

Managing REDCap Data: The R package REDCapDM

João Carmezim, Biostatistics Support and Research Unit
Germans Trias i Pujol Research Institute and Hospital (IGTP) - Badalona, Spain

useR! 2024 - July


REDCap

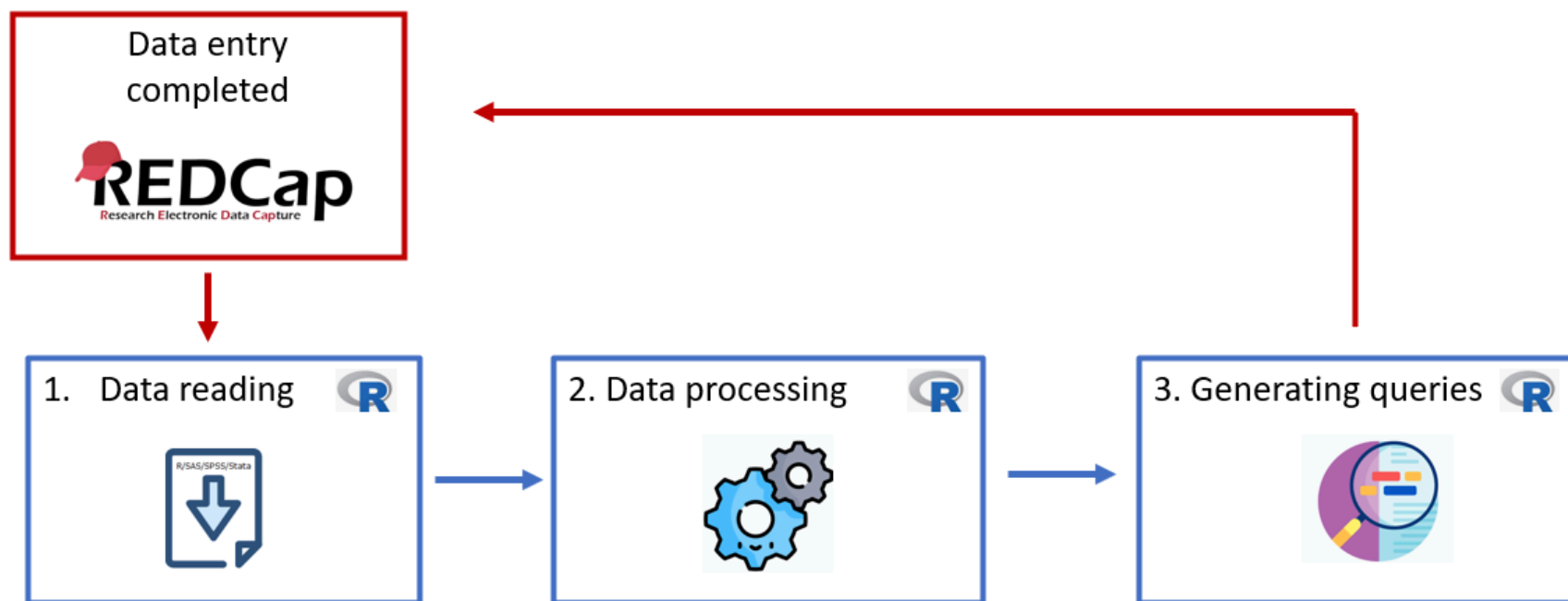


Secure web application for building and managing databases of research studies.

- Vanderbilt University (2004)
- License agreement at no charge to non-profit organizations

Motivation

Streamline and automate the processes that occur between completing data entry and obtaining a database ready for statistical analysis in :

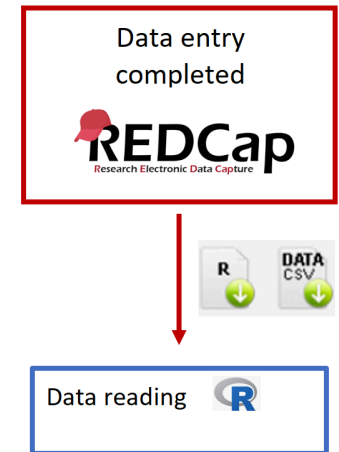


Data reading

(Download from REDCap)

1. Download the data (CSV) and R file from REDCap onto **the same folder**
2. Download the dictionary (codebook)
3. Download the instrument-event mapping (CSV) file
4. Use the `redcap_data` function:

```
dataset <- redcap_data(data_path = "C:/Users/username/example.r",  
                       dic_path = "C:/Users/username/example_dictionary.csv",  
                       event_path = "C:/Users/username/example_instrument_event.csv")
```



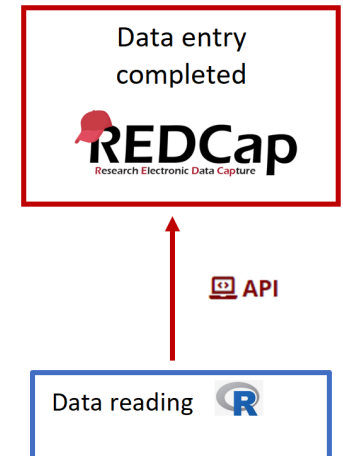
Data reading

(API connection)

1. Generate an API token for the project in REDCap
2. Use the `redcap_data` function:

```
dataset <- redcap_data(uri = "https://wsm.ics.gencat.cat/",  
                      token = "1234567890ABCDEFGHIJ")
```

Please note that the API token gives you special access to the REDCap project and should not be shared carelessly.



