Package ‘tidywikidatar’

March 10, 2023

Title Explore 'Wikidata' Through Tidy Data Frames

Version 0.5.7

Description Query 'Wikidata' API \(<https://www.wikidata.org/wiki/Wikidata:Main_Page>\) with ease, get tidy data frames in response, and cache data in a local database.

License MIT + file LICENSE

Encoding UTF-8

RoxygenNote 7.2.3

Imports magrittr, dplyr, tidyr, WikidataR, stringr, glue, DBI, RSQLite, tibble, purrr, usethis, WikidataQueryServiceR, fs, rlang (>= 0.1.2), progress, jsonlite, pool, WikipediR, vctrs

Suggests spelling, roxygen2, testthat (>= 3.0.0), knitr, rmarkdown, odbc, assertthat

Config/testthat/edition 3

Language en-US

URL https://edjnet.github.io/tidywikidatar/

VignetteBuilder knitr

Depends R (>= 2.10)

LazyData true

NeedsCompilation no

Author Giorgio Comai [aut, cre, cph] (<https://orcid.org/0000-0002-0515-9542>), EDJNet [fnd]

Maintainer Giorgio Comai <giorgio.comai@cci.tn.it>

Repository CRAN

Date/Publication 2023-03-10 08:50:02 UTC

R topics documented:

tw_check_cache .............................................................. 4

tw_check_cached_items .................................................. 4
<table>
<thead>
<tr>
<th>R topics documented:</th>
</tr>
</thead>
<tbody>
<tr>
<td>tw_check_cache_folder</td>
</tr>
<tr>
<td>tw_check_cache_index</td>
</tr>
<tr>
<td>tw_check_qid</td>
</tr>
<tr>
<td>tw_check_search</td>
</tr>
<tr>
<td>tw_connect_to_cache</td>
</tr>
<tr>
<td>tw_create_cache_folder</td>
</tr>
<tr>
<td>tw_disable_cache</td>
</tr>
<tr>
<td>tw_disconnect_from_cache</td>
</tr>
<tr>
<td>tw_empty_image_metadata</td>
</tr>
<tr>
<td>tw_empty_item</td>
</tr>
<tr>
<td>tw_empty_qualifiers</td>
</tr>
<tr>
<td>tw_empty_search</td>
</tr>
<tr>
<td>tw_empty_wikipedia_category_members</td>
</tr>
<tr>
<td>tw_empty_wikipedia_page</td>
</tr>
<tr>
<td>tw_empty_wikipedia_page_links</td>
</tr>
<tr>
<td>tw_empty_wikipedia_page_sections</td>
</tr>
<tr>
<td>tw_enable_cache</td>
</tr>
<tr>
<td>tw_extract_qualifier</td>
</tr>
<tr>
<td>tw_extract_single</td>
</tr>
<tr>
<td>tw_filter</td>
</tr>
<tr>
<td>tw_filter_first</td>
</tr>
<tr>
<td>tw_filter_people</td>
</tr>
<tr>
<td>tw_get</td>
</tr>
<tr>
<td>tw_get_all_with_p</td>
</tr>
<tr>
<td>tw_get_cached_item</td>
</tr>
<tr>
<td>tw_get_cached_qualifiers</td>
</tr>
<tr>
<td>tw_get_cached_search</td>
</tr>
<tr>
<td>tw_get_cached_wikipedia_category_members</td>
</tr>
<tr>
<td>tw_get_cached_wikipedia_page_links</td>
</tr>
<tr>
<td>tw_get_cached_wikipedia_page_qid</td>
</tr>
<tr>
<td>tw_get_cached_wikipedia_page_sections</td>
</tr>
<tr>
<td>tw_get_cache_db</td>
</tr>
<tr>
<td>tw_get_cache_file</td>
</tr>
<tr>
<td>tw_get_cache_table_name</td>
</tr>
<tr>
<td>tw_get_description</td>
</tr>
<tr>
<td>tw_get_field</td>
</tr>
<tr>
<td>tw_get_image</td>
</tr>
<tr>
<td>tw_get_image_metadata</td>
</tr>
<tr>
<td>tw_get_image_metadata_single</td>
</tr>
<tr>
<td>tw_get_image_same_length</td>
</tr>
<tr>
<td>tw_get_label</td>
</tr>
<tr>
<td>tw_get_p1</td>
</tr>
<tr>
<td>tw_get_property</td>
</tr>
<tr>
<td>tw_get_property_description</td>
</tr>
<tr>
<td>tw_get_property_label</td>
</tr>
<tr>
<td>tw_get_property_label_single</td>
</tr>
<tr>
<td>tw_get_property_same_length</td>
</tr>
<tr>
<td>tw_get_property_with_details</td>
</tr>
</tbody>
</table>
### tw_check_cache

*Check caching status in the current session, and override it upon request*

**Description**

Mostly used internally in functions, exported for reference.

**Usage**

```r
tw_check_cache(cache = NULL)
```

**Arguments**

- `cache`  
  Defaults to NULL. If NULL, checks current cache settings. If given, returns given value, ignoring cache.

**Value**

Either TRUE or FALSE, depending on current cache settings.

**Examples**

```r
if (interactive()) {
  tw_check_cache()
}
```

### tw_check_cached_items

*Check if given items are present in cache*

**Description**

Check if given items are present in cache

**Usage**

```r
tw_check_cached_items(
  id,
  language = tidywikidatar::tw_get_language(),
  cache_connection = NULL,
  disconnect_db = TRUE
)
```
**Arguments**

- **id**: A character vector. Each element must start with Q, and correspond to a Wikidata identifier.
- **language**: Defaults to language set with `tw_set_language()`; if not set, "en". Use "all_available" to keep all languages. For available language values, see [https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all](https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all)
- **cache_connection**: Defaults to NULL. If NULL, and caching is enabled, `tidywikidatar` will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).
- **disconnect_db**: Defaults to TRUE. If FALSE, leaves the connection to cache open.

**Value**

A character vector with IDs of items present in cache. If no item found in cache, returns NULL.

**Examples**

```r
if (interactive()) {
  tw_set_cache_folder(path = tempdir())
  tw_enable_cache()
  tw_create_cache_folder(ask = FALSE)

  # add three items to local cache
  invisible(tw_get(id = "Q180099", language = "en"))
  invisible(tw_get(id = "Q228822", language = "en"))
  invisible(tw_get(id = "Q184992", language = "en"))

  # check if these other items are in cache
  items_in_cache <- tw_check_cached_items(
    id = c(
      "Q180099",
      "Q228822",
      "Q76857"
    ),
    language = "en"
  )
  # it should return only the two items from the current list of id
  # but not other item already in cache
  items_in_cache
}
```

---

**tw_check_cache_folder**  Checks if cache folder exists, if not returns an informative message

**Description**

Checks if cache folder exists, if not returns an informative message
Usage

`tw_check_cache_folder()`

Value

If the cache folder exists, returns TRUE. Otherwise throws an error.

Examples

```r
# If cache folder does not exist, it throws an error
tryCatch(tw_check_cache_folder(),
  error = function(e) {
    return(e)
  }
)

# Create cache folder
tw_set_cache_folder(path = fs::path(
  tempdir(),
  "tw_cache_folder"
))
tw_create_cache_folder(ask = FALSE)
tw_check_cache_folder()
```

---

**tw_check_cache_index**  
Check if cache table is indexed

Description

Tested only with SQLite and MySql. May work with other drivers. Used to check if given cache table is indexed (if created with any version of tidywikidatar before 0.6, they are probably not indexed and less efficient).

Usage

```r
tw_check_cache_index(
  table_name = NULL,
  type = "item",
  show_details = FALSE,
  language = tidywikidatar::tw_get_language(),
  response_language = tidywikidatar::tw_get_language(),
  cache = NULL,
  cache_connection = NULL,
  disconnect_db = TRUE
)
```
Arguments

table_name
Name of the table in the database. If given, it takes precedence over other parameters.

type
Defaults to "item". Type of cache file to output. Values typically used by tidywikidatar include "item", "search_item", "search_property", and "qualifier".

show_details
Logical, defaults to FALSE. If FALSE, return a logical vector of length one (TRUE if the table was indexed, FALSE if it was not). If TRUE, returns a data frame with more details about the index.

language
Defaults to language set with `tw_set_language()``en``if not set. Used to limit the data to be cached. Use "all_available" to keep all data. For available values, see [https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all](https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all)

response_language
Defaults to language set with `tw_set_language()``en``if not set. Relevant only when type is set to "search_item" or "search_property". See `tw_search()` for details.

cache
Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.

cache_connection
Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

disconnect_db
Defaults to TRUE. If FALSE, leaves the connection to cache open.

Value

If `show_details` is set to FALSE, return a logical vector of length one (TRUE if the table was indexed, FALSE if it was not). If `show_details` is set to TRUE, returns a data frame with more details about the index.

Examples

```r
if (interactive()) {
  tw_enable_cache()
  tw_set_cache_folder(path = fs::path(
    fs::path_home_r(),
    "R",
    "tw_data"
  ))
  tw_set_language(language = "en")
  tw_check_cache_index()
}
```
tw_check_qid

Ensures that input appears to be a valid Wikidata id

Description

Mostly used internally by other functions.

Usage

```r
tw_check_qid(id, logical_vector = FALSE, non_id_as_NA = FALSE)
```

Arguments

- `id`: A character vector of one or more Wikidata ids.
- `logical_vector`: Logical, defaults to FALSE. If TRUE, returns a logical vector of the same length as input, where TRUE corresponds to seemingly meaningful Q identifiers.
- `non_id_as_NA`: Logical, defaults to FALSE. If TRUE (and if `logical_vector` is set to FALSE), a vector of the same length is returned, with NA replacing items that are seemingly not meaningful Q identifiers.

Value

A character vector with only strings appearing to be Wikidata identifiers; possibly shorter than input

Examples

```r
tw_check_qid(id = c("Q180099", "q228822", "Not an id", "00180099", NA, "Q5"))

tw_check_qid(
  id = c("Q180099", "q228822", "Not an id", "00180099", NA, "Q5"),
  logical_vector = TRUE)

tw_check_qid(
  id = c("Q180099", "q228822", "Not an id", "00180099", NA, "Q5"),
  non_id_as_NA = TRUE)
```
**tw_check_search** Checks if an input is a search; if not, it tries to return a search

**Description**

Mostly used as a convenience function inside other functions to have consistent inputs.

**Usage**

```r
tw_check_search(
  search,
  type = "item",
  language = tidywikidatar::tw_get_language(),
  limit = 10,
  include_search = FALSE,
  wait = 0,
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE
)
```

**Arguments**

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>search</td>
<td>A string to be searched in Wikidata</td>
</tr>
<tr>
<td>type</td>
<td>Defaults to &quot;item&quot;. Either &quot;item&quot; or &quot;property&quot;.</td>
</tr>
</tbody>
</table>
| language          | Language to be used for the search. Can be set once per session with `tw_set_language()`.
|                   | If not set, defaults to "en". For a full list, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all |
| limit             | Maximum numbers of responses to be given                                    |
| include_search    | Logical, defaults to FALSE. If TRUE, the search is returned as an additional column. |
| wait              | In seconds, defaults to 0. Time to wait between queries to Wikidata. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries. |
| cache             | Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`. |
| overwrite_cache   | Defaults to FALSE. If TRUE, overwrites cache.                               |
| cache_connection  | Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details). |
| disconnect_db     | Defaults to TRUE. If FALSE, leaves the connection to cache open.             |
tw_connect_to_cache

Value
A data frame with three columns, id, label, and description, filtered by the above criteria.

Examples
# The following two lines should give the same result.

tw_check_search("Sylvia Pankhurst")
tw_check_search(tw_search("Sylvia Pankhurst"))

tw_connect_to_cache
Return a connection to be used for caching

Description
Return a connection to be used for caching

Usage
tw_connect_to_cache(
  connection = NULL,
  RSQLite = NULL,
  language = tidywikidatar::tw_get_language(),
  cache = NULL
)

Arguments

collection Defaults to NULL. If NULL, uses local SQLite database. If given, must be a connection object or a list with relevant connection settings (see example).

RSQLite Defaults to NULL, expected either NULL or logical. If set to FALSE, details on the database connection must be given either as a named list in the connection parameter, or with tw_set_cache_db() as environment variables.

language Defaults to language set with tw_set_language(); if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all

cache Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with tw_enable_cache() or tw_disable_cache().

Value
A connection object.
Examples

```r
if (interactive()) {
  cache_connection <- pool::dbPool(
    RSQLite::SQLite(), # or e.g. odbc::odbc(),
    Driver = "::memory:", # or e.g. "MariaDB",
    Host = "localhost",
    database = "example_db",
    UID = "example_user",
    PWD = "example_pwd"
  )
  tw_connect_to_cache(cache_connection)
}

db_settings <- list(
  driver = "MySQL",
  host = "localhost",
  server = "localhost",
  port = 3306,
  database = "tidywikidatar",
  user = "secret_username",
  pwd = "secret_password"
)

tw_connect_to_cache(db_settings)
}
```

---

tw_create_cache_folder

*Creates the base cache folder where tidywikidatar caches data.*

Description

Creates the base cache folder where tidywikidatar caches data.

Usage

```r
tw_create_cache_folder(ask = TRUE)
```

Arguments

- **ask** Logical, defaults to TRUE. If FALSE, and cache folder does not exist, it just creates it without asking (useful for non-interactive sessions).

Value

Nothing, used for its side effects.
Examples

if (interactive()) {
  tw_create_cache_folder()
}

tw_disable_cache  Disable caching for the current session

Description
Disable caching for the current session

Usage
tw_disable_cache()

Value
Nothing, used for its side effects.

Examples

if (interactive()) {
  tw_disable_cache()
}

tw_disconnect_from_cache

Ensure that connection to cache is disconnected consistently

Description
Ensure that connection to cache is disconnected consistently

Usage
tw_disconnect_from_cache(
  cache = NULL,
  cache_connection = NULL,
  disconnect_db = TRUE,
  language = tidywikidatar::tw_get_language()
)
Arguments

**cache**
Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.

**cache_connection**
Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

**disconnect_db**
Defaults to TRUE. If FALSE, leaves the connection to cache open.

**language**
Defaults to language set with `tw_set_language()`; if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all

Value

Nothing, used for its side effects.

Examples

```r
if (interactive()) {
  tw_get(
    id = c("Q180099"),
    language = "en"
  )
  tw_disconnect_from_cache()
}
```

---

**tw_empty_image_metadata**

_A zero-rows tibble used internally when `tw_get_image_metadata()` would not return any value._

---

Description

A zero-rows tibble used internally when `tw_get_image_metadata()` would not return any value.

Usage

`tw_empty_image_metadata`

Format

A data frame with 0 rows and 19 columns
A zero-rows tibble used internally when `tw_get()` would not return any value.

**Usage**

`tw_empty_item`

**Format**

A data frame with 0 rows and 3 columns

---

A zero-rows tibble used internally when `tw_get_qualifiers()` would not return any value.

**Usage**

`tw_empty_qualifiers`

**Format**

A data frame with 0 rows and 8 columns

---

A zero-rows tibble used internally when `tw_search()` would not return any value.

**Usage**

`tw_empty_search`

**Format**

A data frame with 0 rows and 4 columns
**tw_empty_wikipedia_category_members**

A zero-rows tibble used internally when `tw_empty_wikipedia_category_members()` would not return any value.

