Package ‘TCIAPathfinder’

January 25, 2018

Title Client for the Cancer Imaging Archive REST API
Version 1.0.2
Description A wrapper for The Cancer Imaging Archive’s REST API. The Cancer Imaging Archive (TCIA) hosts de-identified medical images of cancer available for public download, as well as rich metadata for each image series. TCIA provides a REST API for programmatic access to the data. This package provides simple functions to access each API endpoint. For more information, see <https://github.com/pamelarussell/TCIAPathfinder> and TCIA’s website.
Depends R (>= 3.4.0)
License MIT + file LICENSE
Encoding UTF-8
LazyData true
Imports httr, jsonlite
RoxygenNote 6.0.1
Suggests knitr, rmarkdown, oro.dicom, testthat
VignetteBuilder knitr
NeedsCompilation no
Author Pamela Russell [aut, cre]
Maintainer Pamela Russell <pamela.russell@ucdenver.edu>
Repository CRAN
Date/Publication 2018-01-24 23:54:33 UTC

R topics documented:

get_body_part_names ................................................................. 2
get_collection_names ............................................................... 3
get_manufacturer_names ............................................................ 4
get_modality_names ................................................................. 5
get_new_patients_in_collection .................................................. 6
get_new_studies_in_collection .................................................. 7
get_body_part_names

get_patients_by_modality .............................................. 8
get_patient_info ......................................................... 9
get_patient_studies ....................................................... 10
get_series_info .......................................................... 11
get_series_size .......................................................... 12
get_sop_instance_uids ...................................................... 13
get_studies_in_collection ............................................... 14
save_image_series ......................................................... 15
save_single_image ........................................................ 16

Index 17

get_body_part_names  Get body part names

Description
Get body part names

Usage
get_body_part_names(collection = NULL, modality = NULL)

Arguments

collection  TCIA collection name. If collection is NULL, body part names from all collections will be returned. To get a list of available collection names, call get_collection_names.

modality  Modality name. If modality is NULL, body part names from all modalities will be returned. To get a list of available modality names, call get_modality_names or see DICOM Modality Abbreviations.

Value
List containing elements:

• body_parts: character vector of body part names
• content: parsed API response content
• response: API response

See Also
get_collection_names, get_modality_names, DICOM Modality Abbreviations, TCIA REST API Usage Guide, TCIA API object definitions
get_collection_names

Examples

## Not run:
get_body_part_names()
get_body_part_names(collection = "TCGA-BRCA")
get_body_part_names(modality = "MR")
get_body_part_names(collection = "TCGA-BRCA", modality = "MR")

## End(Not run)

---

get_collection_names  Get the names of all TCIA collections

Description

Get the names of all TCIA collections

Usage

get_collection_names()

Value

List containing elements:

- collection_names: character vector of TCIA collection names
- content: parsed API response content
- response: API response

See Also

TCIA REST API Usage Guide, TCIA API object definitions

Examples

## Not run:
get_collection_names()

## End(Not run)
get_manufacturer_names

Get manufacturer names

Description
Get manufacturer names

Usage
get_manufacturer_names(collection = NULL, modality = NULL, 
body_part = NULL)

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>collection</td>
<td>TCIA collection name. If collection is NULL, manufacturer names from all</td>
</tr>
<tr>
<td></td>
<td>collections will be returned. To get a list of available collection names, call</td>
</tr>
<tr>
<td></td>
<td>get_collection_names.</td>
</tr>
<tr>
<td>modality</td>
<td>Modality name. If modality is NULL, manufacturer names for all modalities will</td>
</tr>
<tr>
<td></td>
<td>be returned. To get a list of available modality names, call get_modality_names</td>
</tr>
<tr>
<td></td>
<td>or see DICOM Modality Abbreviations.</td>
</tr>
<tr>
<td>body_part</td>
<td>Body part name. If body_part is NULL, manufacturer names for all body parts</td>
</tr>
<tr>
<td></td>
<td>will be returned. To get a list of available body part names, call get_body_part_names.</td>
</tr>
</tbody>
</table>

Value
List containing elements:

- manufacturer_names: character vector of manufacturer names
- content: parsed API response content
- response: API response

See Also
- get_collection_names, get_modality_names, get_body_part_names, DICOM Modality Abbreviations, TCIA REST API Usage Guide, TCIA API object definitions

Examples
```r
## Not run:
geet_manufacturer_names()
geet_manufacturer_names(collection = "TCGA-BRCA")
geet_manufacturer_names(collection = "TCGA-BRCA", modality = "MR", body_part = "BREAST")
## End(Not run)
```
Description

Get modality names

Usage

```r
get_modality_names(collection = NULL, body_part = NULL)
```

Arguments

- `collection` TCIA collection name. If `collection` is `NULL`, modality names from all collections will be returned. To get a list of available collection names, call `get_collection_names`.
- `body_part` Body part name. If `body_part` is `NULL`, modality names for all body parts will be returned. To get a list of available body part names, call `get_body_part_names`.

