Package ‘HuBMAPR’

April 9, 2025

Title Interface to 'HuBMAP'

Version 1.1.8

Description 'HuBMAP' provides an open, global bio-molecular atlas of the human body at the cellular level. The `datasets()`, `samples()`, `donors()`, `publications()`, and `collections()` functions retrieves the information for each of these entity types. `*_details()` are available for individual entries of each entity type. `*_derived()` are available for retrieving derived datasets or samples for individual entries of each entity type. Data files can be accessed using `bulk_data_transfer()`.

License Artistic-2.0

Depends R (>= 4.4.0)

Imports httr2, dplyr, tidyr, tibble, rjsoncons, rlang, utils, stringr, whisker, purrr

Collate query.R datasets.R samples.R donors.R utilities.R publications.R publication_information.R collections.R collection_information.R provenance.R files.R organ.R template.R

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.2

Suggests testthat (>= 3.0.0), knitr, ggplot2, rmarkdown, BiocStyle, pryr

VignetteBuilder knitr

biocViews Software, SingleCell, DataImport, ThirdPartyClient, Spatial, Infrastructure

URL https://christinehou11.github.io/HuBMAPR/, https://github.com/christinehou11/HuBMAPR

BugReports https://github.com/christinehou11/HuBMAPR/issues

Config/testthat/edition 3

git_url https://git.bioconductor.org/packages/HuBMAPR

git_branch devel

git_last_commit 9f529c2

git_last_commit_date 2025-02-12

Repository Bioconductor 3.21

Date/Publication 2025-04-09

Author Christine Hou [aut, cre] (ORCID:
<https://orcid.org/0009-0001-5350-0629>),
Martin Morgan [aut] (ORCID: <https://orcid.org/0000-0002-5874-8148>),
Federico Marini [aut] (ORCID: <https://orcid.org/0000-0003-3252-7758>)

Maintainer Christine Hou <chris2018hou@gmail.com>

Contents

`bulk_data_transfer` *HuBMAP File Querying*

Description

`bulk_data_transfer()` takes a unique dataset or publication UUID to be used to query the HuBMAP API for information about available files.

Usage

```
bulk_data_transfer(uuid)
```

Arguments

`uuid`	character(1) corresponding to the HuBMAP Dataset UUID string. This is expected to be a 32-digit hex number. One unique Dataset UUID corresponds to one specific URL on Globus. Files corresponding to unique Dataset UUID are saved on Globus HuBMAP Public Collection webpage.

Details

Additional details are provided on the HuBMAP consortium webpage, https://software.docs.hubmapconsortium.org/apis

Value

`bulk_data_transfer()` display information about how to bulk download all files related to the dataset or publication.

Examples

```
# Globus available
uuid_globus <- "2d3dbd5e5a3d9d5cfc7a46f06815a4eb"
uuid_sra_dbGaP <- "5ba6a5b81e95c93c26a33980f6e957d7"
uuid_not_available <- "0eb5e457b4855ce28531bc97147196b6"

# bulk_data_transfer(*_globus)
```

`collections` *HuBMAP Collections*

Description

`collections()` returns details about available collections, ordered by last modified dates
`*_columns()` returns a tibble or named character vector describing the content of the tibble returned by `samples()`, `datasets()`, `donors()`, `collections()`, or `publications()`.

`collection_contacts()` takes a unique collection_id and returns contacts information of one specified collection as a tibble

`collection_data()` takes a unique collection_id and returns related datasets of one specified collection as a tibble

`collection_contributors()` takes a unique collection_id and returns contributors information of one specified collection as a tibble

Usage

```
collections()

collections_default_columns(as = c("tibble", "character"))

collection_contacts(uuid)

collection_data(uuid)

collection_contributors(uuid)
```

Arguments

`as`	character(1) return format. One of `"tibble"` (default), or `"character"`.
`uuid`	character(1) corresponding to the HuBMAP Collection UUID string. This is expected to be a 32-digit hex number.

