Insert Cohort Definitions Into Package

2022-03-09

Contents

ROhdsiWebApi is part of HADES. # Introduction From Package Readme

ROhdsiWebApi maybe used to insert cohort definition json and sql from your Atlas (webapi) instance into your study R package. Two functions in this package support `insertCohortDefinitionInPackage` and `insertCohortDefinitionSetInPackage` . Please see documentation for the functions. This vignette focusses on the `insertCohortDefinitionsetInPackage` function. It depends on a reference file called CohortsToCreate.

The `insertCohortDefinitionSetInPackage` function extracts a set of cohort definitions from your Atlas (webapi) instance and inserts them into your R studio package. It requires as an input a csv file with information about the Cohorts that needs to be inserted called 'CohortsToCreate.csv'. This file should have three columns - altasId (The cohort Id in Atlas), cohortId (The cohort ID that will be used when instantiating the cohort (can be different from atlasId)), and name (The name to be used for the cohort. This name will be used to generate file names, so please use letters and numbers only (no spaces)). 'CohortsToCreate.csv' can take additional optional fields such as 'cohortName' but these optional fields will not be used by this function. It maybe used by your study package e.g. 'cohortName' maybe used to represent a custom name to identify the cohort in your study.

We can easily create 'CohortsToCreate.csv', including the optional fields as follows:

```r
library(magrittr)
# Set up
baseUrl <- Sys.getenv("BaseUrl")
# list of cohort ids
cohortIds <- c(18345,18346)

# get specifications for the cohortIds above
webApiCohorts <-
        ROhdsiWebApi::getCohortDefinitionsMetaData(baseUrl = baseUrl) %>%
        dplyr::filter(.data$id %in% cohortIds)

cohortsToCreate <- list()
for (i in (1:nrow(webApiCohorts))) {
  cohortId <- webApiCohorts$id[[i]]
  cohortDefinition <-
    ROhdsiWebApi::getCohortDefinition(cohortId = cohortId,
                                      baseUrl = baseUrl)
  cohortsToCreate[[i]] <- tidyr::tibble(
    atlasId = webApiCohorts$id[[i]],
```

```
    name = webApiCohorts$id[[i]],
    cohortName = stringr::str_trim(stringr::str_squish(cohortDefinition$name))
  )
}

cohortsToCreate <- dplyr::bind_rows(cohortsToCreate)

readr::write_excel_csv(x = cohortsToCreate, na = "",
                       file = "inst/settings/CohortsToCreate.csv",
                       append = FALSE)
```

Note in the step above - the only input needed was the list of atlasId, and ROhdsiWebApi was able to populate the rest of fields by pulling that information from your webApi (atlas). Once this file has been created, then use insert function as follows

```
# Insert cohort definitions from ATLAS into package -----------------------
ROhdsiWebApi::insertCohortDefinitionSetInPackage(
  fileName = "inst/settings/CohortsToCreate.csv",
  baseUrl = baseUrl,
  insertTableSql = TRUE,
  insertCohortCreationR = TRUE,
  generateStats = TRUE,
  packageName = 'yourFavoritePackage'
)
```

Now your package should have all the cohort specification in the 'inst' folder.

IMPORTANT: This function is designed to be used as an 'extra' script for your package maintenance, i.e. please call this function in R Studio in project mode - where the project is your package.