Value

List containing elements:

- modality_names: character vector of modality names
- content: parsed API response content
- response: API response

See Also

`get_collection_names`, `get_body_part_names`, TCIA REST API Usage Guide, TCIA API object definitions

Examples

```r
## Not run:
get_modality_names()
get_modality_names(collection = "TCGA-BRCA")
get_modality_names(body_part = "BREAST")
get_modality_names(collection = "TCGA-BRCA", body_part = "BREAST")

## End(Not run)
```
get_new_patients_in_collection

*Get IDs of patients that have been added to a collection since a specified date*

**Description**

Get IDs of patients that have been added to a collection since a specified date

**Usage**

get_new_patients_in_collection(collection, date)

**Arguments**

- **collection**: TCIA collection name. To get a list of available collection names, call get_collection_names.
- **date**: Date in format "YYYY-MM-DD"

**Value**

List containing elements:

- **patient_ids**: Patient IDs
- **content**: parsed API response content
- **response**: API response

**See Also**

get_collection_names, TCIA REST API Usage Guide, TCIA API object definitions

**Examples**

```
## Not run:
get_new_patients_in_collection("TCGA-BRCA", "2014-01-01")

## End(Not run)
```
get_new_studies_in_collection

Get studies that have been added to a collection and optionally to a patient since a specified date

Description

Get studies that have been added to a collection and optionally to a patient since a specified date.

Usage

get_new_studies_in_collection(collection, date, patient_id = NULL)

Arguments

collection  TCIA collection name. To get a list of available collection names, call get_collection_names.
date        Date in format "YYYY-MM-DD"
patient_id  Patient ID. To get a list of available patient IDs, call get_patient_info. If patient_id is NULL, relevant studies for all patients in the collection will be returned.

Value

List containing elements:

- studies: Data frame of collection, patient ID, and study instance UID
- content: parsed API response content
- response: API response

See Also

get_collection_names, get_patient_info, TCIA REST API Usage Guide, TCIA API object definitions

Examples

```r
## Not run:
geet_new_studies_in_collection("TCGA-BRCA", "2014-01-01")
geet_new_studies_in_collection("TCGA-BRCA", "2014-01-01", "TCGA-0L-A660")
```

## End(Not run)
get_patients_by_modality

Get patient IDs given a collection name and modality

Description

Get patient IDs given a collection name and modality

Usage

get_patients_by_modality(collection, modality)

Arguments

collection: TCIA collection name. To get a list of available collection names, call get_collection_names.
modality: Modality name. To get a list of available modality names, call get_modality_names or see DICOM Modality Abbreviations.

Value

List containing elements:

- patient_ids: Patient IDs
- content: parsed API response content
- response: API response

See Also

get_collection_names, get_modality_names, DICOM Modality Abbreviations, TCIA REST API Usage Guide, TCIA API object definitions

Examples

## Not run:
get_patients_by_modality("TCGA-BRCA", "MR")

## End(Not run)
get_patient_info

Description
Get patient information

Usage
get_patient_info(collection = NULL)

Arguments
collection TCIA collection name. If collection is NULL, patients from all collections will be returned. To get a list of available collection names, call get_collection_names.

Value
List containing elements:

- patients: Data frame of patient ID, name, sex, ethnic group, and collection name
- content: parsed API response content
- response: API response

See Also
general_collection_names, TCIA REST API Usage Guide, TCIA API object definitions

Examples
## Not run:
get_patient_info()
get_patient_info("TCGA-BRCA")

## End(Not run)
**get_patient_studies**

Get patient study information

**Description**

Get patient study information

**Usage**

```r
get_patient_studies(collection = NULL, patient_id = NULL, study_instance_uid = NULL)
```

**Arguments**

- **collection**: TCIA collection name. To get a list of available collection names, call `get_collection_names`. If `collection` is `NULL`, information for all relevant collections will be returned.
- **patient_id**: Patient ID. To get a list of available patient IDs, call `get_patient_info`. If `patient_id` is `NULL`, information for all relevant patients will be returned.
- **study_instance_uid**: Study instance UID. If `study_instance_uid` is `NULL`, information for all relevant study instance UIDs will be returned. To get available study instance UIDs, call `get_studies_in_collection, get_patient_studies, or get_new_studies_in_collection`.