Details

Additional details are provided on the HuBMAP consortium webpage, https://software.docs.hubmapconsortium.org/apis

Additional details are provided on the HuBMAP consortium webpage, https://software.docs.hubmapconsortium.org/apis

Additional details are provided on the HuBMAP consortium webpage, https://software.docs.hubmapconsortium.org/apis

Additional details are provided on the HuBMAP consortium webpage, https://software.docs.hubmapconsortium.org/apis

Value

`*_columns()` returns a named list `name` containing the column name used in the tibble returned by `samples()`, `datasets()`, `donors()`, `collections()`, or `publications()`. When `as = "tibble"`,the return value is a tibble with paths as elements and abbreviations as names.

Examples

```
collections()
collections_default_columns()

uuid <- "381f65e58d5e2c1d16a9cef2cc203aab"
collection_contacts(uuid)
uuid <- "381f65e58d5e2c1d16a9cef2cc203aab"
collection_data(uuid)
uuid <- "590b0485a196956284b8f3344276bc50"
collection_contributors(uuid)
```

`collection_information` *Collection Information Summaries from Collection UUID*

Description

`collection_information()` queries the HuBMAP collection for title, description, DOI, URL, creation date, and last modified time date.

`print.collection_information()` organizes the returned tibble from `collection_information()` in a more legible format.

Usage

```
collection_information(uuid)

## S3 method for class 'collection_information'
print(x, ...)
```

Arguments

`uuid`	character(1) corresponding to the HuBMAP Collection UUID string. This is expected to be a 32-digit hex number.
`x`	an object of class `collection_information`, the result of a call to `collection_information()`.
`...`	additional arguments, required to conform with the `print` generic but not used.

Value

`collection_information()` returns a tibble, and columns containing information about the collection. The tibble is of class `collection_information` and is printed in character.

`print.collection_information()` is invoked automatically when the result of `collection_information()` is displayed for it's side effect of displaying the object.

Examples

```
uuid <- "90399057f6ff0ff18d4c46a378f0b069"
collection_information(uuid)
```

`datasets` *HuBMAP Datasets*

Description

`datasets` returns the details available datasets, ordered by last modified dates
`*_columns()` returns a tibble or named character vector describing the content of the tibble returned by `samples()`, `datasets()`, `donors()`, `collections()`, or `publications()`.

`dataset_detail()` takes a unique dataset_id and returns details about one specified dataset as a tibble

`dataset_derived()` takes a unique sample_id and returns the derived (support) dataset details. Support datasets normally belong to Image Pyramid, with image files available to download via Globus Collection. See details to download in `files_globus_url()`.

`dataset_metadata()` takes a unique dataset_id and returns the metadata of the dataset.

`dataset_contributors()` takes a unique dataset_id and returns the contributors of the dataset. For questions for this dataset, reach out to the individuals listed as contacts, either via the email address listed in the table or contact information provided on their ORCID profile page.

Usage

```
datasets()

datasets_default_columns(as = c("tibble", "character"))

dataset_detail(uuid)
```

```
dataset_derived(uuid)

dataset_metadata(uuid)

dataset_contributors(uuid)
```

Arguments

`as`	character(1) return format. One of `"tibble"` (default), or `"character"`.
`uuid`	character(1) corresponding to the HuBMAP Donor UUID string. This is expected to be a 32-digit hex number.

Details

Additional details are provided on the HuBMAP consortium webpage, https://software.docs.hubmapconsortium.org/apis

Additional details are provided on the HuBMAP consortium webpage, https://software.docs.hubmapconsortium.org/apis

Additional details are provided on the HuBMAP consortium webpage, https://software.docs.hubmapconsortium.org/apis

Additional details are provided on the HuBMAP consortium webpage, https://software.docs.hubmapconsortium.org/apis

Value

`*_columns()` returns a named list `name` containing the column name used in the tibble returned by `samples()`, `datasets()`, `donors()`, `collections()`, or `publications()`. When `as = "tibble"`,the return value is a tibble with paths as elements and abbreviations as names.

