

Package ‘tndata’

March 19, 2025

Title Fetch Datasets from the Official Tunisian Data Catalog

Version 0.1.0

Description Simplifies access to Tunisian government open data from [<https://data.gov.tn/fr/>](https://data.gov.tn/fr/). Queries datasets by theme, author, or keywords, retrieves metadata, and gets structured results ready for analysis; all through the official 'CKAN' API.

License Apache License (>= 2)

Encoding UTF-8

Imports dplyr, fs, glue, httr, httr2, jsonlite, logger, lubridate, purrr, stringr

Depends R (>= 4.1.0)

RoxygenNote 7.3.2

URL <https://github.com/aymennasri/tndata>, <http://aymennasri.me/tndata/>

BugReports <https://github.com/aymennasri/tndata/issues>

NeedsCompilation no

Author Aymen Nasri [aut, cre, cph]

Maintainer Aymen Nasri <aymennasri@proton.me>

Repository CRAN

Date/Publication 2025-03-19 13:20:02 UTC

Contents

download_dataset	2
get_authors	2
get_datasets	3
get_keywords	4
get_organizations	5
get_themes	5
Index	7

download_dataset	<i>Download Dataset</i>
------------------	-------------------------

Description

Downloads a dataset from the Tunisian data catalog API (data.gov.tn).

Usage

```
download_dataset(title, download_dir, format = NULL)
```

Arguments

title	Character. Display name of the dataset to download.
download_dir	Character. Directory to save the downloaded dataset at.
format	Character. Format of the dataset to download.

Value

The demanded dataset in the demanded path.

Examples

```
try({  
  download_dataset(  
    "Ressources en eau- Gouvernorat de Kasserine",  
    format = "xls",  
    download_dir = tempdir()  
  )  
})
```

get_authors	<i>List All Authors in the Dataset Catalog</i>
-------------	--

Description

Retrieves a list of all authors who have contributed datasets to the catalog by fetching datasets and extracting unique author information.

Usage

```
get_authors(max_datasets = 1000)
```

Arguments

max_datasets Numeric. Maximum number of datasets to retrieve for author extraction, defaults to 1000.

Value

A tibble (data frame) with the following columns:

name Character. Name of the author.

count Numeric. Number of datasets contributed by this author.

Examples

```
try({
  authors <- get_authors(max_datasets = 100)
  head(authors)
})
```

get_datasets

List Available Datasets

Description

Fetches datasets by keyword and/or author

Usage

```
get_datasets(
  keyword = NULL,
  author = NULL,
  organization = NULL,
  max_results = 100
)
```

Arguments

keyword Character. Keyword to search for in dataset titles.

author Character. Author name to filter datasets by.

organization Character. Organization name to filter datasets by.

max_results Numeric. Maximum number of datasets to return, defaults to 100.

Value

A tibble (data frame) with the following columns:

title Character. Title of the dataset.

id Character. ID of the dataset.

resources List. A list of tibbles with the following columns:

name Character. Name of the dataset.

format Character. Format of the dataset.

url Character. URL of the dataset.

created Date. Date the dataset was created.

Examples

```
try({  
  datasets <- get_datasets("agriculture")  
  head(datasets)  
})
```

get_keywords

List Dataset Keywords/Tags

Description

Retrieves a list of unique keywords/tags from the Tunisian data catalog API.

Usage

```
get_keywords(limit = 10, query = NULL)
```

Arguments

limit Integer. Maximum number of tags to return (default: 10).

query Character. Optional search string to filter tags.

Value

A data frame of keywords/tags with counts.

Examples

```
try({  
  get_keywords(limit = 10)  
})
```

get_organizations *List Organizations*

Description

Retrieves organizations data from the Tunisian data catalog API (data.gov.tn) using faceted search. This function returns organizations that have published datasets.

Usage

```
get_organizations(min_count = 1)
```

Arguments

min_count Integer. Minimum number of datasets an organization must have to be included in results. Default is 1, meaning only organizations with at least one dataset are returned.

Value

A tibble (data frame) with the following columns:

name Character. Machine-readable name/identifier of the organization.

display_name Character. Human-readable name of the organization.

dataset_count Integer. Number of datasets published by the organization.

Examples

```
try({  
  # Get all organizations with at least 5 datasets  
  orgs <- get_organizations(min_count = 5)  
  head(orgs)  
})
```

get_themes *Summarize Dataset Themes*

Description

Fetches and summarizes themes (groups) alongside the number of datasets in each theme from the Tunisian data catalog API (data.gov.tn).

Usage

```
get_themes()
```

Value

A tibble (data frame) with two columns:

theme Character. Name of the theme/group.

dataset_count Numeric. Number of datasets in the theme.

Examples

```
try({  
  themes_summary <- get_themes()  
  head(themes_summary)  
})
```

Index

`download_dataset`, 2

`get_authors`, 2

`get_datasets`, 3

`get_keywords`, 4

`get_organizations`, 5

`get_themes`, 5