**Value**

List containing elements:

- **patient_studies**: Data frame with columns representing the contents of a PatientStudy object as described in TCIA API Return Values
- **content**: parsed API response content
- **response**: API response

**See Also**

`get_collection_names, get_patient_info, get_studies_in_collection, get_patient_studies, get_new_studies_in_collection`, TCIA REST API Usage Guide, TCIA API object definitions

**Examples**

```r
## Not run:
get_patient_studies()
get_patient_studies(collection = "TCGA-BRCA")
get_patient_studies(patient_id = "TCGA-OL-A6VO")
get_patient_studies(patient_id = "TCGA-OL-A5OA",
                   study_instance_uid = "1.3.6.1.4.1.14519.5.2.1.5382.4002.104582989590517557856962159716")

## End(Not run)
```
**get_series_info**

- **Get image series information**

**Description**

Get image series information

**Usage**

```python
get_series_info(collection = NULL, patient_id = NULL,
    study_instance_uid = NULL, series_instance_uid = NULL, modality = NULL,
    body_part_examined = NULL, manufacturer_model_name = NULL,
    manufacturer = NULL)
```

**Arguments**

- **collection**  
  TCIA collection name. To get a list of available collection names, call `get_collection_names`. If collection is NULL, information for all relevant collections will be returned.

- **patient_id**  
  Patient ID. To get a list of available patient IDs, call `get_patient_info`. If patient_id is NULL, information for all relevant patients will be returned.

- **study_instance_uid**  
  Study instance UID. If study_instance_uid is NULL, information for all relevant study instance UIDs will be returned. To get available study instance UIDs, call `get_studies_in_collection`, `get_patient_studies`, or `get_new_studies_in_collection`.

- **series_instance_uid**  
  Series instance UID. To get a list of available series instance UIDs, call this function leaving out parameter series_instance_uid. If series_instance_uid is NULL, information for all relevant series will be returned.

- **modality**  
  Modality name. To get a list of available modality names, call `get_modality_names` or see DICOM Modality Abbreviations. If modality is NULL, information for all relevant modalities will be returned.

- **body_part_examined**  
  Body part name. To get a list of available body part names, call `get_body_part_names`. If body_part_examined is NULL, information for all relevant body parts will be returned. IMPORTANT: a bug in this query key has been observed in the TCIA API. If queries using this key return zero results, try removing this parameter.

- **manufacturer_model_name**  
  Manufacturer model name. To get a list of available model names, call this function leaving out parameter manufacturer_model_name. If manufacturer_model_name is NULL, information for all relevant model names will be returned.

- **manufacturer**  
  Manufacturer name. To get a list of available manufacturer names, call `get_manufacturer_names`. If manufacturer is NULL, information for all relevant manufacturers will be returned.
get_series_size

Value

List containing elements:

- series: Data frame with columns representing the contents of a Series object as described in TCIA API Return Values
- content: parsed API response content
- response: API response

See Also

get_collection_names, get_patient_info, get_studies_in_collection, get_patient_studies,
get_new_studies_in_collection, get_modality_names, get_body_part_names, get_manufacturer_names,
TCIA REST API Usage Guide, TCIA API object definitions

Examples

```r
## Not run:
get_series_info()
get_series_info(collection = "TCGA-BRCA")
get_series_info(patient_id = "TCGA-OL-A6V0")
get_series_info(modality = "MR", manufacturer = "GE MEDICAL SYSTEMS")
## End(Not run)
```

get_series_size  Get size of image series

Description

Get size of image series

Usage

get_series_size(series_instance_uid)

Arguments

- `series_instance_uid`  
  Series instance UID. To get a list of available series instance UIDs, call `get_series_info`.

Value

List containing elements:

- `size_bytes`: Total size of image series in bytes
- `object_count`: Number of objects in image series
- content: parsed API response content
- response: API response
get_sop_instance_uids

See Also

get_series_info, TCIA REST API Usage Guide, TCIA API object definitions

Examples

```r
## Not run:
get_series_size("1.3.6.1.4.1.14519.5.2.1.5382.4002.272234209223992578700978260744")

## End(Not run)
```

---

get_sop_instance_uids Get SOP instance UIDs (individual DICOM image IDs) for an image series

**Description**

Get SOP instance UIDs (individual DICOM image IDs) for an image series

**Usage**

```r
get_sop_instance_uids(series_instance_uid)
```

**Arguments**

- `series_instance_uid`
  - Series instance UID. To get a list of available series instance UIDs, call `get_series_info`.