**Description**
A zero-rows tibble used internally when `tw_empty_wikipedia_category_members()` would not return any value.

**Usage**
`tw_empty_wikipedia_category_members`

**Format**
A data frame with 0 rows and 3 columns

---

**tw_empty_wikipedia_page**

A zero-rows tibble used internally when `tw_get_wikipedia_page_qid()` would not return any value.

**Description**
A zero-rows tibble used internally when `tw_get_wikipedia_page_qid()` would not return any value.

**Usage**
`tw_empty_wikipedia_page`

**Format**
A data frame with 0 rows and 6 columns
**tw_empty_wikipedia_page_links**

A zero-rows tibble used internally when `tw_get_wikipedia_page_links()` would not return any value.

---

**Description**

A zero-rows tibble used internally when `tw_get_wikipedia_page_links()` would not return any value.

**Usage**

`tw_empty_wikipedia_page_links`

**Format**

A data frame with 0 rows and 8 columns

---

**tw_empty_wikipedia_page_sections**

A zero-rows tibble used internally when `tw_get_wikipedia_page_sections()` would not return any value.

---

**Description**

A zero-rows tibble used internally when `tw_get_wikipedia_page_sections()` would not return any value.

**Usage**

`tw_empty_wikipedia_page_sections`

**Format**

A data frame with 0 rows and 8 columns
tw_enable_cache

Enable caching for the current session

Description
Enable caching for the current session

Usage
tw_enable_cache(SQLite = TRUE)

Arguments
SQLite Logical, defaults to TRUE. Set to FALSE to use custom database options. See tw_set_cache_db() for details.

Value
Nothing, used for its side effects.

Examples

if (interactive()) {
  tw_enable_cache()
}

tw_extract_qualifier

Extract qualifiers from an object of class Wikidata created with WikidataR

Description
This function is mostly used internally and for testing.

Usage
tw_extract_qualifier(id, p, w = NULL)

Arguments
id A character vector of length 1, must start with Q, e.g. "Q254" for Wolfgang Amadeus Mozart.
p A character vector of length 1, a property. Must always start with the capital letter "P", e.g. "P31" for "instance of".
w An object of class Wikidata created with WikidataR, typically created with WikidataR::get_item(id = id)
Value

A data frame (a tibble) with eight columns: id for the input id, property, qualifier_id, qualifier_property, qualifier_value, rank, qualifier_value_type, and set (to distinguish sets of data when a property is present more than once)

Examples

```r
w <- WikidataR::get_item(id = "Q180099")
tw_extract_qualifier(id = "Q180099", p = "P26", w = w)
```

Description

This function is mostly used internally and for testing.

Usage

```r
tw_extract_single(w, language = tidywikidatar::tw_get_language())
```

Arguments

- **w**: An object of class Wikidata created with WikidataR, typically created with WikidataR::get_item(id = id)
- **language**: Defaults to language set with tw_set_language(); if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all

Value

A data frame (a tibble) with four columns, such as the one created by tw_get.

Examples

```r
item <- tryCatch(WikidataR::get_item(id = "Q180099"),
                 error = function(e) {
                   as.character(e[[1]])
                 })

tidywikidatar::tw_extract_single(w = item)
```
**tw_filter**

Filter search result and keep only items with matching property and Q identifier

**Description**

Filter search result and keep only items with matching property and Q identifier

**Usage**

```r
tw_filter(  
  search,  
  p,  
  q,  
  language = tidywikidatar::tw_get_language(),  
  limit = 10,  
  include_search = FALSE,  
  wait = 0,  
  cache = NULL,  
  overwrite_cache = FALSE,  
  cache_connection = NULL,  
  disconnect_db = TRUE  
)
```

**Arguments**

- **search**  
  A data frame generated by `tw_search()`, or a search query. If a data frame is given, language and limits are ignore.

- **p**  
  A character vector of length 1, a property. Must always start with the capital letter "P", e.g. "P31" for "instance of".

- **q**  
  A character vector of length 1, a wikidata id. Must always start with the capital letter "Q", e.g. "Q5" for "human being".

- **language**  
  Language to be used for the search. Can be set once per session with `tw_set_language()`. If not set, defaults to "en". For a full list, see [https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all](https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all).

- **limit**  
  Maximum numbers of responses to be given.

- **include_search**  
  Logical, defaults to FALSE. If TRUE, the search is returned as an additional column.

- **wait**  
  In seconds, defaults to 0. Time to wait between queries to Wikidata. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.

- **cache**  
  Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.

- **overwrite_cache**  
  Defaults to FALSE. If TRUE, overwrites cache.
cache_connection

Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

disconnect_db

Defaults to TRUE. If FALSE, leaves the connection to cache open.

Value

A data frame with three columns, id, label, and description, filtered by the above criteria.

Examples

```r
tw_search(search = "Margaret Mead", limit = 3) %>%
tw_filter(p = "P31", q = "Q5")
```

Description

Same as `tw_filter()`, but consistently returns data frames with a single row.

Usage

```r
tw_filter_first(
  search,
  p,
  q,
  language = tidywikidatar::tw_get_language(),
  limit = 10,
  include_search = FALSE,
  wait = 0,
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE
)
```

Arguments

- **search**
  A data frame generated by `tw_search()`, or a search query. If a data frame is given, language and limits are ignore.

- **p**
  A character vector of length 1, a property. Must always start with the capital letter "P", e.g. "P31" for "instance of".

- **q**
  A character vector of length 1, a wikidata id. Must always start with the capital letter "Q", e.g. "Q5" for "human being".

- **language**
  Language to be used for the search.
tw_filter_people

```r
limit
include_search
wait
cache
overwrite_cache
cache_connection
disconnect_db

Value
A data frame with one row and three columns, id, label, and description, filtered by the above criteria.

Examples
```r
tw_search("Margaret Mead") %>%
  tw_filter_first(p = "P31", q = "Q5")
```r

Description
A wrapper of `tw_filter()` that defaults to keep only "instance of" (P31) "human being" (Q5).

Usage
```r
tw_filter_people(
  search,
  language = tidywikidatar::tw_get_language(),
  limit = 10,
  stop_at_first = TRUE,
  wait = 0,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE
)
Arguments

- **search**: A data frame generated by `tw_search()`, or a search query. If a data frame is given, language and limits are ignored.
- **language**: Language to be used for the search.
- **limit**: Maximum numbers of responses to be given.
- **stop_at_first**: Logical, defaults to `TRUE`. If `TRUE`, returns only the first match from the search that satisfies the criteria.
- **wait**: In seconds, defaults to 0. Time to wait between queries to Wikidata. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.
- **overwrite_cache**: Defaults to `FALSE`. If `TRUE`, overwrites cache.
- **cache_connection**: Defaults to `NULL`. If `NULL`, and caching is enabled, `tidywikidatar` will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).
- **disconnect_db**: Defaults to `TRUE`. If `FALSE`, leaves the connection to cache open.

Value

A data frame with three columns, `id`, `label`, and `description`; all rows refer to a human being.

Examples

```r
tw_search("Ruth Benedict")
```

```r
tw_search("Ruth Benedict") %>%
  tw_filter_people()
```

Description

Return (most) information from a Wikidata item in a tidy format

Usage

```r
tw_get(
  id,
  language = tidywikidatar::tw_get_language(),
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE,
  wait = 0,
  id_l = NULL
)
```
tw_get

Arguments

id  A character vector, must start with Q, e.g. "Q180099" for the anthropologist Margaret Mead. Can also be a data frame of one row, typically generated with \texttt{tw_search()} or a combination of \texttt{tw_search()} and \texttt{tw_filter_first()}.

language  Defaults to language set with \texttt{tw_set_language()}; if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all.

cache  Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with \texttt{tw_enable_cache()} or \texttt{tw_disable_cache()}.

overwrite_cache  Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.

cache_connection  Defaults to NULL. If NULL, and caching is enabled, \texttt{tidywikidatar} will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

disconnect_db  Defaults to TRUE. If FALSE, leaves the connection to cache open.

wait  In seconds, defaults to 0. Time to wait between queries to Wikidata. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.

id_l  Defaults to NULL. If given, must be an object or list such as the one generated with \texttt{WikidataR::get_item()}. If given, and the requested id is actually present in \texttt{id_l}, then no query to Wikidata servers is made.

Value  

A data.frame (a tibble) with three columns (id, property, and value).

Examples

```r
if (interactive()) {
  tw_get(
    id = c("Q180099", "Q228822"),
    language = "en"
  )
}
```

## using \texttt{\textquotesingle tw_test_items\textquotesingle} in examples in order to show output without calling
## on Wikidata servers

```r
tw_get(
  id = c("Q180099", "Q228822"),
  language = "en",
  id_l = tw_test_items
)
```
tw_get_all_with_p

Get all items that have a given property (irrespective of the value)

Description

This function does not cache results.

Usage

tw_get_all_with_p(
  p,
  fields = c("item", "itemLabel", "itemDescription"),
  language = tidywikidatar::tw_get_language(),
  method = "SPARQL",
  wait = 0.1,
  limit = Inf,
  return_as_tw_search = TRUE
)

Arguments

p
A character vector, a property. Must always start with the capital letter "P", e.g. "P31" for "instance of".

fields
A character vector of Wikidata fields. Ignored if return_as_tw_search is set to TRUE (as per default). Defaults to ("item", "itemLabel", "itemDescription")

language
Defaults to language set with tw_set_language(); if not set, "en". If more than one, can be set in order of preference, e.g. c("it", "fr", "en"). Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all

method
Defaults to "SPARQL". The only accepted alternative value is "JSON", to use instead json-based API.

wait
Defaults to 0.1. Used only in method is set to "JSON".

limit
Defaults to Inf. Set to smaller values for testing and cache locally when possible to reduce load on servers.

return_as_tw_search
Logical, defaults to TRUE. If TRUE, returns a data frame with three columns (id, label, and description) that can be piped to other tw_ functions. If FALSE, a data frame with as many columns as fields.

Value

A data frame with three columns is method is set to "SPARQL", or as many columns as fields if more are given and return_as_tw_search is set to FALSE. A single column with Wikidata identifier if method is set to "JSON".
Examples

```r
if (interactive()) {
  # get all Wikidata items with an ICAO airport code ("P239")
  tw_get_all_with_p(p = "P239", limit = 10)
}
```

**tw_get_cached_item**  
*Retrieve cached item*

**Description**

Retrieve cached item

**Usage**

```r
tw_get_cached_item(
  id,
  language = tidywikidatar::tw_get_language(),
  cache = NULL,
  cache_connection = NULL,
  disconnect_db = TRUE
)
```

**Arguments**

- **id**
  - A character vector, must start with Q, e.g. "Q180099" for the anthropologist Margaret Mead. Can also be a data frame of one row, typically generated with `tw_search()` or a combination of `tw_search()` and `tw_filter_first()`.

- **language**
  - Defaults to language set with `tw_set_language()`; if not set, "en". Use "all_available" to keep all languages. For available language values, see [https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all](https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all).

- **cache**
  - Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.

- **cache_connection**
  - Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

- **disconnect_db**
  - Defaults to TRUE. If FALSE, leaves the connection open.

**Value**

If data present in cache, returns a data frame with cached data.
tw_getCachedQualifiers

Retrieve cached qualifier

**Usage**

tw_getCachedQualifiers(
    id,
    p,
    language = tidywikidatar::tw_get_language(),
    cache = NULL,
    cache_connection = NULL,
    disconnect_db = TRUE
)

**Arguments**

- **id**  A character vector, must start with Q, e.g. "Q180099" for the anthropologist Margaret Mead. Can also be a data frame of one row, typically generated with `tw_search()` or a combination of `tw_search()` and `tw_filter_first()`.
- **p**  A character vector of length 1, a property. Must always start with the capital letter "P", e.g. "P31" for "instance of".
- **language**  Defaults to language set with `tw_set_language()`; if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all.
- **cache**  Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.
tw_get_cached_search

Usage

```
tw_get_cached_search(
  search,
  type = "item",
  language = tidywikidatar::tw_get_language(),
  response_language = tidywikidatar::tw_get_language(),
  cache = NULL,
  include_search = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE
)
```

Value

If data present in cache, returns a data frame with cached data.

Examples

```
tw_set_cache_folder(path = tempdir())
tw_enable_cache()
tw_create_cache_folder(ask = FALSE)

df_from_api <- tw_get_qualifiers(id = "Q180099", p = "P26", language = "en")

df_from_cache <- tw_get_cached_qualifiers(
  id = "Q180099",
  p = "P26",
  language = "en"
)

df_from_cache
```
Arguments

search  A string to be searched in Wikidata

type     Defaults to "item". Either "item" or "property".

language Language to be used for the search. Can be set once per session with tw_set_language(). If not set, defaults to "en''. For a full list, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes

response_language Language to be used for the returned labels and descriptions. Corresponds to the useLang parameter of the MediaWiki API: https://www.wikidata.org/w/api.php?action=help&modules=wbsearchentities. Can be set once per session with tw_set_language(). If not set, defaults to "en''. For a full list, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes

cache Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with tw_enable_cache() or tw_disable_cache().

include_search Logical, defaults to FALSE. If TRUE, the search is returned as an additional column.

cache_connection Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

disconnect_db Defaults to TRUE. If FALSE, leaves the connection to cache open.

Value

If data present in cache, returns a data frame with cached data.

Examples

    tw_set_cache_folder(path = tempdir())
    tw_enable_cache()
    tw_create_cache_folder(ask = FALSE)

    search_from_api <- tw_search("Sylvia Pankhurst")
    search_from_api

    df_from_cache <- tw_get_cached_search("Sylvia Pankhurst")
    df_from_cache

    tw_getCachedWikipediaCategoryMembers

    Gets members of Wikipedia categories from local cache

Description

    Mostly used internally.
Usage

tw_get_cached_wikipedia_category_members(
    category,
    type = "page",
    language = tidywikidatar::tw_get_language(),
    cache = NULL,
    cache_connection = NULL,
    disconnect_db = TRUE
)

Arguments

category Title of a Wikipedia category page or final parts of its url. Must include "Category:“, or equivalent in other languages. If given, url can be left empty, but language must be provided.

type Defaults to "page", defines which kind of members of a category are returned. Valid values include "page", "file", and "subcat" (for sub-category). Corresponds to cmtype. For details, see https://www.mediawiki.org/wiki/API:Categorymembers

language Two-letter language code used to define the Wikipedia version to use. Defaults to language set with tw_set_language(); if not set, "en". If url given, this can be left empty.

cache Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with tw_enable_cache() or tw_disable_cache().

cache_connection Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

disconnect_db Defaults to TRUE. If FALSE, leaves the connection to cache open.

Value

If data present in cache, returns a data frame with cached data.

Examples

if (interactive()) {
    tw_set_cache_folder(path = tempdir())
    tw_enable_cache()
    tw_create_cache_folder(ask = FALSE)

    df_from_api <- tw_get_wikipedia_page_qid(category = "Margaret Mead", language = "en")

    df_from_cache <- tw_get_cached_wikipedia_category_members(
        category = "Margaret Mead",
        language = "en"
    )
}
Description

Mostly used internally.

