

Package ‘CdmInspection’

February 9, 2021

Title R Package to support quality control inspection of an OMOP-CDM instance

Version 1.0.2

Description

The CdmInspection R Package performs additional checks on top of the Data Quality Dashboard.

License Apache License 2.0

VignetteBuilder knitr

URL <https://github.com/EHDEN/CdmInspection>

BugReports <https://github.com/EHDEN/CdmInspection/issue>

Encoding UTF-8

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.1.1

Depends R (>= 3.5.0),
DatabaseConnector (>= 3.0.0),
ROhdsiWebApi (>= 1.1.2),
dplyr,
officer,
data.table,
flextable,
lubridate,
ggplot2

Imports parallel,
benchmarkme,
prettyunits,
magrittr,
tools

R topics documented:

cdmInspection	2
dataTablesChecks	3
performanceChecks	4
validateSchema	6
vocabularyChecks	7

Index	9
--------------	----------

cdmInspection	<i>The main CdmInspection analyses (for v5.x)</i>
---------------	---

Description

CdmInspection runs a list of checks as part of the CDM inspection procedure

Usage

```
cdmInspection(
  connectionDetails,
  cdmDatabaseSchema,
  resultsDatabaseSchema = cdmDatabaseSchema,
  scratchDatabaseSchema = resultsDatabaseSchema,
  vocabDatabaseSchema = cdmDatabaseSchema,
  oracleTempSchema = resultsDatabaseSchema,
  databaseName = "",
  analysisIds = "",
  createTable = TRUE,
  smallCellCount = 5,
  runVocabularyChecks = TRUE,
  runDataTablesChecks = TRUE,
  runPerformanceChecks = TRUE,
  runWebAPIChecks = TRUE,
  baseUrl,
  sqlOnly = FALSE,
  outputFolder = "output",
  verboseMode = TRUE
)
```

Arguments

connectionDetails

An R object of type connectionDetails created using the function createConnectionDetails in the DatabaseConnector package.

cdmDatabaseSchema

Fully qualified name of database schema that contains OMOP CDM schema. On SQL Server, this should specify both the database and the schema, so for example, on SQL Server, 'cdm_instance.dbo'.

resultsDatabaseSchema

Fully qualified name of database schema that we can write final results to. Default is cdmDatabaseSchema. On SQL Server, this should specify both the database and the schema, so for example, on SQL Server, 'cdm_results.dbo'.

vocabDatabaseSchema

String name of database schema that contains OMOP Vocabulary. Default is cdmDatabaseSchema. On SQL Server, this should specify both the database and the schema, so for example 'results.dbo'.

oracleTempSchema

For Oracle only: the name of the database schema where you want all temporary tables to be managed. Requires create/insert permissions to this database.

databaseName	String name of the database name. If blank, CDM_SOURCE table will be queried to try to obtain this.
smallCellCount	To avoid patient identifiability, cells with small counts (\leq smallCellCount) are deleted. Set to NULL if you don't want any deletions.
runVocabularyChecks	Boolean to determine if vocabulary checks need to be run. Default = TRUE
runPerformanceChecks	Boolean to determine if performance checks need to be run. Default = TRUE
runWebAPIChecks	Boolean to determine if WebAPI checks need to be run. Default = TRUE
baseUrl	WebAPI url, example: http://server.org:80/WebAPI
sqlOnly	Boolean to determine if Achilles should be fully executed. TRUE = just generate SQL files, don't actually run, FALSE = run Achilles
outputFolder	Path to store logs and SQL files
verboseMode	Boolean to determine if the console will show all execution steps. Default = TRUE
runSchemaChecks	Boolean to determine if CDM Schema Validation should be run. Default = TRUE

Details

CdmInspection runs a list of checks as part of the CDM inspection procedure

Value

An object of type `\code{achillesResults}` containing details for connecting to the

dataTablesChecks	<i>The vocabulary checks (for v5.x)</i>
------------------	---

Description

CdmInspection runs a list of checks on the vocabulary as part of the CDM inspection procedure