Examples

```
datasets()
datasets_default_columns()

uuid <- "7754aa5ebde628b5e92705e33e74a4ef"
dataset_detail(uuid)
# no derived dataset
uuid <- "3acdb3ed962b2087fbe325514b098101"
dataset_derived(uuid)

# with derived dataset
uuid <- "2c77b1cdf33dbed3dbfb74e4b578300e"
dataset_derived(uuid)
uuid <- "564167adbbb2fdd64c24e7ea409c23f1"
dataset_metadata(uuid)

uuid <- "564167adbbb2fdd64c24e7ea409c23f1"
dataset_contributors(uuid)
```

`donors` *HuBMAP Donors*

Description

`donors()` returns details about available samples, ordered by last modified dates
`*_columns()` returns a tibble or named character vector describing the content of the tibble returned by `samples()`, `datasets()`, `donors()`, `collections()`, or `publications()`.

`donor_detail()` takes a unique donor_id and returns details about one specified sample as a tibble

`donor_derived()` takes a unique donor_id and returns the derived dataset or/and sample details.

`donor_metadata()` takes a unique donor_id and returns the metadata of the donor.

Usage

```
donors()

donors_default_columns(as = c("tibble", "character"))

donor_detail(uuid)

donor_derived(uuid, entity_type = c("Dataset", "Sample"))

donor_metadata(uuid)
```

Arguments

`as`	character(1) return format. One of `"tibble"` (default), or `"character"`.
`uuid`	character(1) corresponding to the HuBMAP Donor UUID string. This is expected to be a 32-digit hex number.
`entity_type`	character(1) selected derived entity type. One of `"Sample"` or `"Dataset"` (default).

Details

Additional details are provided on the HuBMAP consortium webpage, https://software.docs.hubmapconsortium.org/apis

Additional details are provided on the HuBMAP consortium webpage, https://software.docs.hubmapconsortium.org/apis

Additional details are provided on the HuBMAP consortium webpage, https://software.docs.hubmapconsortium.org/apis

Additional details are provided on the HuBMAP consortium webpage, https://software.docs.hubmapconsortium.org/apis

Value

`*_columns()` returns a named list `name` containing the column name used in the tibble returned by `samples()`, `datasets()`, `donors()`, `collections()`, or `publications()`. When `as = "tibble"`,the return value is a tibble with paths as elements and abbreviations as names.

Examples

```
donors()
donors_default_columns()

uuid <- "1dcde05aea3509b2cf89a41ceb3d700f"
donor_detail(uuid)
uuid <- "1dcde05aea3509b2cf89a41ceb3d700f"
donor_derived(uuid, "Sample")

uuid <- "d37df2cad4e80dc368763caefccf7140"
donor_metadata(uuid)
```

`organ` *HuBMAP Organs*

Description

`organ` retrieves available organ categories and abbreviations

Usage

```
organ()
```

Value

`organ` returns the tibble with all available organs' abbreviations and corresponding full name.

Examples

```
organs <- organ()
organs |>
    dplyr::select("name") |>
    dplyr::distinct()
```

`publications` *HuBMAP Samples*

Description

`publications()` returns details about available samples, ordered by last modified dates. `publication_status` == `TRUE`: peer reviewed publication; `publication_status` == `FALSE`: pre-print publication.
`*_columns()` returns a tibble or named character vector describing the content of the tibble returned by `samples()`, `datasets()`, `donors()`, `collections()`, or `publications()`.

`publication_data()` takes a unique publication_id and returns details about one specified publication.