Usage

tw_get_cached_wikipedia_page_links(
    title,
    language = tidywikidatar::tw_get_language(),
    cache = NULL,
    cache_connection = NULL,
    disconnect_db = TRUE
)

Arguments

title Title of a Wikipedia page or final parts of its url. If given, url can be left empty, but language must be provided.

language Defaults to language set with tw_set_language(); if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all

cache Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with tw_enable_cache() or tw_disable_cache().

cache_connection Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

disconnect_db Defaults to TRUE. If FALSE, leaves the connection open.

Value

If data present in cache, returns a data frame with cached data.

Examples

if (interactive()) {
  tw_set_cache_folder(path = tempdir())
  tw_enable_cache()
  tw_create_cache_folder(ask = FALSE)

df_from_api <- tw_get_wikipedia_page_qid(title = "Margaret Mead", language = "en")
}
**tw_get_cached_wikipedia_page_qid**

```

df_from_cache <- tw_get_cached_wikipedia_page_links(  
  title = "Margaret Mead",  
  language = "en"
)

df_from_cache

```

---

**tw_get_cached_wikipedia_page_qid**

*Gets id of Wikipedia pages from local cache*

**Description**

Mostly used internally.

**Usage**

```
tw_get_cached_wikipedia_page_qid(  
  title,  
  language = tidywikidatar::tw_get_language(),  
  cache = NULL,  
  cache_connection = NULL,  
  disconnect_db = TRUE
)
```

**Arguments**

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>title</code></td>
<td>Title of a Wikipedia page or final parts of its url. If given, url can be left empty, but language must be provided.</td>
</tr>
<tr>
<td><code>language</code></td>
<td>Defaults to language set with <code>tw_set_language()</code>; if not set, &quot;en&quot;. Use &quot;all_available&quot; to keep all languages. For available language values, see <a href="https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all">https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all</a></td>
</tr>
<tr>
<td><code>cache</code></td>
<td>Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with <code>tw_enable_cache()</code> or <code>tw_disable_cache()</code>.</td>
</tr>
<tr>
<td><code>cache_connection</code></td>
<td>Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).</td>
</tr>
<tr>
<td><code>disconnect_db</code></td>
<td>Defaults to TRUE. If FALSE, leaves the connection open.</td>
</tr>
</tbody>
</table>

**Value**

If data present in cache, returns a data frame with cached data.
Examples

if (interactive()) {
  tw_set_cache_folder(path = tempdir())
  tw_enable_cache()
  tw_create_cache_folder(ask = FALSE)

  df_from_api <- tw_get_wikipedia_page_qid(title = "Margaret Mead", language = "en")

  df_from_cache <- tw_get_cached_wikipedia_page_qid(
    title = "Margaret Mead",
    language = "en"
  )

  df_from_cache
}

---

tw_get_cached_wikipedia_page_sections

*Gets sections of Wikipedia pages from local cache*

Description

Mostly used internally.

Usage

```r
  tw_get_cached_wikipedia_page_sections(
    title,
    language = tidywikidatar::tw_get_language(),
    cache = NULL,
    cache_connection = NULL,
    disconnect_db = TRUE
  )
```

Arguments

title Title of a Wikipedia page or final parts of its url. If given, url can be left empty, but language must be provided.

language Defaults to language set with `tw_set_language()`; if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all

cache Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.

cache_connection Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

disconnect_db Defaults to TRUE. If FALSE, leaves the connection open.
If data present in cache, returns a data frame with cached data.

Examples

```r
if (interactive()) {
  tw_set_cache_folder(path = tempdir())
  tw_enable_cache()
  tw_create_cache_folder(ask = FALSE)

  df_from_api <- tw_get_wikipedia_page_qid(title = "Margaret Mead", language = "en")

  df_from_cache <- tw_get_cached_wikipedia_page_sections(
    title = "Margaret Mead",
    language = "en"
  )

  df_from_cache
}
```

---

**tw_get_cache_db**

Get database connection settings from the environment

**Description**

Typically set with `tw_set_cache_db()`

**Usage**

```r
tw_get_cache_db()
```

**Value**

A list with all database parameters as stored in environment variables.

**Examples**

```r
tw_get_cache_db()
```
\begin{itemize}
  \item \texttt{tw_get_cache_file} \textit{Gets location of cache file}
  \item \texttt{tw_get_cache_table_name} \textit{Gets name of table inside the database}
\end{itemize}

**Description**

- \texttt{tw_get_cache_file}: Gets location of cache file
- \texttt{tw_get_cache_table_name}: Gets name of table inside the database

**Usage**

- \texttt{tw_get_cache_file(type = NULL, language = tidywikidatar::tw_get_language())}
- \texttt{tw_get_cache_table_name(type = "item", language = tidywikidatar::tw_get_language(), response_language = tidywikidatar::tw_get_language())}

**Arguments**

- **type**: Defaults to NULL. Deprecated. If given, type of cache file to output. Values typically used by tidywikidatar in versions up to 4.2 include "item", "search", and "qualifier".
- **language**: Defaults to language set with \texttt{tw_set_language()}; if not set, "en". Use "all_available" to keep all languages. For available language values, see \url{https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all}

**Value**

- A character vector of length one with location of item cache file.

**Examples**

```r
  tw_set_cache_folder(path = tempdir())
  sqlite_cache_file_location <- tw_get_cache_file() # outputs location of cache file
  tw_get_cache_table_name
```

- A character vector of length one with location of item cache file.
Arguments

- **type**: Defaults to "item". Type of cache file to output. Values typically used by tidywikidatar include "item", "search_item", "search_property", and "qualifier".
- **language**: Defaults to language set with `tw_set_language()`; "en" if not set. Used to limit the data to be cached. Use "all_available" to keep all data. For available values, see [https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all](https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all)
- **response_language**: Defaults to language set with `tw_set_language()`; "en" if not set. Relevant only when type is set to "search_item" or "search_property". See `tw_search()` for details.

Value

A character vector of length one with the name of the relevant table in the cache file.

Examples

```r
# outputs name of table used in the cache database
tw_get_cache_table_name(type = "item", language = "en")
```

**Description**

Get Wikidata description in given language

**Usage**

```r
tw_get_description(
id,
language = tidywikidatar::tw_get_language(),
id_df = NULL,
cache = NULL,
overwrite_cache = FALSE,
cache_connection = NULL,
disconnect_db = TRUE,
wait = 0
)
```

**Arguments**

- **id**: A character vector, must start with Q, e.g. "Q254" for Wolfgang Amadeus Mozart
- **language**: Defaults to language set with `tw_set_language()`; if not set, "en". Use "all_available" to keep all languages. For available language values, see [https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all](https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all)
### tw_get_field

**Description**

Gets a field such a label or description from a dataframe typically generated with `tw_get()`

**Usage**

```r
tw_get_field(df, field, id, language = tidywikidatar::tw_get_language())
```
**tw_get_image**

Get image from Wikimedia Commons

**Arguments**

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>df</td>
<td>A data frame typically generated with tw_get(). It should include data for the id included in the dedicated parameter.</td>
</tr>
<tr>
<td>field</td>
<td>A character vector of length one. Typically, either &quot;label&quot; or &quot;description&quot;.</td>
</tr>
<tr>
<td>id</td>
<td>A character vector, typically of Wikidata identifiers. The output will be of the same length and in the same order as the identifiers provided with this parameter.</td>
</tr>
<tr>
<td>language</td>
<td>Defaults to language set with tw_set_language(); if not set, &quot;en&quot;. Use &quot;all_available&quot; to keep all languages. For available language values, see <a href="https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all">https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all</a></td>
</tr>
</tbody>
</table>

**Value**

A character vector of the same length, and with data in the same order, as id.

**Examples**

```r
tw_get("Q180099") %>%
tw_get_field(field = "label", id = "Q180099")
```

**Description**

Please consult the relevant documentation for reusing content outside Wikimedia: https://commons.wikimedia.org/wiki/Commons:Reusing_content_outside_Wikimedia/technical

**Usage**

```r
tw_get_image(  
  id,  
  format = "filename",  
  width = NULL,  
  language = tidywikidatar::tw_get_language(),  
  id_df = NULL,  
  cache = NULL,  
  overwrite_cache = FALSE,  
  cache_connection = NULL,  
  disconnect_db = TRUE,  
  wait = 0  
)
```
Arguments

id
A character vector of length 1, must start with Q, e.g. "Q254" for Wolfgang Amadeus Mozart.

format
A character vector, defaults to 'filename'. If set to 'commons', outputs the link to the Wikimedia Commons page. If set to 'embed', outputs a link that can be used to embed.

width
A numeric value, defaults to NULL, relevant only if format is set to 'embed'. If not given, defaults to full resolution image.

language
Needed for caching, defaults to language set with `tw_set_language()`; if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all

id_df
Default to NULL. If given, it should be a dataframe typically generated with `tw_get_()`, and is used instead of calling Wikidata or using SQLite cache. Ignored when id is of length more than one.

cache
Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.

overwrite_cache
Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.

cache_connection
Defaults to NULL. If NULL, and caching is enabled, tidywikidata will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

disconnect_db
Defaults to TRUE. If FALSE, leaves the connection to cache open.

wait
In seconds, defaults to 0. Time to wait between queries to Wikidata. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.

Value

A data frame of two columns, id and image, corresponding to reference to the image in the requested format.

Examples

tw_get_image("Q180099",
  format = "filename"
)

if (interactive()) {
  tw_get_image("Q180099",
    format = "commons"
  )
}

tw_get_image("Q180099",
  format = "embed",
  width = 300
)
tw_get_image_metadata

Get metadata for images from Wikimedia Commons

Description

Please consult the relevant documentation for reusing content outside Wikimedia: https://commons.wikimedia.org/wiki/Commons:Reusing_content_outside_Wikimedia/technical

Usage

tw_get_image_metadata(id,
  id, image_filename = NULL,
  only_first = TRUE,
  language = tidywikidatar::tw_get_language(),
  id_df = NULL,
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE,
  wait = 1,
  attempts = 10)

Arguments

id A character vector of length 1, must start with Q, e.g. "Q254" for Wolfgang Amadeus Mozart.
image_filename Defaults to NULL. If NULL, image_filename is obtained from the Wikidata id. If given, must be of the same length as id.
only_first Defaults to TRUE. If TRUE, returns metadata only for the first image associated with a given Wikidata id. If FALSE, returns all images available.
language Needed for caching, defaults to language set with tw_set_language(); if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all
id_df Default to NULL. If given, it should be a dataframe typically generated with tw_get_(), and is used instead of calling Wikidata or using SQLite cache. Ignored when id is of length more than one.
cache Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with tw_enable_cache() or tw_disable_cache().
overwrite_cache Logical, defaults to FALSE. If TRUE, it overwrites the table in the local SQLite database. Useful if the original Wikidata object has been updated.
cache_connection
Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

disconnect_db
Defaults to TRUE. If FALSE, leaves the connection to cache open.

wait
In seconds, defaults to 1. Time to wait between queries to the APIs. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.

attempts
Defaults to 10. Number of times it re-attempts to reach the API before failing.

Value
A character vector, corresponding to reference to the image in the requested format.

Examples

if (interactive()) {
  tw_get_image_metadata("Q180099")
}

---

tw_get_image_metadata_single
Get metadata for images from Wikimedia Commons

Description
Please consult the relevant documentation for reusing content outside Wikimedia: https://commons.wikimedia.org/wiki/Commons:Reusing_content_outside_Wikimedia/technical

Usage

tw_get_image_metadata_single(
  id,
  image_filename = NULL,
  only_first = TRUE,
  language = tidywikidatar::tw_get_language(),
  id_df = NULL,
  cache = NULL,
  overwrite_cache = FALSE,
  read_cache = TRUE,
  cache_connection = NULL,
  disconnect_db = TRUE,
  wait = 1,
  attempts = 10
)
**Arguments**

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>id</td>
<td>A character vector of length 1, must start with Q, e.g. &quot;Q254&quot; for Wolfgang Amadeus Mozart.</td>
</tr>
<tr>
<td>image_filename</td>
<td>Defaults to NULL. If NULL, image_filename is obtained from the Wikidata id. If given, must be of the same length as id.</td>
</tr>
<tr>
<td>only_first</td>
<td>Defaults to TRUE. If TRUE, returns metadata only for the first image associated with a given Wikidata id. If FALSE, returns all images available.</td>
</tr>
<tr>
<td>language</td>
<td>Needed for caching, defaults to language set with <code>tw_set_language()</code>; if not set, &quot;en&quot;. Use &quot;all_available&quot; to keep all languages. For available language values, see <a href="https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all">https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all</a></td>
</tr>
<tr>
<td>id_df</td>
<td>Default to NULL. If given, it should be a dataframe typically generated with <code>tw_get_()</code>, and is used instead of calling Wikidata or using SQLite cache. Ignored when id is of length more than one.</td>
</tr>
<tr>
<td>cache</td>
<td>Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with <code>tw_enable_cache()</code> or <code>tw_disable_cache()</code>.</td>
</tr>
<tr>
<td>overwrite_cache</td>
<td>Logical, defaults to FALSE. If TRUE, it overwrites the table in the local SQLite database. Useful if the original Wikidata object has been updated.</td>
</tr>
<tr>
<td>read_cache</td>
<td>Logical, defaults to TRUE. Mostly used internally to prevent checking if an item is in cache if it is already known that it is not in cache.</td>
</tr>
<tr>
<td>cache_connection</td>
<td>Defaults to NULL. If NULL, and caching is enabled, <code>tidywikidatar</code> will use a local SQLite database. A custom connection to other databases can be given (see vignette caching for details).</td>
</tr>
<tr>
<td>disconnect_db</td>
<td>Defaults to TRUE. If FALSE, leaves the connection to cache open.</td>
</tr>
<tr>
<td>wait</td>
<td>In seconds, defaults to 1. Time to wait between queries to the APIs. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.</td>
</tr>
<tr>
<td>attempts</td>
<td>Defaults to 10. Number of times it re-attempts to reach the API before failing.</td>
</tr>
</tbody>
</table>

**Value**

A character vector, corresponding to reference to the image in the requested format.

**Examples**

```r
if (interactive()) {
  tw_get_image_metadata_single("Q180099")
}
```
tw_get_image_same_length

Get image from Wikimedia Commons

Description

Please consult the relevant documentation for reusing content outside Wikimedia: https://commons.wikimedia.org/wiki/Commons:Reusing_content_outside_Wikimedia/technical

Usage

tw_get_image_same_length(
  id,
  format = "filename",
  as_tibble = FALSE,
  only_first = TRUE,
  width = NULL,
  language = tidywikidatar::tw_get_language(),
  id_df = NULL,
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE,
  wait = 0
)

Arguments

id A character vector of length 1, must start with Q, e.g. "Q254" for Wolfgang Amadeus Mozart.

format A character vector, defaults to 'filename'. If set to 'commons', outputs the link to the Wikimedia Commons page. If set to 'embed', outputs a link that can be used to embed.

as_tibble Defaults to FALSE. If TRUE, returns a data frame instead of a character vector.

only_first Defaults to TRUE. If TRUE, returns only the first image associated with a given Wikidata id. If FALSE, returns all images available.

width A numeric value, defaults to NULL, relevant only if format is set to 'embed'. If not given, defaults to full resolution image.

language Needed for caching, defaults to language set with tw_set_language(); if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all

id_df Default to NULL. If given, it should be a dataframe typically generated with tw_get_( ), and is used instead of calling Wikidata or using SQLite cache. Ignored when id is of length more than one.

cache Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with tw_enable_cache() or tw_disable_cache().
overwrite_cache

Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.