Usage

```
dataTablesChecks(
  connectionDetails,
  cdmDatabaseSchema,
  resultsDatabaseSchema = cdmDatabaseSchema,
  vocabDatabaseSchema = cdmDatabaseSchema,
  oracleTempSchema = resultsDatabaseSchema,
  sourceName = "",
  sqlOnly = FALSE,
  outputFolder = "output",
  verboseMode = TRUE
)
```

Arguments

connectionDetails	An R object of type connectionDetails created using the function createConnectionDetails in the DatabaseConnector package.
cdmDatabaseSchema	Fully qualified name of database schema that contains OMOP CDM schema. On SQL Server, this should specify both the database and the schema, so for example, on SQL Server, 'cdm_instance.dbo'.
resultsDatabaseSchema	Fully qualified name of database schema that we can write final results to. Default is cdmDatabaseSchema. On SQL Server, this should specify both the database and the schema, so for example, on SQL Server, 'cdm_results.dbo'.
vocabDatabaseSchema	String name of database schema that contains OMOP Vocabulary. Default is cdmDatabaseSchema. On SQL Server, this should specify both the database and the schema, so for example 'results.dbo'.
oracleTempSchema	For Oracle only: the name of the database schema where you want all temporary tables to be managed. Requires create/insert permissions to this database.
sourceName	String name of the data source name. If blank, CDM_SOURCE table will be queried to try to obtain this.
sqlOnly	Boolean to determine if Achilles should be fully executed. TRUE = just generate SQL files, don't actually run, FALSE = run Achilles
outputFolder	Path to store logs and SQL files
verboseMode	Boolean to determine if the console will show all execution steps. Default = TRUE

Details

CdmInspection runs a list of checks on the vocabulary as part of the CDM inspection procedure

Value

An object of type `\code{achillesResults}` containing details for connecting to the

performanceChecks	<i>The performane checks (for v5.x)</i>
-------------------	---

Description

PerformanceChecks runs a list of performance checks as part of the CDM inspection procedure

Usage

```
performanceChecks(
  connectionDetails,
  cdmDatabaseSchema,
  resultsDatabaseSchema = cdmDatabaseSchema,
  vocabDatabaseSchema = cdmDatabaseSchema,
```

```

    oracleTempSchema = resultsDatabaseSchema,
    sourceName = "",
    sqlOnly = FALSE,
    outputFolder = "output",
    verboseMode = TRUE
  )

```

Arguments

connectionDetails

An R object of type `connectionDetails` created using the function `createConnectionDetails` in the `DatabaseConnector` package.

cdmDatabaseSchema

Fully qualified name of database schema that contains OMOP CDM schema. On SQL Server, this should specify both the database and the schema, so for example, on SQL Server, `'cdm_instance.dbo'`.

resultsDatabaseSchema

Fully qualified name of database schema that we can write final results to. Default is `cdmDatabaseSchema`. On SQL Server, this should specify both the database and the schema, so for example, on SQL Server, `'cdm_results.dbo'`.

vocabDatabaseSchema

String name of database schema that contains OMOP Vocabulary. Default is `cdmDatabaseSchema`. On SQL Server, this should specify both the database and the schema, so for example `'results.dbo'`.

oracleTempSchema

For Oracle only: the name of the database schema where you want all temporary tables to be managed. Requires create/insert permissions to this database.

sourceName

String name of the data source name. If blank, `CDM_SOURCE` table will be queried to try to obtain this.

sqlOnly

Boolean to determine if Achilles should be fully executed. `TRUE` = just generate SQL files, don't actually run, `FALSE` = run Achilles

outputFolder

Path to store logs and SQL files

verboseMode

Boolean to determine if the console will show all execution steps. Default = `TRUE`