**Value**

List containing elements:

- `sop_instance_uids`: character vector of SOP instance UIDs (individual DICOM image IDs)
- `content`: parsed API response content
- `response`: API response

See Also

get_series_info, TCIA REST API Usage Guide, TCIA API object definitions

Examples

```r
## Not run:
get_sop_instance_uids("1.3.6.1.4.1.14519.5.2.1.5382.4002.272234209223992578700978260744")

## End(Not run)
```
get_studies_in_collection

Get studies in a collection and optionally for a specific patient

Description

Get studies in a collection and optionally for a specific patient

Usage

get_studies_in_collection(collection, patient_id = NULL)

Arguments

collection  TCIA collection name. To get a list of available collection names, call get_collection_names.
patient_id  Patient ID. To get a list of available patient IDs, call get_patient_info. If patient_id is NULL, studies for all patients in the collection will be returned.

Value

List containing elements:

- studies: Data frame of collection, patient ID, and study instance UID
- content: parsed API response content
- response: API response

See Also

get_collection_names, get_patient_info, TCIA REST API Usage Guide, TCIA API object definitions

Examples

```r
## Not run:
get_studies_in_collection("TCGA-BRCA")
get_studies_in_collection("TCGA-BRCA", "TCGA-OL-A660")
## End(Not run)
```
save_image_series  

**Save a series of DICOM image files as a zip file**

**Description**

Save a series of DICOM image files as a zip file.

**Usage**

```bash
save_image_series(series_instance_uid, out_dir, out_file_name = NULL)
```

**Arguments**

- `series_instance_uid`: Series instance UID. To get a list of available series instance UIDs, call `get_series_info`. Note: if `series_instance_uid` is invalid, the API may still successfully return an empty zip file.
- `out_dir`: Directory to write zip file to.
- `out_file_name`: Name of zip file to write. If `out_file_name` is `null`, the original file name will be used.

**Value**

List containing elements:

- `out_file`: The output zip file that was written
- `response`: API response

**See Also**

`get_series_info`, TCIA REST API Usage Guide, TCIA API object definitions

**Examples**

```bash
## Not run:
save_image_series(
    series_instance_uid = "1.3.6.1.4.1.14519.5.2.1.5382.4002.806935685832642465081499816867",
    out_dir = "~/Desktop")
save_image_series(
    series_instance_uid = "1.3.6.1.4.1.14519.5.2.1.5382.4002.806935685832642465081499816867",
    out_dir = "~/Desktop", out_file_name = "file.zip")

## End(Not run)
```
save_single_image  Save a single DICOM image file

Description
Save a single DICOM image file

Usage
save_single_image(series_instance_uid, sop_instance_uid, out_dir,
             out_file_name = NULL)

Arguments
series_instance_uid
  Series instance UID. To get a list of available series instance UIDs, call get_series_info.
sop_instance_uid
  SOP instance UID. To get a list of SOP instance UIDs for an image series, call get_sop_instance_uids.
out_dir
  Directory to write DICOM file to
out_file_name
  Name of DICOM file to write, with .dcm extension. If out_file_name is NULL, the original file name will be used.

Value
List containing elements:
  • out_file: The output file that was written
  • response: API response

See Also
  get_series_info, get_sop_instance_uids, TCIA REST API Usage Guide, TCIA API object definitions

Examples
## Not run:
save_single_image(
  series_instance_uid = "1.3.6.1.4.1.14519.5.2.1.5382.4002.806935685832642465081499816867",
  sop_instance_uid = "1.3.6.1.4.1.14519.5.2.1.5382.4002.257663256941568276393774062283",
  out_dir = "/Desktop"
)

save_single_image(
  series_instance_uid = "1.3.6.1.4.1.14519.5.2.1.5382.4002.806935685832642465081499816867",
  sop_instance_uid = "1.3.6.1.4.1.14519.5.2.1.5382.4002.257663256941568276393774062283",
  out_dir = "/Desktop", out_file_name = "file.dcm"
)

## End(Not run)
Index

get_body_part_names, 2, 4, 5, 11, 12
get_collection_names, 2, 3, 4–12, 14
get_manufacturer_names, 4, 11, 12
get_modality_names, 2, 4, 5, 8, 11, 12
get_new_patients_in_collection, 6
get_new_studies_in_collection, 7, 10–12
get_patient_info, 7, 9, 10–12, 14
get_patient_studies, 10, 10, 11, 12
get_patients_by_modality, 8
get_series_info, 11, 12, 13, 15, 16
get_series_size, 12
get_sop_instance_uids, 13, 16
get_studies_in_collection, 10–12, 14

save_image_series, 15
save_single_image, 16