(cache_connection)

Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

disconnect_db

Defaults to TRUE. If FALSE, leaves the connection to cache open.

wait

In seconds, defaults to 0. Time to wait between queries to Wikidata. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.

Value

A character vector, corresponding to reference to the image in the requested format.

Examples

tw_get_image_same_length("Q180099",
  format = "filename"
)

if (interactive()) {
  tw_get_image_same_length("Q180099",
    format = "commons"
  )
  
  tw_get_image_same_length("Q180099",
    format = "embed",
    width = 300
  )
}

---

tw_get_label

Get Wikidata label in given language

Description

Get Wikidata label in given language

Usage

tw_get_label(
  id,
  language = tidywikidatar::tw_get_language(),
  id_df = NULL,
  cache = NULL,
  overwrite_cache = FALSE,
)
Arguments

id A character vector, must start with Q, e.g. "Q254" for Wolfgang Amadeus Mozart
language Defaults to language set with `tw_set_language()`; if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all
id_df Default to NULL. If given, it should be a dataframe typically generated with `tw_get_()`, and is used instead of calling Wikidata or using SQLite cache. Ignored when `id` is of length more than one.
cache Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.
overwrite_cache Logical, defaults to FALSE. If TRUE, it overwrites the table in the local SQLite database. Useful if the original Wikidata object has been updated.
cache_connection Defaults to NULL. If NULL, and caching is enabled, `tidywikidatar` will use a local SQLite database. A custom connection to other databases can be given (see vignette caching for details).
disconnect_db Defaults to TRUE. If FALSE, leaves the connection to cache open.
wait In seconds, defaults to 0. Time to wait between queries to Wikidata. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.

Value

A character vector of the same length as the vector of `id` given, with the Wikidata label in the requested language.

Examples

tw_get_label(
  id = c(
    "Q180099",
    "Q228822"
  ),
  language = "en"
)

# If a label is not available, a NA value is returned
if (interactive()) {
  tw_get_label(
    id = c(
      "Q180099",
      "Q228822"
    ),
    language = "en"
  )
}
Get Wikidata property of an item as a character vector of the same length as input

**Description**

This function wraps `tw_get_p()`, but always sets `only_first` and `preferred` to `TRUE` in order to give back always a character vector.

**Usage**

```r
tw_get_p1(
  id,
  p,
  latest_start_time = FALSE,
  language = tidywikidatar::tw_get_language(),
  id_df = NULL,
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE,
  wait = 0
)
```

**Arguments**

- **id**  
  A character vector, must start with Q, e.g. "Q254" for Wolfgang Amadeus Mozart.

- **p**  
  A character vector, a property. Must always start with the capital letter "P", e.g. "P31" for "instance of".

- **latest_start_time**  
  Logical, defaults to `FALSE`. If `TRUE`, returns the property that has the most recent start time ("P580") as qualifier. If no such qualifier is found, then it is ignored.

- **language**  
  Defaults to language set with `tw_set_language()`, if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all.

- **id_df**  
  Default to `NULL`. If given, it should be a dataframe typically generated with `tw_get_()` and is used instead of calling Wikidata or replying on cache.
Describe the function `tw_get_property` and its parameters:

- **cache**
  - Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.

- **overwrite_cache**
  - Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.

- **cache_connection**
  - Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

- **disconnect_db**
  - Defaults to TRUE. If FALSE, leaves the connection to cache open.

- **wait**
  - In seconds, defaults to 0. Time to wait between queries to Wikidata. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.

**Value**

A character vector of the same length as the input.

**Examples**

```
tw_get_p1(id = "Q180099", "P26")
```

**Description**

Get Wikidata property of one or more items as a tidy data frame.

**Usage**

```
tw_get_property(
  id,
  p,
  language = tidywikidatar::tw_get_language(),
  id_df = NULL,
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE,
  wait = 0
)
```
Arguments

id    A character vector, must start with Q, e.g. "Q254" for Wolfgang Amadeus Mozart.

p     A character vector, a property. Must always start with the capital letter "P", e.g. "P31" for "instance of".

language  Defaults to language set with `tw_set_language()`; if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all

id_df  Default to NULL. If given, it should be a dataframe typically generated with `tw_get_()`, and is used instead of calling Wikidata or using SQLite cache. Ignored when `id` is of length more than one.

cache Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.

overwrite_cache Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.

cache_connection  Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

disconnect_db  Defaults to TRUE. If FALSE, leaves the connection to cache open.

wait  In seconds, defaults to 0. Time to wait between queries to Wikidata. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.

Value

A tibble, corresponding to the value for the given property. A tibble of zero rows if no relevant property found.

Examples

# Who were the doctoral advisors - P184 - of Margaret Mead - Q180099?
advisors <- tw_get_property(id = "Q180099", p = "P184")
advisors

# tw_get_label(advisors)

# It is also possible to get one property for many id

if (interactive()) {
  tw_get_property(
    id = c("Q180099", "Q228822"),
    p = "P31"
  )
}
# Or many properties for a single id

tw_get_property(
  id = "Q180099",
  p = c("P21", "P31")
)

## tw_get_property_description

*Get description of a Wikidata property in a given language*

### Description

Get description of a Wikidata property in a given language

### Usage

```r
tw_get_property_description(
  property,
  language = tidywikidatar::tw_get_language(),
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE,
  wait = 0
)
```

### Arguments

- **property**
  - A character vector of length 1, must start with P, e.g. "P31".

- **language**
  - Defaults to language set with `tw_set_language()`; if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all

- **cache**
  - Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.

- **overwrite_cache**
  - Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.

- **cache_connection**
  - Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

- **disconnect_db**
  - Defaults to TRUE. If FALSE, leaves the connection to cache open.

- **wait**
  - In seconds, defaults to 0. Time to wait between queries to Wikidata. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.
tw_get_property_label

Value

A character vector of length 1, with the Wikidata label in the requested language.

Examples

tw_get_property_description(property = "P31")

Get label of a Wikidata property in a given language

Usage

tw_get_property_label(
  property,
  language = tidywikidatar::tw_get_language(),
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE,
  wait = 0
)

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>property</td>
<td>A character vector. Each element must start with P, e.g. &quot;P31&quot;.</td>
</tr>
<tr>
<td>language</td>
<td>Defaults to language set with tw_set_language(); if not set, &quot;en&quot;. Use &quot;all_available&quot; to keep all languages. For available language values, see <a href="https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all">https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all</a>.</td>
</tr>
<tr>
<td>cache</td>
<td>Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with tw_enable_cache() or tw_disable_cache().</td>
</tr>
<tr>
<td>overwrite_cache</td>
<td>Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.</td>
</tr>
<tr>
<td>cache_connection</td>
<td>Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).</td>
</tr>
<tr>
<td>disconnect_db</td>
<td>Defaults to TRUE. If FALSE, leaves the connection to cache open.</td>
</tr>
<tr>
<td>wait</td>
<td>In seconds, defaults to 0. Time to wait between queries to Wikidata. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.</td>
</tr>
</tbody>
</table>
tw_get_property_label_single

Value
A character vector, with the Wikidata label in the requested language.

Examples

tw_get_property_label(property = "P31")

Description
Get label of a Wikidata property in a given language

Usage

tw_get_property_label_single(
  property,
  language = tidywikidatar::tw_get_language(),
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE,
  wait = 0
)

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>property</td>
<td>A character vector. Each element must start with P, e.g. &quot;P31&quot;.</td>
</tr>
<tr>
<td>language</td>
<td>Defaults to language set with tw_set_language(); if not set, &quot;en&quot;. Use &quot;all_available&quot; to keep all languages. For available language values, see <a href="https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all">https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all</a>.</td>
</tr>
<tr>
<td>cache</td>
<td>Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with tw_enable_cache() or tw_disable_cache().</td>
</tr>
<tr>
<td>overwrite_cache</td>
<td>Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.</td>
</tr>
<tr>
<td>cache_connection</td>
<td>Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).</td>
</tr>
<tr>
<td>disconnect_db</td>
<td>Defaults to TRUE. If FALSE, leaves the connection to cache open.</td>
</tr>
<tr>
<td>wait</td>
<td>In seconds, defaults to 0. Time to wait between queries to Wikidata. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.</td>
</tr>
</tbody>
</table>
tw_get_property_same_length

Value

A character vector of length 1, with the Wikidata label in the requested language.

Examples

```r
 tidywikidatar::tw_get_property_label_single(property = "P31")
```

---

tw_get_property_same_length

Get Wikidata property of an item as a vector or list of the same length as input

Description

Get Wikidata property of an item as a vector or list of the same length as input

Usage

```r
 tw_get_property_same_length(
   id,
   p,
   only_first = FALSE,
   preferred = FALSE,
   latest_start_time = FALSE,
   language = tidywikidatar::tw_get_language(),
   id_df = NULL,
   cache = NULL,
   overwrite_cache = FALSE,
   cache_connection = NULL,
   disconnect_db = TRUE,
   wait = 0
 )
```

```r
 tw_get_p(
   id,
   p,
   only_first = FALSE,
   preferred = FALSE,
   latest_start_time = FALSE,
   language = tidywikidatar::tw_get_language(),
   id_df = NULL,
   cache = NULL,
   overwrite_cache = FALSE,
   cache_connection = NULL,
   disconnect_db = TRUE,
   wait = 0
 )
```
Arguments

- **id**: A character vector, must start with Q, e.g. "Q254" for Wolfgang Amadeus Mozart.
- **p**: A character vector, a property. Must always start with the capital letter "P", e.g. "P31" for "instance of".
- **only_first**: Logical, defaults to FALSE. If TRUE, it just keeps the first relevant property value for each id (or NA if none is available), and returns a character vector. Warning: this likely discards valid values, so make sure this is really what you want. If FALSE, returns a list of the same length as input, with all values for each id stored in a list if more than one is found.
- **preferred**: Logical, defaults to FALSE. If TRUE, returns properties that have rank "preferred" if available; if no "preferred" property is found, then it is ignored.
- **latest_start_time**: Logical, defaults to FALSE. If TRUE, returns the property that has the most recent start time ("P580") as qualifier. If no such qualifier is found, then it is ignored.
- **language**: Defaults to language set with `tw_set_language()`; if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all
- **id_df**: Default to NULL. If given, it should be a dataframe typically generated with `tw_get_()` and is used instead of calling Wikidata or replying on cache.
- **cache**: Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.
- **overwrite_cache**: Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.
- **cache_connection**: Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).
- **disconnect_db**: Defaults to TRUE. If FALSE, leaves the connection to cache open.
- **wait**: In seconds, defaults to 0. Time to wait between queries to Wikidata. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.

Value

A list of the same length of input (or a character vector is only_first is set to TRUE)

Examples

# By default, it returns a list of the same length as input,
# no matter how many values for each id/property
if (interactive()) {
  tw_get_property_same_length(
    id = c(
      "Q180099",
      "Q228822",
      "Q76857"
    ),
    p = "P26"
  )
  # Notice that if no relevant match is found, it returns a NA
  # This is useful for piped operations

tibble::tibble(id = c(
  "Q180099",
  "Q228822",
  "Q76857"
)) %>%
  dplyr::mutate(spouse = tw_get_property_same_length(id, "P26"))

  # Consider unnesting for further analysis

tibble::tibble(id = c(
  "Q180099",
  "Q228822",
  "Q76857"
)) %>%
  dplyr::mutate(spouse = tw_get_property_same_length(id, "P26")) %>%
  tidyr::unnest(cols = spouse)

  # If you are sure that you are interested only in the first return value,
  # consider setting only_first=TRUE to get a character vector rather than a list
  # Be mindful: you may well be discarding valid values.

tibble::tibble(id = c(
  "Q180099",
  "Q228822",
  "Q76857"
)) %>%
  dplyr::mutate(spouse = tw_get_property_same_length(id, "P26",
                                                     only_first = TRUE
    ))
}
tw_get_p(id = "Q180099", "P26")

---

tw_get_property_with_details

*Gets all details of a property*

**Description**

Gets all details of a property
Usage

\texttt{tw_get_property_with_details(id, p, wait = 0)}

Arguments

\begin{itemize}
  \item \textbf{id} A character vector, must start with Q, e.g. "Q254" for Wolfgang Amadeus Mozart.
  \item \textbf{p} A character vector, a property. Must always start with the capital letter "P", e.g. "P31" for "instance of".
  \item \textbf{wait} In seconds, defaults to 0. Time to wait between queries to Wikidata. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.
\end{itemize}

Value

A tibble, corresponding to the details for the given property. NULL if no relevant property found.

Examples

\begin{verbatim}
# Get "female form of label", including language
tw_get_property_with_details(id = "Q64733534", p = "P2521")
\end{verbatim}

\texttt{tw_get_property_with_details_single

\textit{Gets all details of a property}}

Description

Gets all details of a property

Usage

\texttt{tw_get_property_with_details_single(id, p)}

Arguments

\begin{itemize}
  \item \textbf{id} A character vector, must start with Q, e.g. "Q254" for Wolfgang Amadeus Mozart.
  \item \textbf{p} A character vector, a property. Must always start with the capital letter "P", e.g. "P31" for "instance of".
\end{itemize}

Value

A tibble, corresponding to the details for the given property. NULL if no relevant property found.

Examples

\begin{verbatim}
# Get "female form of label", including language
tidywikidatar::tw_get_property_with_details_single(id = "Q64733534", p = "P2521")
\end{verbatim}
tw_get_p_wide

Description

Efficiently get a wide table with various properties of a given set of Wikidata identifiers

Usage

tw_get_p_wide(
  id,
  p,
  label = FALSE,
  property_label_as_column_name = FALSE,
  both_id_and_label = FALSE,
  only_first = FALSE,
  preferred = FALSE,
  unlist = FALSE,
  collapse = ";",
  language = tidywikidatar::tw_get_language(),
  id_df = NULL,
  id_df_label = NULL,
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE,
  wait = 0
)

Arguments

id       A character vector, must start with Q, e.g. "Q254" for Wolfgang Amadeus Mozart.
p       A character vector, a property. Must always start with the capital letter "P", e.g. "P31" for "instance of".
label       Logical, defaults to FALSE. If TRUE labels of Wikidata Q identifiers are reported instead of the identifiers themselves (or labels are presented along of them, if both_id_and_label is set to TRUE)
property_label_as_column_name       Logical, defaults to FALSE. If FALSE, names of columns with properties are the "P" identifiers of the property. If TRUE, the label of the correspondent property is assigned as column name.
both_id_and_label       Logical, defaults to FALSE. Relevant only if label is set to TRUE, otherwise ignored. If TRUE, the label is added as a separate column along the original one. Column name is the same as the property column, followed by "_label".
only_first Logical, defaults to FALSE. If TRUE, it just keeps the first relevant property value for each id (or NA if none is available), and returns a character vector. Warning: this likely discards valid values, so make sure this is really what you want. If FALSE, returns a list of the same length as input, with all values for each id stored in a list if more than one is found.

preferred Logical, defaults to FALSE. If TRUE, returns properties that have rank "preferred" if available; if no "preferred" property is found, then it is ignored.

unlist Logical, defaults to FALSE. Typically used sharing or exporting data as csv files. Collapses all properties in a single string. The separator is defined by the collapse parameter. Relevant only when only_first is set to FALSE.

collapse Defaults to ",". Character used to separate results when unlist is set to TRUE.

tw_set_language() if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all

id_df Default to NULL. If given, it should be a dataframe typically generated with tw_get_(), and is used instead of calling Wikidata or replying on cache.

id_df_label Defaults to NULL. If given, it should be a dataframe typically generated with tw_get() with all items for which labels will be requested. It is used instead of calling Wikidata or relying on cache.

cache Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with tw_enable_cache() or tw_disable_cache().

overwrite_cache Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.

cache_connection Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

disconnect_db Defaults to TRUE. If FALSE, leaves the connection to cache open.

wait In seconds, defaults to 0. Time to wait between queries to Wikidata. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.