Process

(Data transformation)

The data processing is done using the `rd_transform` function:

```
dataset_transformed <- rd_transform(dataset)
```

Possible steps:

- Remove specific variables or variables with a certain pattern
- Recalculate every calculated field
- Transform checkboxes options into No/Yes
- Replace original variables for their factor version
- Convert each branching logic in the dictionary into R logic

Generate queries

(Function versatility)

We can decide to make multiple queries at the same time:

```
example <- rd_query(dataset,  
  variables = c("age", "resp_rate", "dm"),  
  expression = c("x<30", "x>50", "is.na(x)"))
```

Or apply the same query to several variables:

```
example <- rd_query(dataset,  
  variables = c("age", "resp_rate", "dm"),  
  expression = "is.na(x)")
```

Follow-up of queries

To perform a follow-up on the generated queries, we can use the `check_queries` function:

```
check <- check_queries(old = example$queries,  
                       new = new_example$queries)
```

The function compares the old query dataset with the newly created one and classifies each query as:

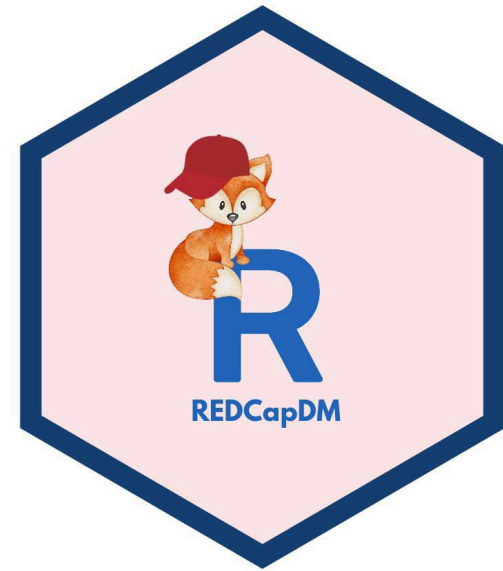
- New
- Solved
- Miscorrected
- Pending

Warning: queries classified as *Miscorrected* belong to the same combination of record identifier and variable name in both the old and new datasets, but with a different reason.

Conclusions

The REDCapDM package:

- Provides a set of useful functions for importing, organising and checking the quality of REDCap data.
- Facilitates a study's data management workflow and helps ensure reliable, high quality data that is ready for analysis.
- Fills a gap in the tools available in R for query generation of REDCap data.



Vignette

Link: <https://ubidi.github.io/REDCapDM/articles/REDCapDM.html>

REDCapDM 0.9.9

Get started

Reference

Articles ▾

Changelog

Search for



REDCapDM

Source: [vignettes/REDCapDM.Rmd](#)



On this page

Introduction

Functions

Built-in dataset

Examples

Introduction

The REDCapDM package allows users to read data exported directly from REDCap or via API connection. It also allows users to process the previously downloaded data, create reports of queries such as outliers or missing values and track the identified queries.

References

- Carmezim, J., Satorra, P., Peñafiel, J. et al. REDCapDM: An R package with a set of data management tools for a REDCap project, <https://doi.org/10.1186/s12874-024-02178-6>
-
- Carmezim J, Satorra P, Peñafiel J, García E, Pallarés N, Tebé C (2024). *REDCapDM: 'REDCap' Data Management*. R package version 0.9.9, <https://CRAN.R-project.org/package=REDCapDM>
 - Beasley W (2022). *REDCapR: Interaction Between R and REDCap*. R package version 1.1.0, <https://CRAN.R-project.org/package=REDCapR>
 - Wickham H, François R, Henry L, Müller K, Vaughan D (2023). *dplyr: A Grammar of Data Manipulation*. R package version 1.1.4, <https://CRAN.R-project.org/package=dplyr>
 - Wickham H (2023). *stringr: Simple, Consistent Wrappers for Common String Operations*. R package version 1.5.1, <https://CRAN.R-project.org/package=stringr>
 - Wickham H, Henry L (2023). *purrr: Functional Programming Tools*. R package version 1.0.2, <https://CRAN.R-project.org/package=purrr>