a filtered query by passing in `flist` as a list with the column name values to filter on.

Usage

```
rcax_table_query(
  tablename = NULL,
  flist = NULL,
  qlist = NULL,
  cols = NULL,
  sortcols = NULL,
  type = c("data.frame", "colnames"),
  GETargs = list(table_id = NULL, recordloc = "records", key = NULL, parse = TRUE),
  ...
)
```

Arguments

`tablename` The name of the table in the CAX API. See `rcax_datasets()` for the names.

`flist` A filter for the query. See details and `rcax_filter()`

`qlist` Additional query parameters. See `rcax_table_query()` for details.

`cols` column names to return. Use cols=NULL if you want all columns. Names are case insensitive.

`sortcols` The columns to sort on for the returned table. The order of sortcols indicates the order of sorting. Start with the smallest group and work outwards, e.g. `c("spawningyear", "popid", "MPG")`. Names are case insensitive.

`type` whether to return the table ("data.frame") or colnames with definitions ("colnames"). Default is to return the table as a data frame.

`GETargs` A list of arguments for the `rcax_GET()` call. These do not need to be specified by the user unless the user wants to change the default values. `table_id` This is the CAX table id. It is looked up using `tablename` and `rCAX:::caxpops` (an internal dataset). If `table_id` is passed in via `GETargs`, it will override the default table id lookup. `recordloc` This is the name of the record we want in the list returned from the GET call. `key` is the API key.

`...` Curl options passed to `HttpClient`

Details

The required query parameters are table_id and XAPIKEY. These are set from the information in `GETargs`. table_id is normally looked up from `rcax_datasets()` using the `tablename`. The `qlist` argument is any additional query parameters. The following are some of the available extra (meaning not required) query parameters.

- page=num Integer of the page selected.
- per_page=num Integer for the number of records per page.
- limit=num Maximum number of records to return. Default is 1000.
- agency=character Agency that submitted the record. You can designate one or more (comma-separated) agency acronyms: IDFG, WDFW, ODFW, CRITFC, CCT, NOAA.
- updated_since Date since epoch in seconds: 2012-04-24 15:05:22 -0400 = 1335294322.

- popid Population id in NOSA tables.

Filtering, i.e. subsetting, the query is specified using the flist argument. This is a list with the column name and value. See rcax_filter() for the filtering code which writes the filter query parameter that is added to the GET query. See rcax_filter() for examples of how one passes in flist.

Value

data frame if type="data.frame" (default) and the colnames if type="colnames"

References

API docs at https://www.streamnet.org/resources/exchange-tools/rest-api-documentation/

See Also

rcax_GET(), rcax_hli(), rcax_escdata(), rcax_superpops()

Examples

```
a <- rcax_table_query(tablename="NOSA", qlist = list(limit=5))
a[, c("popid", "spawningyear", "nosaij", "nosaej")]
```

rcax_table_sortcols *Returns the default sort columns for tables*

Description

- NOSA = c("spawningyear", "popid")
- SAR = c("outmigrationyear", "popid")
- PNI = c("spawningyear", "popid")
- JuvOut = c("outmigrationyear", "popid")
- PreSmolt = c("surveyyear", "popid")
- RperS = c("broodyear", "popid")

Usage

```
rcax_table_sortcols(hli, type = c("xport", "base"))
```

Arguments

hli	The HLI short name: NOSA, SAR, PNI, RperS, JuvOut, or PreSmolt
type	XPort table or base table

Value

A vector of column names for sorting

Examples

```
rcax_table_sortcols("NOSA")
```

`rcax_termsofuse` *Get CAX HLI data Terms of Use*

Description

Get CAX HLI data Terms of Use

Usage

```
rcax_termsofuse()
```

Value

Text of the terms of us

Examples

```
rcax_termsofuse()
```

`rcax_ua` *Make the user agent string*

Description

Make the user agent string

Usage

```
rcax_ua()
```

References

This function is modeled off `rl_ua()` in https://github.com/ropensci/rredlist

Examples

```
rcax_ua()
```

`rcax_version` *Look up the rCAX versions*

Description

Get the version installed and the version on GitHub.

Usage

```
rcax_version()
```

Value

returns the GitHub version invisibly

Examples

```
rcax_version()
```

Index