Value

A data frame, with a column for each given property.

Examples

if (interactive()) {
  tw_get_p_wide(  
    id = c("Q180099", "Q228822", "Q191095"),  
    p = c("P27", "P19", "P20"),  
    label = TRUE,  
    only_first = TRUE
  )
}

tw_get_qualifiers  

Get Wikidata qualifiers for a given property of a given item

Description

N.B. In order to provide for consistently structured output, this function outputs either id or value for each qualifier. The user should keep in mind that some of these come with additional detail (e.g. the unit, precision, or reference calendar).

Usage

tw_get_qualifiers(
  id,
  p,
  language = tidywikidatar::tw_get_language(),
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE,
  wait = 0,
  id_l = NULL
)

Arguments

id  A character vector of length 1, must start with Q, e.g. "Q254" for Wolfgang Amadeus Mozart.
p  A character vector of length 1, a property. Must always start with the capital letter "P", e.g. "P31" for "instance of".
language  Defaults to language set with tw_set_language(); if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all

cache  Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with tw_enable_cache() or tw_disable_cache().
overwrite_cache  Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.
cache_connection  Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).
disconnect_db  Defaults to TRUE. If FALSE, leaves the connection to cache open.
wait  In seconds, defaults to 0. Time to wait between queries to Wikidata. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.
**tw_get_qualifiers_single**

Get Wikidata qualifiers for a given property of a given item

### Description

N.B. In order to provide for consistently structured output, this function outputs either id or value for each qualifier. The user should keep in mind that some of these come with additional detail (e.g. the unit, precision, or reference calendar).

### Usage

```r
tw_get_qualifiers_single(
  id,
  p,
  language = tidywikidatar::tw_get_language(),
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE,
  wait = 0,
  id_l = NULL
)
```

### Examples

```r
if (interactive()) {
  tidywikidatar::tw_get_qualifiers(id = "Q180099", p = "P26", language = "en")
}
```

# using `tw_test_items` in examples in order to show output without calling on Wikidata servers

```r
tidywikidatar::tw_get_qualifiers(
  id = "Q180099",
  p = "P26",
  language = "en",
  id_l = tw_test_items
)
```
Arguments

- **id**: A character vector of length 1, must start with Q, e.g. "Q254" for Wolfgang Amadeus Mozart.
- **p**: A character vector of length 1, a property. Must always start with the capital letter "P", e.g. "P31" for "instance of".
- **language**: Defaults to language set with `tw_set_language()`; if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all
- **cache**: Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.
- **overwrite_cache**: Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.
- **cache_connection**: Defaults to NULL. If NULL, and caching is enabled, `tidywikidatar` will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).
- **disconnect_db**: Defaults to TRUE. If FALSE, leaves the connection to cache open.
- **wait**: In seconds, defaults to 0. Time to wait between queries to Wikidata. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.
- **id_l**: Defaults to NULL. If given, must be an object or list such as the one generated with `tidywikidatar:::tw_get_item()`. If given, and the requested id is actually present in id_l, then no query to Wikidata servers is made.

Value

A data frame (a tibble) with eight columns: id for the input id, property, qualifier_id, qualifier_property, qualifier_value, rank, qualifier_value_type, and set (to distinguish sets of data when a property is present more than once)

Examples

```r
if (interactive()) {
  tidywikidatar:::tw_get_qualifiers_single(id = "Q180099", p = "P26", language = "en")
}

#' ## using 'tw_test_items' in examples in order to show output without calling
## on Wikidata servers

 tidywikidatar:::tw_get_qualifiers_single(
  id = "Q180099",
  p = "P26",
  language = "en",
  id_l = tw_test_items
)
```
### tw_get_single

*Return (most) information from a Wikidata item in a tidy format from a single Wikidata identifier*

#### Description

Return (most) information from a Wikidata item in a tidy format from a single Wikidata identifier

#### Usage

```r
tw_get_single(
  id,
  language = tidywikidatar::tw_get_language(),
  cache = NULL,
  overwrite_cache = FALSE,
  read_cache = TRUE,
  cache_connection = NULL,
  disconnect_db = TRUE,
  wait = 0,
  id_l = NULL
)
```

#### Arguments

- **id**: A character vector, must start with Q, e.g. "Q180099" for the anthropologist Margaret Mead. Can also be a data frame of one row, typically generated with `tw_search()` or a combination of `tw_search()` and `tw_filter_first()`.
- **language**: Defaults to language set with `tw_set_language()`; if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all
- **cache**: Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.
- **overwrite_cache**: Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.
- **read_cache**: Logical, defaults to TRUE. Mostly used internally to prevent checking if an item is in cache if it is already known that it is not in cache.
- **cache_connection**: Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).
- **disconnect_db**: Defaults to TRUE. If FALSE, leaves the connection to cache open.
- **wait**: In seconds, defaults to 0. Time to wait between queries to Wikidata. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.
id_l

Defaults to NULL. If given, must be an object or list such as the one generated with `WikiDataR::get_item()`. If given, and the requested id is actually present in `id_l`, then no query to Wikidata servers is made.

Value

A data.frame (a tibble) with four columns (id, property, value, and rank). If item not found or trouble connecting with the server, a data frame with four columns and zero rows is returned, with the warning as an attribute, which can be retrieved with `attr(output, "warning")`.

Examples

```r
if (interactive()) {
  tidywikidatar::tw_get_single(
    id = "Q180099",
    language = "en"
  )
}
```

#' ## using `tw_test_items` in examples in order to show output without calling
#' ## on Wikidata servers
```
```r
tidywikidatar::tw_get_single(
  id = "Q180099",
  language = "en",
  id_l = tw_test_items
)
```

---

### Get URL to a Wikipedia article corresponding to a Wikidata Q identifier in given language

**Description**

Get URL to a Wikipedia article corresponding to a Wikidata Q identifier in given language

**Usage**

```r
tw_get_wikipedia(
  id,
  full_link = TRUE,
  language = tidywikidatar::tw_get_language(),
  id_df = NULL,
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE,
  wait = 0
)
```
**Arguments**

- **id**
  A character vector, must start with Q, e.g., "Q254" for Wolfgang Amadeus Mozart

- **full_link**
  Logical, defaults to TRUE. If FALSE, returns only the part of the url that corresponds to the title.

- **language**
  Defaults to language set with `tw_set_language()`; if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all

- **id_df**
  Default to NULL. If given, it should be a dataframe typically generated with `tw_get_()`, and is used instead of calling Wikidata or using SQLite cache.

- **cache**
  Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.

- **overwrite_cache**
  Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.

- **cache_connection**
  Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

- **disconnect_db**
  Defaults to TRUE. If FALSE, leaves the connection to cache open.

- **wait**
  In seconds, defaults to 0. Time to wait between queries to Wikidata. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.

**Value**

A character vector of the same length as the vector of id given, with the Wikipedia link in the requested language.

**Examples**

```
tw_get_wikipedia(id = "Q180099")
```

---

**tw_get_wikipedia_base_api_url**

*Facilitates the creation of MediaWiki API base URLs*

**Description**

Mostly used internally
Usage

tw_get_wikipedia_base_api_url(
    url = NULL,
    title = NULL,
    language = tidywikidatar::tw_get_language(),
    action = "query",
    type = "page"
)

Arguments

url A character vector with the full URL to one or more Wikipedia pages. If given, title and language can be left empty.
title Title of a Wikipedia page or final parts of its url. If given, url can be left empty, but language must be provided.
language Two-letter language code used to define the Wikipedia version to use. Defaults to language set with tw_set_language(); if not set, "en". If url given, this can be left empty.
action Defaults to "query". Usually either "query" or "parse". In principle, any valid action value, see: https://www.mediawiki.org/w/api.php
type Defaults to "page". Either "page" or "category".

Value

A character vector of base urls to be used with the MediaWiki API

Examples

tw_get_wikipedia_base_api_url(title = "Margaret Mead", language = "en")
tw_get_wikipedia_base_api_url(
    title = "Category:American women anthropologists",
    type = "category",
    language = "en"
)

---

tw_get_wikipedia_category_members

Get all Wikidata Q identifiers of all Wikipedia pages (or files, or subcategories) that are members of the given category,

Description

Get all Wikidata Q identifiers of all Wikipedia pages (or files, or subcategories) that are members of the given category,
Usage

tw_get_wikipedia_category_members(
  url = NULL,
  category = NULL,
  type = "page",
  language = tidywikidatar::tw_get_language(),
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE,
  wait = 1,
  attempts = 10
)

Arguments

url Full URL to a Wikipedia category page. If given, title and language can be left empty.

category Title of a Wikipedia category page or final parts of its url. Must include "Category:", or equivalent in other languages. If given, url can be left empty, but language must be provided.

type Defaults to "page", defines which kind of members of a category are returned. Valid values include "page", "file", and "subcat" (for sub-category). Corresponds to cmtype. For details, see https://www.mediawiki.org/wiki/API:Categorymembers

language Two-letter language code used to define the Wikipedia version to use. Defaults to language set with tw_set_language(); if not set, "en". If url given, this can be left empty.

cache Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with tw_enable_cache() or tw_disable_cache().

overwrite_cache Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.

cache_connection Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

disconnect_db Defaults to TRUE. If FALSE, leaves the connection to cache open.

wait In seconds, defaults to 1 due to time-outs with frequent queries. Time to wait between queries to the APIs. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.

attempts Defaults to 10. Number of times it re-attempts to reach the API before failing.
Value

A data frame (a tibble) with eight columns: `source_title_url`, `source_wikipedia_title`, `source_qid`, `wikipedia_title`, `wikipedia_id`, `qid`, `description`, and `language`.

Examples

```r
if (interactive()) {
  sub_categories <- tw_get_wikipedia_category_members(
    category = "Category:American women anthropologists",
    type = "subcat"
  )

  sub_categories

  tw_get_wikipedia_category_members(
    category = sub_categories$wikipedia_title,
    type = "page"
  )
}
```

---

tw_get_wikipedia_category_members_single

Get all Wikidata Q identifiers of all Wikipedia pages that appear in a given page

Description

Get all Wikidata Q identifiers of all Wikipedia pages that appear in a given page

Usage

```r
tw_get_wikipedia_category_members_single(
  url = NULL,
  category = NULL,
  type = "page",
  language = tidywikidatar::tw_get_language(),
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE,
  wait = 1,
  attempts = 10
)
```
Arguments

url Full URL to a Wikipedia category page. If given, title and language can be left empty.

category Title of a Wikipedia category page or final parts of its url. Must include "Category:", or equivalent in other languages. If given, url can be left empty, but language must be provided.

type Defaults to "page", defines which kind of members of a category are returned. Valid values include "page", "file", and "subcat" (for sub-category). Corresponds to cmtype. For details, see https://www.mediawiki.org/wiki/API:Categorymembers

language Two-letter language code used to define the Wikipedia version to use. Defaults to language set with tw_set_language(); if not set, "en". If url given, this can be left empty.

cache Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with tw_enable_cache() or tw_disable_cache().

overwrite_cache Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.

cache_connection Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

disconnect_db Defaults to TRUE. If FALSE, leaves the connection to cache open.

wait In seconds, defaults to 1 due to time-outs with frequent queries. Time to wait between queries to the APIs. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.

attempts Defaults to 10. Number of times it re-attempts to reach the API before failing.

Value

A data frame (a tibble) with four columns: wikipedia_title, wikipedia_id, wikidata_id, wikidata_description.

Examples

if (interactive()) {
  tidywikidatar::tw_get_wikipedia_category_members_single(
    category = "Category:American women anthropologists",
    type = "subcat"
  )

  tidywikidatar::tw_get_wikipedia_category_members_single(
    category = "Category:Puerto Rican women anthropologists",
    type = "page"
  )
}

tw_get_wikipedia_page_links

Get all Wikidata Q identifiers of all Wikipedia pages that appear in one or more pages

Description

Get all Wikidata Q identifiers of all Wikipedia pages that appear in one or more pages

Usage

```r
tw_get_wikipedia_page_links(
  url = NULL,
  title = NULL,
  language = tidywikidatar::tw_get_language(),
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE,
  wait = 1,
  attempts = 10
)
```

Arguments

- **url**
  Full URL to a Wikipedia page. If given, title and language can be left empty.

- **title**
  Title of a Wikipedia page or final parts of its URL. If given, url can be left empty, but language must be provided.

- **language**
  Two-letter language code used to define the Wikipedia version to use. Defaults to language set with `tw_set_language()`; if not set, "en". If url given, this can be left empty.

- **cache**
  Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.

- **overwrite_cache**
  Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.

- **cache_connection**
  Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

- **disconnect_db**
  Defaults to TRUE. If FALSE, leaves the connection to cache open.

- **wait**
  In seconds, defaults to 1 due to time-outs with frequent queries. Time to wait between queries to the APIs. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.

- **attempts**
  Defaults to 10. Number of times it re-attempts to reach the API before failing.
Value

A data frame (a tibble) with eight columns: `source_title_url`, `source_wikipedia_title`, `source_qid`, `wikipedia_title`, `wikipedia_id`, `qid`, `description`, and `language`.

Examples

```r
if (interactive()) {
  tw_get_wikipedia_page_links(title = "Margaret Mead", language = "en")
}
```

tw_get_wikipedia_page_links_single

Get all Wikidata Q identifiers of all Wikipedia pages that appear in a given page

Description

Get all Wikidata Q identifiers of all Wikipedia pages that appear in a given page

Usage

