Package: mospiR (via r-universe)

October 28, 2026

Type Package

Title Client for the 'MoSPI' Microdata Portal

Version 0.1.1.9000

Description Download unit-level survey microdata from the Ministry of Statistics and Programme Implementation ('MoSPI', Government of India) Microdata Portal <https://microdata.gov.in>. Provides functions to list datasets, list files within a dataset, and download individual files or entire datasets through the portal's 'NADA' API.

License MIT + file LICENSE

Encoding UTF-8

Language en-US

Depends R (>= 4.1.0)

Imports httr2 (>= 1.0.0)

Suggests archive, testthat (>= 3.0.0), withr, knitr, rmarkdown

VignetteBuilder knitr

Config/testthat/edition 3

Config/Needs/website pkgdown

URL https://mospir.saketlab.org, https://github.com/saketkc/mospiR

BugReports https://github.com/saketkc/mospiR/issues

Config/roxygen2/version 8.1.1

Config/pak/sysreqs libssl-dev

Repository https://saketkc.r-universe.dev

Date/Publication 2026-10-18 15:09:04 UTC

RemoteUrl https://github.com/saketkc/mospir

RemoteRef HEAD

RemoteSha 27dbf1522de945fbbe5dcd43517580bfc7a98e06

Contents

download_dataset 2
download_file 3
extract_archive 3
list_datasets 4
list_files 5

Index **6**

`download_dataset` *Download all files for a dataset*

Description

Download all files for a dataset

Usage

```
download_dataset(dataset_id, folder_path, api_key, overwrite = FALSE)
```

Arguments

`dataset_id`	Dataset ‘idno‘.
`folder_path`	Destination directory. The function makes the directory if it does not exist.
`api_key`	MoSPI API key.
`overwrite`	If ‘FALSE‘ (default), the function does not download the files that are already in ‘folder_path‘.

Value

A character vector of the paths of the saved files, as an invisible value.

Examples

```
# This dataset is approximately 270 MB.
paths <- download_dataset(
  "DDI-IND-NSO-ASI-2020-21", tempdir(), Sys.getenv("MOSPI_KEY")
)
```

download_file — *Download a single file from a dataset*

Description

Download a single file from a dataset

Usage

```
download_file(dataset_id, file_name, folder_path, api_key, .files_cache = NULL)
```

Arguments

`dataset_id`	Dataset ‘idno‘.
`file_name`	File name from the ‘name‘ column of [list_files()].
`folder_path`	Destination directory. The function makes the directory if it does not exist.
`api_key`	MoSPI API key.
`.files_cache`	A file list from [list_files()]. If ‘NULL‘ (default), the function gets the file list from the portal. Use this argument when you download many files from one dataset. This prevents more API requests.

Value

The path of the saved file, as an invisible value. ‘NULL‘ if the download is not successful.

Examples

```
path <- download_file(
  "DDI-IND-NSO-ASI-2020-21", "DDI-IND-NSO-ASI-2020-21.xml",
  tempdir(), Sys.getenv("MOSPI_KEY")
)
```

extract_archive — *Extract a downloaded archive*

Description

Extracts a ‘.zip‘, ‘.rar‘, ‘.7z‘ or ‘.tar‘ file from the portal. The function uses the archive package first. If the archive package is not installed or cannot extract the file, the function uses the command line tool ‘unar‘.

Usage

```
extract_archive(path, dir)
```

Arguments

`path`	Path of the archive file.
`dir`	Destination directory. The function makes the directory if it does not exist.

Details

On Linux, the archive package uses the libarchive of the system. libarchive versions before 3.7.4 (for example 3.7.2 in Ubuntu 24.04) give a "File CRC error" on some ‘.rar‘ files from the portal. To extract these files, install ‘unar‘ (‘sudo apt-get install unar‘ on Ubuntu, ‘brew install unar‘ on macOS).

The function extracts into a temporary directory and then moves the files into ‘dir‘. Thus, if the extraction fails, ‘dir‘ does not change, and the function does not make ‘dir‘. Files in ‘dir‘ that have the same names as files in the archive are replaced.

Value

A character vector of the paths of the extracted files, as an invisible value. ‘NULL‘ if the extraction is not successful.

Examples

```
src <- file.path(tempdir(), "example.txt")
writeLines("example", src)
zip_file <- file.path(tempdir(), "example.zip")
archive::archive_write_files(zip_file, src)
extract_archive(zip_file, file.path(tempdir(), "extracted"))
```

`list_datasets` *List datasets from the MoSPI Microdata Portal*

Description

List datasets from the MoSPI Microdata Portal

Usage

```
list_datasets(api_key, page = NULL, query = NULL)
```

Arguments

`api_key`	MoSPI API key.
`page`	Page number to get. If ‘NULL‘ (default), the function gets all the pages.
`query`	Text to find in the dataset titles. Upper case and lower case letters are equal. The comparison occurs on your computer.

Value

A data frame with the columns ‘id‘, ‘idno‘, ‘title‘, and other columns. ‘NULL‘ if the request is not successful.

Examples

```
page1 <- list_datasets(Sys.getenv("MOSPI_KEY"), page = 1)
head(page1[, c("idno", "title")])
```

---

list_files *List files available for a dataset*

Description

List files available for a dataset

Usage

```
list_files(dataset_id, api_key)
```

Arguments

dataset_id	Dataset ‘idno‘, for example ‘"DDI-IND-NSO-ASI-2020-21"‘.
api_key	MoSPI API key.

Value

A data frame with the columns ‘name‘, ‘base64‘, ‘size‘, and other columns. ‘NULL‘ if the request is not successful.

Examples

```
files <- list_files("DDI-IND-NSO-ASI-2020-21", Sys.getenv("MOSPI_KEY"))
```

Index

download_dataset, 2
download_file, 3

extract_archive, 3

list_datasets, 4
list_files, 5