

Package: qcdd (via r-universe)

October 5, 2026

Title Build a Data Dictionary for Use with QAQC

Version 0.0.0.9000

Description Builds concept lookup tables from Connect data dictionaries to support quality assurance and quality control workflows. Combines concept identifiers and definitions across dictionary columns, removes duplicate entries and records with missing identifiers, and reports conflicting definitions for review.

License MIT + file LICENSE

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 8.0.0

Depends R (>= 4.1.0)

Imports dplyr, purrr, rio, rlang, tibble

Config/pak/sysreqs make libssl-dev libx11-dev zlib1g-dev

Repository <https://danielruss.r-universe.dev>

Date/Publication 2026-10-05 16:02:57 UTC

RemoteUrl <https://github.com/danielruss/qcdd>

RemoteRef main

RemoteSha 0d3c396407d078cc3d63cfb16a484ed267fa286d

Contents

add_concept_ids	2
cid_warning	2
load_dictionary_file	3
make_dictionary	3

Index	5
--------------	----------

add_concept_ids	<i>Build a concept lookup from dictionary columns</i>
-----------------	---

Description

Combine paired concept ID and definition columns into a two-column tibble. Column pairs are matched by position and must have equal lengths.

Usage

```
add_concept_ids(df, concept_ids, definitions)
```

Arguments

df	A data frame or other object convertible to a tibble containing the dictionary columns.
concept_ids	Bare concept ID column names, supplied as a single name or combined with <code>c()</code> .
definitions	Bare definition column names, supplied as a single name or combined with <code>c()</code> , in the same order as <code>concept_ids</code> .

Details

Identical ID-definition pairs and rows with missing concept IDs are removed after combining all column pairs. If an ID has conflicting definitions, a warning is issued and the first occurrence is retained, using column-pair order followed by input row order. Missing definitions are retained and may be selected as the first occurrence. When conflicts are found, the result is ordered by the concept ID groups.

Value

A tibble with `concept_id` and `definitions` columns and one row per non-missing concept ID. Conflicting definitions are stored in the `CID_warning` attribute and can be retrieved with [cid_warning\(\)](#).

cid_warning	<i>Retrieve conflicting concept definitions</i>
-------------	---

Description

Retrieve the conflict report stored by [add_concept_ids\(\)](#) without issuing another warning or modifying the lookup.

Usage

```
cid_warning(df)
```

Arguments

df A concept lookup returned by `add_concept_ids()`.

Value

A tibble with `concept_id` and `definitions` columns containing the distinct conflicting ID-definition pairs, including the definitions retained in the lookup. Returns NULL if the `CID_warning` attribute is absent, as when no conflicts were found.

load_dictionary_file	<i>Load the dictionary</i>
----------------------	----------------------------

Description

Load the dictionary

Usage

```
load_dictionary_file(path, ...)
```

Arguments

path Path to the dictionary

... additional parameters passed to `rio::import` and its underlying functions. For example the sheet name for an excel file

Value

a tibble containing the dictionary

make_dictionary	<i>Build a concept dictionary from a file</i>
-----------------	---

Description

Import a dictionary file with `load_dictionary_file()` and combine selected concept ID and definition columns with `add_concept_ids()`.

Usage

```
make_dictionary(path, concept_ids, definitions, ...)
```

Arguments

<code>path</code>	Path to the dictionary file.
<code>concept_ids</code>	Bare concept ID column names, supplied as a single name or combined with <code>c()</code> .
<code>definitions</code>	Bare definition column names, supplied as a single name or combined with <code>c()</code> , in the same order as <code>concept_ids</code> .
<code>...</code>	Additional arguments passed to <code>load_dictionary_file()</code> and ultimately to <code>rio::import()</code> , such as <code>sheet</code> and <code>col_types</code> for Excel.

Details

Column pairs must have equal lengths. Identical ID-definition pairs and rows with missing concept IDs are removed. Leading numeric value labels such as `1 =` are stripped from definitions. Conflicting definitions issue a warning, and the first occurrence is retained, using column-pair order followed by input row order. Missing definitions are retained. See `add_concept_ids()` for details.

Value

A tibble with `concept_id` and `definitions` columns and one row per non-missing concept ID. When conflicts occur, the conflict report is stored in the `CID_warning` attribute and can be retrieved with `cid_warning()`.

Examples

```
path <- tempfile(fileext = ".csv")
utils::write.csv(
  data.frame(
    concept_id_1 = c("72841936", "39508271"),
    definition_1 = c("Age", "Height"),
    concept_id_2 = c("86173409", "54296813"),
    definition_2 = c("Weight", "Blood pressure")
  ),
  path, row.names = FALSE
)
dictionary <- make_dictionary(
  path,
  concept_ids = c(concept_id_1, concept_id_2),
  definitions = c(definition_1, definition_2),
  colClasses = "character"
)
dictionary
unlink(path)
```

Index

`add_concept_ids`, [2](#)

`add_concept_ids()`, [2–4](#)

`cid_warning`, [2](#)

`cid_warning()`, [2, 4](#)

`load_dictionary_file`, [3](#)

`load_dictionary_file()`, [3, 4](#)

`make_dictionary`, [3](#)

`rio::import()`, [4](#)