```r
tw_get_wikipedia_page_links_single(
  url = NULL,
  title = NULL,
  language = tidywikidatar::tw_get_language(),
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE,
  wait = 1,
  attempts = 10,
  wikipedia_page_qid_df = NULL
)
```

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>url</code></td>
<td>Full URL to a Wikipedia page. If given, title and language can be left empty.</td>
</tr>
<tr>
<td><code>title</code></td>
<td>Title of a Wikipedia page or final parts of its url. If given, url can be left empty, but language must be provided.</td>
</tr>
<tr>
<td><code>language</code></td>
<td>Two-letter language code used to define the Wikipedia version to use. Defaults to language set with <code>tw_set_language()</code>; if not set, &quot;en&quot;. If url given, this can be left empty.</td>
</tr>
<tr>
<td><code>cache</code></td>
<td>Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with <code>tw_enable_cache()</code> or <code>tw_disable_cache()</code>.</td>
</tr>
<tr>
<td><code>overwrite_cache</code></td>
<td>Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.</td>
</tr>
</tbody>
</table>
cache_connection  Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a
local sqlite database. A custom connection to other databases can be given (see
vignette caching for details).

disconnect_db  Defaults to TRUE. If FALSE, leaves the connection to cache open.

wait  In seconds, defaults to 1 due to time-outs with frequent queries. Time to wait be-
tween queries to the APIs. If data are cached locally, wait time is not applied. If
you are running many queries systematically you may want to add some waiting
time between queries.

attempts  Defaults to 10. Number of times it re-attempts to reach the API before failing.

wikipedia_page_qid_df  Defaults to NULL. If given, used to reduce calls to cache. A data frame

Value

A data frame (a tibble) with four columns: wikipedia_title, wikipedia_id, wikidata_id,
wikidata_description.

Examples

if (interactive()) {
  tw_get_wikipedia_page_links_single(title = "Margaret Mead", language = "en")
}

---

tw_get_wikipedia_page_qid

*Gets the Wikidata Q identifier of one or more Wikipedia pages*

Description

Gets the Wikidata Q identifier of one or more Wikipedia pages

Usage

tw_get_wikipedia_page_qid(
  url = NULL,
  title = NULL,
  language = tidywikidatar::tw_get_language(),
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE,
  wait = 1,
  attempts = 10
)
Arguments

url A character vector with the full URL to one or more Wikipedia pages. If given, title and language can be left empty.

title Title of a Wikipedia page or final parts of its url. If given, url can be left empty, but language must be provided.

language Two-letter language code used to define the Wikipedia version to use. Defaults to language set with `tw_set_language()`; if not set, "en". If url given, this can be left empty.

cache Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.

overwrite_cache Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.

cache_connection Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

disconnect_db Defaults to TRUE. If FALSE, leaves the connection to cache open.

wait In seconds, defaults to 1 due to time-outs with frequent queries. Time to wait between queries to the APIs. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.

attempts Defaults to 10. Number of times it re-attempts to reach the API before failing.

Value

A a data frame with six columns, including qid with Wikidata identifiers, and a logical disambiguation to flag when disambiguation pages are returned.

Examples

```r
if (interactive()) {
  tw_get_wikipedia_page_qid(title = "Margaret Mead", language = "en")

  # check when Wikipedia returns disambiguation page
tw_get_wikipedia_page_qid(title = c("Rome", "London", "New York", "Vienna"))
}
```

---

tw_get_wikipedia_page_qid_single

*Gets the Wikidata id of a Wikipedia page*

---

Description

Gets the Wikidata id of a Wikipedia page
tw_get_wikipedia_page_qid_single

Usage

tw_get_wikipedia_page_qid_single(
    title = NULL,
    url = NULL,
    language = tidywikidatar::tw_get_language(),
    cache = NULL,
    overwrite_cache = FALSE,
    cache_connection = NULL,
    disconnect_db = TRUE,
    wait = 1,
    attempts = 10
)

Arguments

title Title of a Wikipedia page or final parts of its url. If given, url can be left empty, but language must be provided.
url Full URL to a Wikipedia page. If given, title and language can be left empty.
language Two-letter language code used to define the Wikipedia version to use. Defaults to language set with tw_set_language(); if not set, "en". If url given, this can be left empty.
cache Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with tw_enable_cache() or tw_disable_cache().
overwrite_cache Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.
cache_connection Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).
disconnect_db Defaults to TRUE. If FALSE, leaves the connection to cache open.
wait In seconds, defaults to 1 due to time-outs with frequent queries. Time to wait between queries to the APIs. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.
attempts Defaults to 10. Number of times it re-attempts to reach the API before failing.

Value

A data frame (a tibble) with eight columns: title, wikipedia_title, wikipedia_id, qid, description, disambiguation, and language.

Examples

if (interactive()) {
  tw_get_wikipedia_page_qid_single(title = "Margaret Mead", language = "en")
}
tw_get_wikipedia_page_sections

Get sections of a Wikipedia page

Description

Get sections of a Wikipedia page

Usage

tw_get_wikipedia_page_sections(
  url = NULL,
  title = NULL,
  language = tidywikidatar::tw_get_language(),
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE,
  wait = 1,
  attempts = 10
)

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>url</td>
<td>Full URL to a Wikipedia page. If given, title and language can be left empty.</td>
</tr>
<tr>
<td>title</td>
<td>Title of a Wikipedia page or final parts of its url. If given, url can be left empty, but language must be provided.</td>
</tr>
<tr>
<td>language</td>
<td>Two-letter language code used to define the Wikipedia version to use. Defaults to language set with tw_set_language(); if not set, &quot;en&quot;. If url given, this can be left empty.</td>
</tr>
<tr>
<td>cache</td>
<td>Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with tw_enable_cache() or tw_disable_cache().</td>
</tr>
<tr>
<td>overwrite_cache</td>
<td>Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.</td>
</tr>
<tr>
<td>cache_connection</td>
<td>Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).</td>
</tr>
<tr>
<td>disconnect_db</td>
<td>Defaults to TRUE. If FALSE, leaves the connection to cache open.</td>
</tr>
<tr>
<td>wait</td>
<td>In seconds, defaults to 1 due to time-outs with frequent queries. Time to wait between queries to the APIs. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.</td>
</tr>
<tr>
<td>attempts</td>
<td>Defaults to 10. Number of times it re-attempts to reach the API before failing.</td>
</tr>
</tbody>
</table>
Value

A data frame (a tibble), with the same columns as `tw_empty_wikipedia_page_sections`.

Examples

```r
if (interactive()) {
  tw_get_wikipedia_page_sections(title = "Margaret Mead", language = "en")
}
```

---

tw_get_wikipedia_page_sections_single

Get all Wikidata Q identifiers of all Wikipedia pages that appear in a given page

Description

Get all Wikidata Q identifiers of all Wikipedia pages that appear in a given page

Usage

```r
tw_get_wikipedia_page_sections_single(
  url = NULL,
  title = NULL,
  language = tidywikidatar::tw_get_language(),
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE,
  wait = 1,
  attempts = 10,
  wikipedia_page_qid_df = NULL
)
```

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>url</code></td>
<td>Full URL to a Wikipedia page. If given, title and language can be left empty.</td>
</tr>
<tr>
<td><code>title</code></td>
<td>Title of a Wikipedia page or final parts of its url. If given, url can be left empty, but language must be provided.</td>
</tr>
<tr>
<td><code>language</code></td>
<td>Two-letter language code used to define the Wikipedia version to use. Defaults to language set with <code>tw_set_language()</code>; if not set, &quot;en&quot;. If url given, this can be left empty.</td>
</tr>
<tr>
<td><code>cache</code></td>
<td>Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with <code>tw_enable_cache()</code> or <code>tw_disable_cache()</code>.</td>
</tr>
<tr>
<td><code>overwrite_cache</code></td>
<td>Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.</td>
</tr>
</tbody>
</table>
tw_get_wikipedia_page_section_links

Cache connection
 Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

disconnect_db
 Defaults to TRUE. If FALSE, leaves the connection to cache open.

wait
 In seconds, defaults to 1 due to time-outs with frequent queries. Time to wait between queries to the APIs. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.

attempts
 Defaults to 10. Number of times it re-attempts to reach the API before failing.

wikipedia_page_qid_df
 Defaults to NULL. If given, used to reduce calls to cache. A data frame

Value

A data frame (a tibble) with four columns: wikipedia_title, wikipedia_id, wikidata_id, wikidata_description.

Examples

if (interactive()) {
  tw_get_wikipedia_page_sections_single(title = "Margaret Mead", language = "en")
}

tw_get_wikipedia_page_section_links

Get links from a specific section of a Wikipedia page

Description

Get links from a specific section of a Wikipedia page

Usage

tw_get_wikipedia_page_section_links(
  url = NULL,
  title = NULL,
  section_title = NULL,
  section_index = NULL,
  language = tidywikidatar::tw_get_language(),
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE,
  wait = 1,
  attempts = 10,
  wikipedia_page_qid_df = NULL
)
### Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>url</code></td>
<td>Full URL to a Wikipedia page. If given, title and language can be left empty.</td>
</tr>
<tr>
<td><code>title</code></td>
<td>Title of a Wikipedia page or final parts of its url. If given, url can be left empty, but language must be provided.</td>
</tr>
<tr>
<td><code>section_title</code></td>
<td>Defaults to NULL. If given, it should correspond to the human-readable title of a section of the relevant Wikipedia page. See also <code>tw_get_wikipedia_page_sections()</code></td>
</tr>
<tr>
<td><code>section_index</code></td>
<td>Defaults to NULL. If given, it should correspond to the ordinal of a section of the relevant Wikipedia page. See also <code>tw_get_wikipedia_page_sections()</code></td>
</tr>
<tr>
<td><code>language</code></td>
<td>Two-letter language code used to define the Wikipedia version to use. Defaults to language set with <code>tw_set_language()</code>; if not set, &quot;en&quot;. If url given, this can be left empty.</td>
</tr>
<tr>
<td><code>cache</code></td>
<td>Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with <code>tw_enable_cache()</code> or <code>tw_disable_cache()</code>.</td>
</tr>
<tr>
<td><code>overwrite_cache</code></td>
<td>Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.</td>
</tr>
<tr>
<td><code>cache_connection</code></td>
<td>Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).</td>
</tr>
<tr>
<td><code>disconnect_db</code></td>
<td>Defaults to TRUE. If FALSE, leaves the connection to cache open.</td>
</tr>
<tr>
<td><code>wait</code></td>
<td>In seconds, defaults to 1 due to time-outs with frequent queries. Time to wait between queries to the APIs. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.</td>
</tr>
<tr>
<td><code>attempts</code></td>
<td>Defaults to 10. Number of times it re-attempts to reach the API before failing.</td>
</tr>
<tr>
<td><code>wikipedia_page_qid_df</code></td>
<td>Defaults to NULL. If given, used to reduce calls to cache. A data frame</td>
</tr>
</tbody>
</table>

### Value

A data frame (a tibble).

### Examples

```r
if (interactive()) {
  tw_get_wikipedia_page_section_links(title = "Margaret Mead", language = "en", section_index = 1)
}
```
tw_get_wikipedia_sections_api_url

Facilitates the creation of MediaWiki API base URLs to retrieve sections of a page

Description

Mostly used internally

Usage

tw_get_wikipedia_sections_api_url(
  url = NULL,
  title = NULL,
  language = tidywikidatar::tw_get_language()
)

Arguments

url A character vector with the full URL to one or more Wikipedia pages. If given, title and language can be left empty.
title Title of a Wikipedia page or final parts of its url. If given, url can be left empty, but language must be provided.
language Two-letter language code used to define the Wikipedia version to use. Defaults to language set with tw_set_language(): if not set, "en". If url given, this can be left empty.

Value

A character vector of base urls to be used with the MediaWiki API

Examples

tw_get_wikipedia_sections_api_url(title = "Margaret Mead", language = "en")
Usage

tw_get_wikipedia_section_links_api_url(
  url = NULL,
  title = NULL,
  section_index,
  language = tidywikidatar::tw_get_language()
)

Arguments

url A character vector with the full URL to one or more Wikipedia pages. If given, title and language can be left empty.
title Title of a Wikipedia page or final parts of its url. If given, url can be left empty, but language must be provided.
section_index Required. It should correspond to the ordinal of a section of the relevant Wikipedia page. See also tw_get_wikipedia_page_sections().
language Two-letter language code used to define the Wikipedia version to use. Defaults to language set with tw_set_language(); if not set, "en". If url given, this can be left empty.

Value

A character vector of base urls to be used with the MediaWiki API

Examples

tw_get_wikipedia_section_links_api_url(title = "Margaret Mead", section_index = 1, language = "en")

---

tw_index_cache_item Add index to caching table for search queries for increased speed

Description

Tested only with SQLite and MySql. May work with other drivers.

Usage

tw_index_cache_item(
  table_name = NULL,
  check_first = TRUE,
  type = "item",
  show_details = FALSE,
  language = tidywikidatar::tw_get_language(),
  cache = NULL,
  cache_connection = NULL,
  disconnect_db = TRUE
)
Arguments

**table_name**  
Name of the table in the database. If given, it takes precedence over other parameters.

**check_first**  
Logical, defaults to TRUE. If TRUE, then before executing anything on the database it checks if the given table has already been indexed. If it has, it does nothing and returns only an informative message.

**type**  
Defaults to "item". Type of cache file to output. Values typically used by tidywikidatar include "item", "search_item", "search_property", and "qualifier".

**show_details**  
Logical, defaults to FALSE. If FALSE, return the function adds the index to the database, but does not return anything. If TRUE, returns a data frame with more details about the index.

**language**  
Defaults to language set with `tw_set_language()`; "en" if not set. Used to limit the data to be cached. Use "all_available" to keep all data. For available values, see [https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all](https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all)

**cache**  
Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.

**cache_connection**  
Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

**disconnect_db**  
Defaults to TRUE. If FALSE, leaves the connection to cache open.

Details

To ensure smooth functioning, the search column in the cache table is transformed into a column of type varchar and length 255.

Value

If `show_details` is set to FALSE, nothing, used only for its side effects (add index to caching table). If TRUE, a data frame, same as the output of `tw_check_cache_index(show_details = TRUE)`.

Examples

```r
if (interactive()) {
  tw_enable_cache()
  tw_set_cache_folder(path = fs::path(fs::path_home_r(), "R", "tw_data"))

  tw_index_cache_search()
}
```
**Description**

Tested only with SQLite and MySql. May work with other drivers.

**Usage**

```r
tw_index_cache_search(
  table_name = NULL,
  check_first = TRUE,
  type = "item",
  show_details = FALSE,
  language = tidywikidatar::tw_get_language(),
  response_language = tidywikidatar::tw_get_language(),
  cache = NULL,
  cache_connection = NULL,
  disconnect_db = TRUE
)
```

**Arguments**

- `table_name`: Name of the table in the database. If given, it takes precedence over other parameters.
- `check_first`: Logical, defaults to TRUE. If TRUE, then before executing anything on the database it checks if the given table has already been indexed. If it has, it does nothing and returns only an informative message.
- `type`: Defaults to "item". Type of cache file to output. Values typically used by tidywikidatar include "item", "search_item", "search_property", and "qualifier".
- `show_details`: Logical, defaults to FALSE. If FALSE, return the function adds the index to the database, but does not return anything. If TRUE, returns a data frame with more details about the index.
- `language`: Defaults to language set with `tw_set_language()`; "en" if not set. Used to limit the data to be cached. Use "all_available" to keep all data. For available values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all
- `response_language`: Defaults to language set with `tw_set_language()`; "en" if not set. Relevant only when type is set to "search_item" or "search_property". See `tw_search()` for details.
- `cache`: Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.

cache_connection

Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

disconnect_db

Defaults to TRUE. If FALSE, leaves the connection to cache open.

Details
To ensure smooth functioning, the search column in the cache table is transformed into a column of type varchar and length 255.

Value
If show_details is set to FALSE, nothing, used only for its side effects (add index to caching table). If TRUE, a data frame, same as the output of tw_check_cache_index(show_details = TRUE).

Examples
if (interactive()) {
tw_enable_cache()
tw_set_cache_folder(path = fs::path(
    fs::path_home_r(),
    "R",
    "tw_data"
))

tw_index_cache_search()
}

---

tw_label

*Gets labels for all columns with names such as "id" and "property".*

**Description**

Gets labels for all columns with names such as "id" and "property".

**Usage**

tw_label(
    df, 
    value = TRUE, 
    language = tidywikidatar::tw_get_language(), 
    cache = NULL, 
    overwrite_cache = FALSE, 
    cache_connection = NULL, 
    disconnect_db = TRUE, 
    wait = 0
)
Arguments

df
A data frame, typically generated with other tidywikidatar functions such as `tw_get_property()`

value
Logical, defaults to TRUE. If TRUE, it tries to get labels for all supposed id in the column called value. May break if the columns include some value which starts with Q and some digits, but is not a wikidata id.

language
Defaults to language set with `tw_set_language()`; if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all

cache
Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.

overwrite_cache
Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.

cache_connection
Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

disconnect_db
Defaults to TRUE. If FALSE, leaves the connection to cache open.

wait
In seconds, defaults to 0. Time to wait between queries to Wikidata. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.

Value

A data frame, with the same shape as the input data frame, but with labels instead of identifiers.

Examples

