Package: mapmetadata (via r-universe)

October 29, 2026

Type Package

Title Map health metadata onto predefined research domains

Version 4.0.1

Description Prior to gaining full access to health datasets, explore publicly available metadata and map metadata onto predefined research domains. This package uses structural metadata files downloaded from the Health Data Research Gateway (https://healthdatagateway.org/en). In theory, any metadata file with the same structure as the files downloaded from this gateway can be used with this package, but the package has been developed and tested on metadata files from this gateway only.

License GPL (>= 3)

URL https://docs.ropensci.org/mapmetadata/

BugReports https://github.com/ropensci/mapmetadata/issues

Depends R (>= 2.10)

Imports cli, dplyr, ggplot2, gridExtra, htmlwidgets, plotly, tidyr

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

VignetteBuilder knitr

Config/Needs/build moodymudskipper/devtag

Config/testthat/edition 3

Encoding UTF-8

LazyData true

Roxygen list(markdown = TRUE, roclets = c(``collate'', ``rd'', ``namespace'', ``devtag::dev_roclet''))

RoxygenNote 7.3.3

Config/pak/sysreqs cmake make libicu-dev libuv1-dev libssl-dev

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

Date/Publication 2026-01-20 09:59:54 UTC

RemoteUrl https://github.com/ropensci/mapmetadata

RemoteRef HEAD

RemoteSha 2aba785d1bd8f58acdf391a35944710f63bb16e3

Contents

map_compare 2
map_convert 3
metadata_map 4

Index **7**

`map_compare` *map_compare*

Description

This function is to be used after running the metadata_map function.

It compares csv outputs from two sessions, finds their differences, and asks for a consensus.

Usage

```
map_compare(
  session_dir,
  session1_base,
  session2_base,
  metadata_file,
  domain_file,
  output_dir = session_dir,
  quiet = FALSE
)
```

Arguments

`session_dir`	This directory should contain 2 csv files for each session (LOG_ and OUTPUT_), 4 csv files in total.
`session1_base`	Base file name for session 1, see Example below.
`session2_base`	Base file name for session 2, see Example below.
`metadata_file`	The full path to the metadata file used when running metadata_map (should be the same for session 1 and session 2)
`domain_file`	The full path to the domain file used when running metadata_map (should be the same for session 1 and session 2)
`output_dir`	The path to the directory where the consensus output file will be saved. By default, the session_dir is used.
`quiet`	Default is FALSE. Change to TRUE to quiet the cli_alert_info and cli_alert_success messages.

Value

It returns a csv output, which represents the consensus decisions between session 1 and session 2

Examples

```
# Demo run requires no function inputs but requires user interaction.
# See package documentation to guide user inputs.
if(interactive()) {
    temp_output_dir <- tempdir()
    # Locate file paths for the example files in the package
    demo_session_dir <- system.file("outputs", package = "mapmetadata")
    demo_session1_base <- "360_NCCHD_CHILD_2025-02-14-18-14-01"
    demo_session2_base <- "360_NCCHD_CHILD_2025-02-14-18-17-47"
    demo_metadata_file <- system.file("inputs","360_NCCHD_Metadata.csv",
    package = "mapmetadata")
    demo_domain_file <- system.file("inputs","domain_list_demo.csv",
    package = "mapmetadata")

    map_compare(
    session_dir = demo_session_dir,
    session1_base = demo_session1_base,
    session2_base = demo_session2_base,
    metadata_file = demo_metadata_file,
    domain_file = demo_domain_file,
    output_dir = temp_output_dir
    )}
```

`map_convert` *map_convert*

Description

The 'MAPPING_' file groups multiple categorisations onto one line e.g. Domain_code could read '1,3'

This function creates a new longer output 'L-MAPPING_' which gives each categorisation its own row.

This 'L-MAPPING_' may be useful when using these csv files in later analyses

Usage

```
map_convert(
  csv_to_convert,
  csv_to_convert_dir,
  output_dir = csv_to_convert_dir,
  quiet = FALSE
)
```

Arguments

csv_to_convert	Name of 'MAPPING_' csv file created from metadata_map
csv_to_convert_dir	Location of csv_to_convert
output_dir	Location where the 'L-MAPPING_' csv file will be saved.
quiet	Default is FALSE. Change to TRUE to quiet the cli_alert_info and cli_alert_success messages. Default is csv_to_convert_dir.

Value

Returns 'L-MAPPING_' file in specified directory

Examples

```
# Locate file path and file name for the example files in the package
demo_csv_to_convert_dir <- system.file("outputs", package = "mapmetadata")
demo_csv_to_convert <- "MAPPING_360_NCCHD_CHILD_2025-02-14-18-14-01.csv"
temp_output_dir <- tempdir()
# Run the function
map_convert(
csv_to_convert = demo_csv_to_convert,
csv_to_convert_dir = demo_csv_to_convert_dir,
output_dir = temp_output_dir)
```

---

metadata_map *metadata_map*

Description

This function will read in the metadata file for a chosen dataset and create a summary plot. It will ask a user to select a table from this dataset to process, and loop through all the variables in this table, asking the user to map (categorise) each variable to one or more domains. The domains will appear in the Plots tab for the user's reference.

These categorisations will be saved to a csv file, alongside a log file which summarises the session details. To speed up this process, some auto-categorisations will be made by the function for commonly occurring variables, and categorisations for the same variable can be copied from one table to another.

Example inputs are provided within the package data, for the user to run this function in a demo mode. Refer to the package website for more guidance.

Usage

```
metadata_map(
  metadata_file = NULL,
  domain_file = NULL,
  look_up_file = NULL,
  output_dir = getwd(),
  table_copy = TRUE,
  long_output = TRUE,
  demo_number = 5,
  quiet = FALSE
)
```

Arguments

metadata_file	This should be a csv download from HDRUK gateway (in the form of ID_Dataset_Metadata.csv). Run '?mapmetadata::metadata' to see how the metadata_file for the demo was created.
domain_file	This should be a csv file created by the user, with two columns (Domain_Code and Domain_Name). Run '?mapmetadata::domain_list' to see how the domain_file for the demo was created.
look_up_file	The lookup file makes auto-categorisations intended for variables that appear regularly in health datasets. It only works for 1:1 mappings right now, i.e. variable should only be listed once in the file. Run '?mapmetadata::look_up' to see how the default look_up was created.
output_dir	The path to the directory where the two csv output files will be saved. Default is the current working directory.
table_copy	Turn on copying between tables (default TRUE). If TRUE, categorisations you made for all other tables in this dataset will be copied over (if 'OUTPUT_' files are found in output_dir). This can be useful when the same variables appear across multiple tables within one dataset; copying from one table to the next will save the user time, and ensure consistency of categorisations across tables.
long_output	Run map_convert.R to create a new longer output. Default is TRUE.
demo_number	How many table variables to loop through in the demo. Default is 5. 'LOUTPUT_' which gives each categorisation its own row. Default is TRUE.
quiet	Default is FALSE. Change to TRUE to quiet the cli_alert_info and cli_alert_success messages.

Value

A html plot summarising the dataset. Various csv and png outputs to summarise the user's mapping session for a specific table in the dataset.

Examples

```
# Demo run requires no function inputs but requires user interaction.
# See package documentation to guide user inputs.
if(interactive()) {
```

```
    temp_output_dir <- tempdir()
    metadata_map(output_dir = temp_output_dir)
}
```

Index

map_compare, 2
map_convert, 3
metadata_map, 4