

Package ‘ucsfindustrydocs’

April 29, 2025

Type Package

Title UCSF Industry Documents Library API Wrapper

Version 0.1.0

Description

Serves as a R wrapper for the University of California San Francisco's [Industry Documents Digital Library]<<https://www.industrydocuments.ucsf.edu/>> API. The API, and this wrapper, serve to pull metadata about of items within the digital library. For more information the API, see the [API's documentation]<https://www.industrydocuments.ucsf.edu/wp-content/uploads/2020/08/IndustryDocumentsDataAPI_v7.pdf>.

License MIT + file LICENSE

Imports arrow, data.table, httr, jsonlite, magrittr, dplyr, R6, stringr

Suggests mockery, testthat

Config/testthat/edition 3

Encoding UTF-8

RoxygenNote 7.3.2

NeedsCompilation no

Author Rolando Rodriguez [aut, cre]

Maintainer Rolando Rodriguez <rolando@ad.unc.edu>

Repository CRAN

Date/Publication 2025-04-29 09:10:02 UTC

Contents

IndustryDocsSearch	2
Index	5

IndustryDocsSearch	<i>UCSF Industry Documents Library Solr API</i>
--------------------	---

Description

UCSF Industry Documents Library Solr API

UCSF Industry Documents Library Solr API

Public fields

results placeholder for storing query results

Methods

Public methods:

- [IndustryDocsSearch\\$new\(\)](#)
- [IndustryDocsSearch\\$query\(\)](#)
- [IndustryDocsSearch\\$save\(\)](#)
- [IndustryDocsSearch\\$clone\(\)](#)

Method new(): Create a new IndustryDocsSearch instance

Usage:

```
IndustryDocsSearch$new()
```

Arguments:

NONE No parameters for initialization

Method query(): Query the UCSF Industry Documents Solr Library

Usage:

```
IndustryDocsSearch$query(  
  q = NULL,  
  case = NULL,  
  collection = NULL,  
  doc_type = NULL,  
  industry = NULL,  
  brand = NULL,  
  availability = NULL,  
  date = NULL,  
  id = NULL,  
  author = NULL,  
  source = NULL,  
  bates = NULL,  
  box = NULL,  
  originalformat = NULL,  
  wt = "json",  
  cursor_mark = "*",
```

```

    sort = "id%20asc",
    n = 1000
  )

```

Arguments:

q The query text that may incorporate the rest of the parameters. The function will not use the rest of the parameters if q is not NULL.

case The case the collection is related to.

collection The collection the results are found in.

doc_type The document type(s) to filter the results.

industry The industry the documents are located within.

brand The brand the documents are related to.

availability The availability status of the documents.

date The date of the documents.

id The id of the document(s).

author The author or originator of the contents of the document(s).

source The source of the document(s); usually the institution that deposited the documents.

bates The bates number(s) of the document(s) to be retrieved.

box The box id of the document(s) to be retrieved.

originalformat The original format of the document(s) to be retrieved.

wt The format the results should come in. Defaults to json. Functions depend on the results being returned as a JSON object.

cursor_mark Initial placeholder for cursormark within the API URL

sort The results will be sorted by ID in ascending order.

n The number of results we want to capture. Defaults to 1000. If n is set to -1 then all documents available related to the query will be retrieved.

Method save(): Save results to file

Usage:

```
IndustryDocsSearch$save(filename, format)
```

Arguments:

filename Output filename

format Output format ('parquet' or 'json' or 'csv')

Examples:

```

ids = IndustryDocsSearch$new()
ids$query(
  industry='tobacco',
  case='State of North Carolina',
  collection='JUUL labs Collection',
  n=100)
ids$save('query_results.csv', format='csv')
file.remove('query_results.csv')

```

Method clone(): The objects of this class are cloneable with this method.

Usage:

```
IndustryDocsSearch$clone(deep = FALSE)
```

Arguments:

deep Whether to make a deep clone.

Examples

```
## -----  
## Method `IndustryDocsSearch$save`  
## -----  
  
ids = IndustryDocsSearch$new()  
ids$query(  
  industry='tobacco',  
  case='State of North Carolina',  
  collection='JUUL labs Collection',  
  n=100)  
ids$save('query_results.csv', format='csv')  
file.remove('query_results.csv')
```

Index

IndustryDocsSearch, [2](#)