```r
if (interactive()) {
  tw_get_qualifiers(id = "Q180099", p = "P26", language = "en") %>%
  head(2) %>%
  tw_label()
}
```

---

tw_qid_airports

The Wikidata Q identifier of European airports found in Eurostat’s avia_par_ dataset

description

The Wikidata Q identifier of European airports found in Eurostat’s avia_par_ dataset
Usage
tw_qid_airports

Format
A data frame with 429 rows and 1 column:

id Q identifiers

Source
https://www.wikidata.org/wiki/Wikidata:Main_Page

---

tw_qid_meps  The Wikidata Q identifier of all members of the European Parliament since its establishment

Description
A dataset with all the Wikidata items that have "Q27169" (member of the European Parliament) for the property "P39" (position held).

Usage
tw_qid_meps

Format
A data frame with 4581 rows and 1 column:

id Q identifiers

Source
https://www.wikidata.org/wiki/Wikidata:Main_Page
tw_query

Perform simple Wikidata queries

Description

This function aims to facilitate only the most basic type of queries: return which items have the following property pairs. For more details on Wikidata queries, consult: https://www.wikidata.org/wiki/Wikidata:SPARQL_query_service/examples. For complex queries, use WikidataQueryServiceR::query_wikidata().

Usage

tw_query(
  query,
  fields = c("item", "itemLabel", "itemDescription"),
  language = tidywikidatar::tw_get_language(),
  return_as_tw_search = TRUE
)

Arguments

query A list of named vectors, or a data frame (see example and readme).
fields A character vector of Wikidata fields. Ignored if return_as_tw_search is set to TRUE (as per default). Defaults to ("item", "itemLabel", "itemDescription")
language Defaults to language set with tidywikidatar::tw_set_language(); if not set, "en". If more than one, can be set in order of preference, e.g. c("it", "fr", "en"). Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all
return_as_tw_search Logical, defaults to TRUE. If TRUE, returns a data frame with three columns (id, label, and description) that can be piped to other tw_ functions. If FALSE, a data frame with as many columns as fields.

Details

Consider tw_get_all_with_p() if you want to get all items with a given property, irrespective of the value.

Value

A data frame

Examples

if (interactive()) {
  query <- list(
    c(p = "P106", q = "Q1397808"),
    c(p = "P21", q = "Q6581072")
  )
}
tw_reset_item_cache

Reset qualifiers cache

Description

Removes the table where qualifiers are cached

Usage

tw_reset_item_cache(
  language = tidywikidatar::tw_get_language(),
  cache = NULL,
  cache_connection = NULL,
  disconnect_db = TRUE,
  ask = TRUE
)

Arguments

language Defaults to language set with tw_set_language(); if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all

cache Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with tw_enable_cache() or tw_disable_cache().

cache_connection Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

disconnect_db Defaults to TRUE. If FALSE, leaves the connection to cache open.

ask Logical, defaults to TRUE. If FALSE, and cache folder does not exist, it just creates it without asking (useful for non-interactive sessions).

Value

Nothing, used for its side effects.

Examples

if (interactive()) {
  tw_reset_item_cache()
}

**tw_reset_qualifiers_cache**

Reset qualifiers cache

**Description**

Removes the table where qualifiers are cached

**Usage**

```r
tw_reset_qualifiers_cache(
  language = tidywikidatar::tw_get_language(),
  cache = NULL,
  cache_connection = NULL,
  disconnect_db = TRUE,
  ask = TRUE
)
```

**Arguments**

- **language**
  - Defaults to language set with `tw_set_language()`; if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all

- **cache**
  - Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.

- **cache_connection**
  - Defaults to NULL. If NULL, and caching is enabled, `tidywikidatar` will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

- **disconnect_db**
  - Defaults to TRUE. If FALSE, leaves the connection to cache open.

- **ask**
  - Logical, defaults to TRUE. If FALSE, and cache folder does not exist, it just creates it without asking (useful for non-interactive sessions).

**Value**

Nothing, used for its side effects.

**Examples**

```r
if (interactive()) {
  tw_reset_qualifiers_cache()
}
```
tw_reset_wikipedia_category_members_cache

Reset Wikipedia category members cache

Description

Removes from cache the table where data typically gathered with `tw_get_wikipedia_category_members()` are stored.

Usage

```r
tw_reset_wikipedia_category_members_cache(
  language = tidywikidatar::tw_get_language(),
  type = "page",
  cache = NULL,
  cache_connection = NULL,
  disconnect_db = TRUE,
  ask = TRUE
)
```

Arguments

- **language**: Defaults to language set with `tw_set_language()`; if not set, "en". Use "all_available" to keep all languages. For available language values, see [https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all](https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all)
- **type**: Defaults to "page", defines which kind of members of a category are returned. Valid values include "page", "file", and "subcat" (for sub-category). Corresponds to `cmtype`. For details, see [https://www.mediawiki.org/wiki/API:Categorymembers](https://www.mediawiki.org/wiki/API:Categorymembers)
- **cache**: Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.
- **cache_connection**: Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database by default. A custom connection to other databases can be given (see vignette `caching` for details).
- **disconnect_db**: Defaults to TRUE. If FALSE, leaves the connection to cache open.
- **ask**: Logical, defaults to TRUE. If FALSE, and cache folder does not exist, it just creates it without asking (useful for non-interactive sessions).

Value

Nothing, used for its side effects.

Examples

```r
if (interactive()) {
  tw_reset_wikipedia_category_members_cache()
}
```
Description

Removes the table where data typically gathered with `tw_get_wikipedia_page_qid()` from cache.

Usage

```r
tw_reset_wikipedia_page_cache(
    language = tidywikidatar::tw_get_language(),
    cache = NULL,
    cache_connection = NULL,
    disconnect_db = TRUE,
    ask = TRUE
)
```

Arguments

- **language** Defaults to language set with `tw_set_language()`; if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all.
- **cache** Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.
- **cache_connection** Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).
- **disconnect_db** Defaults to TRUE. If FALSE, leaves the connection to cache open.
- **ask** Logical, defaults to TRUE. If FALSE, and cache folder does not exist, it just creates it without asking (useful for non-interactive sessions).

Value

Nothing, used for its side effects.

Examples

```r
if (interactive()) {
    tw_reset_wikipedia_page_cache()
}
```
**tw_reset_wikipedia_page_links_cache**

*Reset Wikipedia page link cache*

**Description**

Removes from cache the table where data typically gathered with `tw_get_wikipedia_page_links()` are stored.

**Usage**

```r
tw_reset_wikipedia_page_links_cache(
  language = tidywikidatar::tw_get_language(),
  cache = NULL,
  cache_connection = NULL,
  disconnect_db = TRUE,
  ask = TRUE
)
```

**Arguments**

- `language` Defaults to language set with `tw_set_language()`; if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all
- `cache` Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.
- `cache_connection` Defaults to NULL. If NULL, and caching is enabled, `tidywikidatar` will use a local sqlite database by default. A custom connection to other databases can be given (see vignette caching for details).
- `disconnect_db` Defaults to TRUE. If FALSE, leaves the connection to cache open.
- `ask` Logical, defaults to TRUE. If FALSE, and cache folder does not exist, it just creates it without asking (useful for non-interactive sessions).

**Value**

Nothing, used for its side effects.

**Examples**

```r
if (interactive()) {
  tw_reset_wikipedia_page_links_cache()
}
```
tw_reset_wikipedia_page_sections_cache

Reset Wikipedia page link cache

Description

Removes from cache the table where data typically gathered with `tw_get_wikipedia_page_sections()` are stored.

Usage

```r
# R

tw_reset_wikipedia_page_sections_cache(
  language = tidywikidatar::tw_get_language(),
  cache = NULL,
  cache_connection = NULL,
  disconnect_db = TRUE,
  ask = TRUE
)
```

Arguments

- **language**: Defaults to language set with `tw_set_language()`; if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all.
- **cache**: Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.
- **cache_connection**: Defaults to NULL. If NULL, and caching is enabled, `tidywikidatar` will use a local sqlite database by default. A custom connection to other databases can be given (see vignette caching for details).
- **disconnect_db**: Defaults to TRUE. If FALSE, leaves the connection to cache open.
- **ask**: Logical, defaults to TRUE. If FALSE, and cache folder does not exist, it just creates it without asking (useful for non-interactive sessions).

Value

Nothing, used for its side effects.

Examples

```r
if (interactive()) {
  tw_reset_wikipedia_page_sections_cache()
}
```
tw_search  

Search for Wikidata items or properties and return Wikidata id, label, and description.

Description

By defaults, this search returns items. Set type to property or use `tw_search_property()` for properties.

Usage

```r
tw_search(
  search,
  type = "item",
  language = tidywikidatar::tw_get_language(),
  response_language = tidywikidatar::tw_get_language(),
  limit = 10,
  include_search = FALSE,
  wait = 0,
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE
)
```

Arguments

- **search**  
  A string to be searched in Wikidata

- **type**  
  Defaults to "item". Either "item" or "property".

- **language**  
  Language to be used for the search. Can be set once per session with `tw_set_language()`. If not set, defaults to "en". For a full list, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all

- **response_language**  
  Language to be used for the returned labels and descriptions. Corresponds to the `uselang` parameter of the MediaWiki API: https://www.wikidata.org/w/api.php?action=help&modules=wbsearchentities. Can be set once per session with `tw_set_language()`. If not set, defaults to "en". For a full list, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all

- **limit**  
  Maximum numbers of responses to be given.

- **include_search**  
  Logical, defaults to FALSE. If TRUE, the search is returned as an additional column.

- **wait**  
  In seconds, defaults to 0. Time to wait between queries to Wikidata. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.

- **cache**  
  Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.
tw_search_item

overwrite_cache
Defaults to FALSE. If TRUE, overwrites cache.

cache_connection
Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

disconnect_db
Defaults to TRUE. If FALSE, leaves the connection to cache open.

Value
A data frame (a tibble) with three columns (id, label, and description), and as many rows as there are results (by default, limited to 10). Four columns when include_search is set to TRUE.

Examples

tw_search(search = c("Margaret Mead", "Ruth Benedict"))

Description
This search returns only items, use tw_search_property() for properties.

Usage

tw_search_item(
  search,
  language = tidywikidatar::tw_get_language(),
  response_language = tidywikidatar::tw_get_language(),
  limit = 10,
  include_search = FALSE,
  wait = 0,
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE
)

Arguments

search
A string to be searched in Wikidata

language
Language to be used for the search. Can be set once per session with tw_set_language(). If not set, defaults to "en". For a full list, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes.
tw_search_property

response_language
  Language to be used for the returned labels and descriptions. Corresponds to the use\text{lang} parameter of the MediaWiki API: https://www.wikidata.org/w/api.php?action=help&modules=wbsearchentities. Can be set once per session with \texttt{tw_set_language()}. If not set, defaults to "en". For a full list, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all

limit
  Maximum numbers of responses to be given.

include_search
  Logical, defaults to FALSE. If TRUE, the search is returned as an additional column.

wait
  In seconds, defaults to 0. Time to wait between queries to Wikidata. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.

cache
  Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with \texttt{tw_enable_cache()} or \texttt{tw_disable_cache()}.

overwrite_cache
  Defaults to FALSE. If TRUE, overwrites cache.

cache_connection
  Defaults to NULL. If NULL, and caching is enabled, \texttt{tidywikidatar} will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

disconnect_db
  Defaults to TRUE. If FALSE, leaves the connection to cache open.

Value
  A data frame (a tibble) with three columns (id, label, and description), and as many rows as there are results (by default, limited to 10).

Examples
  \begin{verbatim}
  tw_search_item(search = "Sylvia Pankhurst")
  \end{verbatim}

Description
  This search returns only properties, use \texttt{tw_search_items()} for properties.

Usage
  \begin{verbatim}
  tw_search_property(
    search,
    language = tidywikidatar::tw_get_language(),
    response_language = tidywikidatar::tw_get_language(),
    limit = 10,
    include_search = FALSE,
  )
  \end{verbatim}
tw_search_property

wait = 0,
cache = NULL,
overwrite_cache = FALSE,
cache_connection = NULL,
disconnect_db = TRUE
)

Arguments

- **search** A string to be searched in Wikidata
- **language** Language to be used for the search. Can be set once per session with `tw_set_language()`. If not set, defaults to "en". For a full list, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all
- **response_language** Language to be used for the returned labels and descriptions. Corresponds to the `uselang` parameter of the MediaWiki API: https://www.wikidata.org/w/api.php?action=help&modules=wbsearchentities. Can be set once per session with `tw_set_language()`. If not set, defaults to "en". For a full list, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all
- **limit** Maximum numbers of responses to be given.
- **include_search** Logical, defaults to FALSE. If TRUE, the search is returned as an additional column.
- **wait** In seconds, defaults to 0. Time to wait between queries to Wikidata. If data are cached locally, wait time is not applied. If you are running many queries systemically you may want to add some waiting time between queries.
- **cache** Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.
- **overwrite_cache** Defaults to FALSE. If TRUE, overwrites cache.
- **cache_connection** Defaults to NULL. If NULL, and caching is enabled, `tidywikidata` will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).
- **disconnect_db** Defaults to TRUE. If FALSE, leaves the connection to cache open.

Value

A data frame (a tibble) with three columns (id, label, and description), and as many rows as there are results (by default, limited to 10).

Examples

```r
tw_search_property(search = "gender")
```
tw_search_single

Search for Wikidata items or properties and return Wikidata id, label, and description.

Description

This search returns only items, use `tw_search_property()` for properties.

Usage

```r
tw_search_single(
  search,
  type = "item",
  language = tidywikidatar::tw_get_language(),
  response_language = tidywikidatar::tw_get_language(),
  limit = 10,
  include_search = FALSE,
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE,
  wait = 0
)
```

Arguments

- **search**: A string to be searched in Wikidata
- **type**: Defaults to "item". Either "item" or "property".
- **language**: Language to be used for the search. Can be set once per session with `tw_set_language()`.
- **response_language**: Language to be used for the returned labels and descriptions. Corresponds to the `uselang` parameter of the MediaWiki API: https://www.wikidata.org/w/api.php?action=help&modules=wbsearchentities. Can be set once per session with `tw_set_language()`. If not set, defaults to "en". For a full list, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all
- **limit**: Maximum numbers of responses to be given.
- **include_search**: Logical, defaults to FALSE. If TRUE, the search is returned as an additional column.
- **cache**: Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.
- **overwrite_cache**: Defaults to FALSE. If TRUE, overwrites cache.
- **cache_connection**: Defaults to NULL. If NULL, and caching is enabled, `tidywikidatar` will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).
### tw_set_cache_db

**disconnect_db**  Defaults to TRUE. If FALSE, leaves the connection to cache open.

**wait**  In seconds, defaults to 0. Time to wait between queries to Wikidata. If data are cached locally, wait time is not applied. If you are running many queries systematically you may want to add some waiting time between queries.

#### Value

A data frame (a tibble) with three columns (id, label, and description), and as many rows as there are results (by default, limited to 10). Four columns when `include_search` is set to TRUE.

#### Examples