Details

`PerformanceChecks` runs a list of performance checks as part of the CDM inspection procedure

Value

An object of type `\code{achillesResults}` containing details for connecting to the

validateSchema	<i>Validate the CDM schema</i>
----------------	--------------------------------

Description

Validate the CDM schema

Usage

```
validateSchema(
  connectionDetails,
  cdmDatabaseSchema,
  resultsDatabaseSchema = cdmDatabaseSchema,
  cdmVersion,
  runCostAnalysis,
  outputFolder,
  sqlOnly = FALSE,
  verboseMode = TRUE
)
```

Arguments

connectionDetails	An R object of type connectionDetails created using the function createConnectionDetails in the DatabaseConnector package.
cdmDatabaseSchema	string name of database schema that contains OMOP CDM. On SQL Server, this should specify both the database and the schema, so for example 'cdm_instance.dbo'.
resultsDatabaseSchema	Fully qualified name of database schema that the cohort table is written to. Default is cdmDatabaseSchema. On SQL Server, this should specify both the database and the schema, so for example, on SQL Server, 'cdm_results.dbo'.
cdmVersion	Define the OMOP CDM version used: currently supports v5 and above. Use major release number or minor number only (e.g. 5, 5.3)
runCostAnalysis	Boolean to determine if cost analysis should be run. Note: only works on CDM v5 and v5.1.0+ style cost tables.
outputFolder	Path to store logs and SQL files
sqlOnly	TRUE = just generate SQL files, don't actually run, FALSE = run Achilles
verboseMode	Boolean to determine if the console will show all execution steps. Default = TRUE

Details

Runs a validation script to ensure the CDM is valid based on v5.x (NOT USED CURRENTLY)

vocabularyChecks	<i>The vocabulary checks (for v5.x)</i>
------------------	---

Description

CdmInspection runs a list of checks on the vocabulary as part of the CDM inspection procedure

Usage

```
vocabularyChecks(
  connectionDetails,
  cdmDatabaseSchema,
  resultsDatabaseSchema = cdmDatabaseSchema,
  vocabDatabaseSchema = cdmDatabaseSchema,
  oracleTempSchema = resultsDatabaseSchema,
  smallCellCount = 5,
  sourceName = "",
  sqlOnly = FALSE,
  outputFolder = "output",
  verboseMode = TRUE
)
```

Arguments

connectionDetails	An R object of type connectionDetails created using the function createConnectionDetails in the DatabaseConnector package.
cdmDatabaseSchema	Fully qualified name of database schema that contains OMOP CDM schema. On SQL Server, this should specify both the database and the schema, so for example, on SQL Server, 'cdm_instance.dbo'.
resultsDatabaseSchema	Fully qualified name of database schema that we can write final results to. Default is cdmDatabaseSchema. On SQL Server, this should specify both the database and the schema, so for example, on SQL Server, 'cdm_results.dbo'.
vocabDatabaseSchema	String name of database schema that contains OMOP Vocabulary. Default is cdmDatabaseSchema. On SQL Server, this should specify both the database and the schema, so for example 'results.dbo'.
oracleTempSchema	For Oracle only: the name of the database schema where you want all temporary tables to be managed. Requires create/insert permissions to this database.
sourceName	String name of the data source name. If blank, CDM_SOURCE table will be queried to try to obtain this.
sqlOnly	Boolean to determine if Achilles should be fully executed. TRUE = just generate SQL files, don't actually run, FALSE = run Achilles
outputFolder	Path to store logs and SQL files
verboseMode	Boolean to determine if the console will show all execution steps. Default = TRUE

Details

`CdmInspection` runs a list of checks on the vocabulary as part of the CDM inspection procedure

Value

An object of type `\code{achillesResults}` containing details for connecting to the

Index

cdmInspection, [2](#)
dataTablesChecks, [3](#)
performanceChecks, [4](#)
validateSchema, [6](#)
vocabularyChecks, [7](#)