```r
tidywikidatar::tw_search_single(search = "Sylvia Pankhurst")
```

---

### Description

Set database connection settings for the session

### Usage

```r
tw_set_cache_db(
  db_settings = NULL,
  driver = NULL,
  host = NULL,
  server = NULL,
  port = NULL,
  database = NULL,
  user = NULL,
  pwd = NULL
)
```

### Arguments

- **db_settings**  A list of database connection settings (see example)
- **driver**  A database driver. Common database drivers include MySQL, PostgreSQL, and MariaDB. See `unique(odbc::odbcListDrivers()[[1]])` for a list of locally available drivers.
- **host**  Host address, e.g. "localhost". Different drivers use server or host parameter, only one of them is likely needed.
- **server**  Server address, e.g. "localhost". Different drivers use server or host parameter, only one of them is likely needed.
- **port**  Port to use to connect to the database.
- **database**  Database name.
- **user**  Database user name.
- **pwd**  Password for the database user.
tw_set_cache_folder

Value

A list with all given parameters (invisibly).

Examples

```r
if (interactive()) {
  # Settings can be provided either as a list
  db_settings <- list(
    driver = "MySQL",
    host = "localhost",
    server = "localhost",
    port = 3306,
    database = "tidywikidatar",
    user = "secret_username",
    pwd = "secret_password"
  )

  tw_set_cache_db(db_settings)

  # or as parameters

  tw_set_cache_db(
    driver = "MySQL",
    host = "localhost",
    server = "localhost",
    port = 3306,
    database = "tidywikidatar",
    user = "secret_username",
    pwd = "secret_password"
  )

  # or ignoring fields that can be left to default values, such as "localhost" and port 3306

  tw_set_cache_db(
    driver = "MySQL",
    database = "tidywikidatar",
    user = "secret_username",
    pwd = "secret_password"
  )
}
```

---

**tw_set_cache_folder**     Set folder for caching data

Description

Consider using a folder out of your current project directory, e.g. `tw_set_cache_folder("~/R/tw_data/")`: you will be able to use the same cache in different projects, and prevent cached files from being sync-ed if you use services such as Nextcloud or Dropbox.
### tw_set_cache_folder

**Usage**

```r
tw_set_cache_folder(path = NULL)
```

```r
tw_get_cache_folder(path = NULL)
```

**Arguments**

- `path` A path to a location used for caching data. If the folder does not exist, it will be created.

**Value**

The path to the caching folder, if previously set; the same path as given to the function; or the default, `tw_data` is none is given.

**Examples**

```r
if (interactive()) {
  tw_set_cache_folder(fs::path(fs::path_home_r(), "R", "tw_data"))
}

tw_get_cache_folder()
```

### tw_set_language

**Description**

Defaults to "en".

**Usage**

```r
tw_set_language(language = NULL)
```

```r
tw_get_language(language = NULL)
```

**Arguments**

- `language` A character vector of length one, with a string of two letters such as "en". For a full list of available values, see: https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all

**Value**

A two letter code for the language, if previously set; the same language as given to the function; or the default, en is none is given.
Examples

```r
if (interactive()) {
  tw_set_language(language = "en")
}

tw_get_language()
```

tw_test_items  

---

**tw_test_items**  

A list mostly used for testing with some Wikidata items in the format resulting from `WikidataR::get_item()`

---

**Description**

A list mostly used for testing with some Wikidata items in the format resulting from `WikidataR::get_item()`

**Usage**

```r
tw_test_items
```

**Format**

A list, an object such as the one resulting from `WikidataR::get_item()`

---

**tw_write_item_to_cache**  

Writes item to cache

---

**Description**

Writes item to cache. Typically used internally, but exported to enable custom caching solutions.

**Usage**

```r
tw_write_item_to_cache(
  item_df,
  language = tidywikidatar::tw_get_language(),
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE
)
```
tw_write_item_to_cache

Arguments

item_df A data frame with three columns typically generated with `tw_get()`.  
language Defaults to language set with `tw_set_language()`; if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all  
cache Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.  
overwrite_cache Logical, defaults to FALSE. If TRUE, it first deletes all rows associated with the item(s) included in `item_df`. Useful if the original Wikidata object has been updated.  
cache_connection Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).  
disconnect_db Defaults to TRUE. If FALSE, leaves the connection to cache open.

Value

Nothing, used for its side effects.

Examples

tw_set_cache_folder(path = fs::path(tempdir(), paste(sample(letters, 24), collapse = "")))
tw_create_cache_folder(ask = FALSE)
tw_disable_cache()

df_from_api <- tw_get(id = "Q180099", language = "en")

df_from_cache <- tw_get_cached_item(id = "Q180099", language = "en")

is.null(df_from_cache) # expect TRUE, as nothing has yet been stored in cache

tw_write_item_to_cache(  
  item_df = df_from_api,  
  language = "en",  
  cache = TRUE
)

df_from_cache <- tw_get_cached_item(  
  id = "Q180099",  
  language = "en",  
  cache = TRUE
)

is.null(df_from_cache) # expect a data frame, same as df_from_api
Write WikiData identifier (qid) of Wikipedia page to cache

Description
Mostly used internally by tidywikidatar, use with caution to keep caching consistent.

Usage
```r
tw_write_qid_of_wikipedia_page_to_cache(
  df,
  language = tidywikidatar::tw_get_language(),
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE
)
```

Arguments
- `df`: A data frame typically generated with `tw_get_wikipedia_page_qid()`.
- `language`: Defaults to language set with `tw_set_language()`; if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all
- `cache`: Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.
- `overwrite_cache`: Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.
- `cache_connection`: Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).
- `disconnect_db`: Defaults to TRUE. If FALSE, leaves the connection to cache open.

Value
Silently returns the same data frame provided as input. Mostly used internally for its side effects.

Examples
```r
if (interactive()) {
  df <- tw_get_wikipedia_page_qid(
    title = "Margaret Mead",
    language = "en",
  )
  }
**tw_write_qualifiers_to_cache**

Write qualifiers to cache

**Description**

Mostly to be used internally by `tidywikidatar`, use with caution to keep caching consistent.

**Usage**

```r
tw_write_qualifiers_to_cache(
  qualifiers_df,
  language = tidywikidatar::tw_get_language(),
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE
)
```

**Arguments**

- `qualifiers_df` A data frame typically generated with `tw_get_qualifiers()`.
- `language` Defaults to language set with `tw_set_language()`; if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all
- `cache` Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.
- `overwrite_cache` Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.
- `cache_connection` Defaults to NULL. If NULL, and caching is enabled, `tidywikidatar` will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).
- `disconnect_db` Defaults to TRUE. If FALSE, leaves the connection to cache open.

**Value**

Silently returns the same data frame provided as input. Mostly used internally for its side effects.
Examples

```r
q_df <- tw_get_qualifiers(
  id = "Q180099",
  p = "P26",
  language = "en",
  cache = FALSE
)

tw_write_qualifiers_to_cache(
  qualifiers_df = q_df,
  language = "en",
  cache = TRUE
)
```

```
tw_write_search_to_cache

Writes search to cache
```

Description

Writes search to cache. Typically used internally, but exported to enable custom caching solutions.

Usage

```r
tw_write_search_to_cache(
  search_df,
  type = "item",
  language = tidywikidatar::tw_get_language(),
  response_language = tidywikidatar::tw_get_language(),
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE
)
```

Arguments

- **search_df**: A data frame with four columns typically generated with `tw_search(include_search = TRUE)`.
- **type**: Defaults to "item". Either "item" or "property".
- **language**: Language to be used for the search. Can be set once per session with `tw_set_language()`.
  If not set, defaults to "en". For a full list, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all.
- **response_language**: Language to be used for the returned labels and descriptions. Corresponds to the `uselang` parameter of the MediaWiki API: https://www.wikidata.org/w/api.php?action=help&modules=wbsearchentities.
  Can be set once per session with `tw_set_language()`.
  If not set, defaults to "en". For a full list, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all.
cache Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.

overwrite_cache Defaults to FALSE. If TRUE, overwrites cache.

cache_connection Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

disconnect_db Defaults to TRUE. If FALSE, leaves the connection to cache open.

Value

Nothing, used for its side effects.

Examples

```r
tw_set_cache_folder(path = fs::path(tempdir(), paste(sample(letters, 24), collapse = "")))
tw_create_cache_folder(ask = FALSE)
tw_disable_cache()

search_from_api <- tw_search(search = "Sylvia Pankhurst", include_search = TRUE)
search_from_cache <- tw_get_cached_search("Sylvia Pankhurst")
nrow(search_from_cache) == 0 # expect TRUE, as nothing has yet been stored in cache
tw_write_search_to_cache(search_df = search_from_api)
search_from_cache <- tw_get_cached_search("Sylvia Pankhurst")
search_from_cache
```

Description

Mostly used internally by tidywikidatar, use with caution to keep caching consistent.

Usage

```r
tw_write_wikipedia_category_members_to_cache(
  df,
  language = tidywikidatar::tw_get_language(),
  type = "page",
  cache = NULL,
```

Write Wikipedia page links to cache
overwrite_cache = FALSE,
cache_connection = NULL,
disconnect_db = TRUE
)

Arguments

df A data frame typically generated with `tw_get_wikipedia_category_members()`.
language Defaults to language set with `tw_set_language()`; if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all

type Defaults to "page", defines which kind of members of a category are returned. Valid values include "page", "file", and "subcat" (for sub-category). Corresponds to cmtype. For details, see https://www.mediawiki.org/wiki/API:Categorymembers

cache Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with `tw_enable_cache()` or `tw_disable_cache()`.

overwrite_cache Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.

cache_connection Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).

disconnect_db Defaults to TRUE. If FALSE, leaves the connection to cache open.

Value

Silently returns the same data frame provided as input. Mostly used internally for its side effects.

Examples

```r
if (interactive()) {
  df <- tw_get_wikipedia_category_members(
    category = "American women anthropologists",
    language = "en",
    cache = FALSE
  )

tw_write_wikipedia_category_members_to_cache(
  df = df,
  language = "en"
)
}
```
tw_write_wikipedia_page_links_to_cache

Write Wikipedia page links to cache

Description

Mostly used internally by tidywikidatar, use with caution to keep caching consistent.

Usage

tw_write_wikipedia_page_links_to_cache(
  df,
  language = tidywikidatar::tw_get_language(),
  cache = NULL,
  overwrite_cache = FALSE,
  cache_connection = NULL,
  disconnect_db = TRUE
)

Arguments

df A data frame typically generated with tw_get_wikipedia_page_links().
language Defaults to language set with tw_set_language(); if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all
Cache Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with tw_enable_cache() or tw_disable_cache().
overwrite_cache Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.
cache_connection Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).
disconnect_db Defaults to TRUE. If FALSE, leaves the connection to cache open.

Value

Silently returns the same data frame provided as input. Mostly used internally for its side effects.

Examples

if (interactive()) {
  df <- tw_get_wikipedia_page_links(
    title = "Margaret Mead",
    language = "en",
    cache = FALSE
  )
}

Description

Mostly used internally by tidywikidatar, use with caution to keep caching consistent.

Usage

tw_write_wikipedia_page_sections_to_cache(
    df,  
    language = tidywikidatar::tw_get_language(),  
    cache = NULL,  
    overwrite_cache = FALSE,  
    cache_connection = NULL,  
    disconnect_db = TRUE  
)

Arguments

df  A data frame typically generated with tw_get_wikipedia_page_sections().
language  Defaults to language set with tw_set_language(); if not set, "en". Use "all_available" to keep all languages. For available language values, see https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all.
cache  Defaults to NULL. If given, it should be given either TRUE or FALSE. Typically set with tw_enable_cache() or tw_disable_cache().
overwrite_cache  Logical, defaults to FALSE. If TRUE, it overwrites the table in the local sqlite database. Useful if the original Wikidata object has been updated.
cache_connection  Defaults to NULL. If NULL, and caching is enabled, tidywikidatar will use a local sqlite database. A custom connection to other databases can be given (see vignette caching for details).
disconnect_db  Defaults to TRUE. If FALSE, leaves the connection to cache open.

Value

Silently returns the same data frame provided as input. Mostly used internally for its side effects.
Examples

```r
if (interactive()) {
  df <- tw_get_wikipedia_page_sections(
    title = "Margaret Mead",
    language = "en",
    cache = FALSE
  )

  tw_write_wikipedia_page_sections_to_cache(
    df = df,
    language = "en"
  )
}
```
Index

* datasets
  tw_empty_image_metadata, 13
  tw_empty_item, 14
  tw_empty_qualifiers, 14
  tw_empty_search, 14
  tw_empty_wikipedia_category_members, 15
  tw_empty_wikipedia_page, 15
  tw_empty_wikipedia_page_links, 16
  tw_empty_wikipedia_page_sections, 16
  tw_qid_airports, 81
  tw_qid_meps, 82
  tw_test_items, 98

tw_check_cache, 4
  tw_check_cache_folder, 5
  tw_check_cache_index, 6
  tw_check_cached_items, 4
  tw_check_qid, 8
  tw_check_search, 9
  tw_connect_to_cache, 10
  tw_create_cache_folder, 11
  tw_disable_cache, 12
  tw_disconnect_from_cache, 12
  tw_empty_image_metadata, 13
  tw_empty_item, 14
  tw_empty_qualifiers, 14
  tw_empty_search, 14
  tw_empty_wikipedia_category_members, 15
  tw_empty_wikipedia_page, 15
  tw_empty_wikipedia_page_links, 16
  tw_empty_wikipedia_page_sections, 16
  tw_enable_cache, 17
  tw_extract_qualifier, 17
  tw_extract_single, 18
  tw_filter, 19
  tw_filter_first, 20
  tw_filter_people, 21
  tw_get, 22
  tw_get_all_with_p, 24
  tw_get_cache_db, 33
  tw_get_cache_file, 34
  tw_get_cache_folder
    (tw_set_cache_folder), 96
  tw_get_cache_table_name, 34
  tw_get_cached_item, 25
  tw_get_cached_qualifiers, 26
  tw_get_cached_search, 27
  tw_get_cached_wikipedia_category_members, 28
  tw_get_cached_wikipedia_page_links, 30
  tw_get_cached_wikipedia_page_qid, 31
  tw_get_cached_wikipedia_page_sections, 32
  tw_get_description, 35
  tw_get_field, 36
  tw_get_image, 37
  tw_get_image_metadata, 39
  tw_get_image_metadata_single, 40
  tw_get_image_same_length, 42
  tw_get_label, 43
  tw_get_language (tw_set_language), 97
  tw_get_p (tw_get_property_same_length), 51
  tw_get_p1, 45
  tw_get_p_wide, 55
  tw_get_property, 46
  tw_get_property_description, 48
  tw_get_property_label, 49
  tw_get_property_label_single, 50
  tw_get_property_same_length, 51
  tw_get_property_with_details, 53
  tw_get_property_with_details_single, 54
  tw_get_qualifiers, 57
  tw_get_qualifiers_single, 58
  tw_get_single, 60