`publication_authors()` takes a unique publication_id and returns authors information of one specified collection as a tibble

Usage

```
publications()

publications_default_columns(as = c("tibble", "character"))

publication_data(uuid, entity_type = c("Dataset", "Sample", "Donor"))

publication_authors(uuid)
```

Arguments

`as`	character(1) return format. One of `"tibble"` (default), or `"character"`.
`uuid`	character(1) corresponding to the HuBMAP Publication UUID string. This is expected to be a 32-digit hex number.
`entity_type`	character(1) selected derived entity type. One of `"Sample"`, `"Donor"` or `"Dataset"` (default).

Details

Additional details are provided on the HuBMAP consortium webpage, https://software.docs.hubmapconsortium.org/apis

Additional details are provided on the HuBMAP consortium webpage, https://software.docs.hubmapconsortium.org/apis

Additional details are provided on the HuBMAP consortium webpage, https://software.docs.hubmapconsortium.org/apis

Value

`*_columns()` returns a named list `name` containing the column name used in the tibble returned by `samples()`, `datasets()`, `donors()`, `collections()`, or `publications()`. When `as = "tibble"`,the return value is a tibble with paths as elements and abbreviations as names.

Examples

```
publications()
publications_default_columns()

uuid <- "3c7273660cdf9ab91a7901533b2cd9a5"
publication_data(uuid)

uuid <- "3c7273660cdf9ab91a7901533b2cd9a5"
publication_authors(uuid)
```

`publication_information` *Publication Information Summaries from Publication UUID*

Description

`publication_information()` queries the HuBMAP publication for title, abstract, manuscript,corresponding authors, dataset types, and organ.

`publication_information()` organizes the returned tibble from `publication_information()` in a more legible format.

Usage

```
publication_information(uuid)

## S3 method for class 'publication_information'
print(x, ...)
```

Arguments

`uuid`	character(1) corresponding to the HuBMAP Publication UUID string. This is expected to be a 32-digit hex number.
`x`	an object of class `publication_information`, the result of a call to `publication_information()`.
`...`	additional arguments, required to conform with the `print` generic but not used.

Value

`publication_information()` returns a tibble, and columns containing information about the publication. The tibble is of class `publication_information` and is printed in character.

`print.publication_information()` is invoked automatically when the result of `publication_information()` is displayed for it's side effect of displaying the object.

Examples

```
uuid <- "e8338966c69e759157d6c6ad24847989"
publication_information(uuid)
```

`samples` *HuBMAP Samples*

Description

`samples()` returns details about available samples, ordered by last modified dates. There are multiple tissue sample types displayed in `sample_category` column which are block, organ, suspension, and section.
`*_columns()` returns a tibble or named character vector describing the content of the tibble returned by `samples()`, `datasets()`, `donors()`, `collections()`, or `publications()`.

`sample_detail()` takes a unique sample_id and returns details about one specified sample as a tibble

`sample_derived()` takes a unique sample_id and returns the derived dataset or/and sample details.

`sample_metadata()` takes a unique donor_id and returns the metadata of the sample.

Usage

```
samples()

samples_default_columns(as = c("tibble", "character"))

sample_detail(uuid)

sample_derived(uuid, entity_type = c("Dataset", "Sample"))

sample_metadata(uuid)
```

Arguments

`as`	character(1) return format. One of `"tibble"` (default), or `"character"`.
`uuid`	character(1) corresponding to the HuBMAP Donor UUID string. This is expected to be a 32-digit hex number.
`entity_type`	character(1) selected derived entity type. One of `"Sample"` or `"Dataset"` (default).

Details

Additional details are provided on the HuBMAP consortium webpage, https://software.docs.hubmapconsortium.org/apis

Additional details are provided on the HuBMAP consortium webpage, https://software.docs.hubmapconsortium.org/apis

Additional details are provided on the HuBMAP consortium webpage, https://software.docs.hubmapconsortium.org/apis

Additional details are provided on the HuBMAP consortium webpage, https://software.docs.hubmapconsortium.org/apis

Value

`*_columns()` returns a named list `name` containing the column name used in the tibble returned by `samples()`, `datasets()`, `donors()`, `collections()`, or `publications()`. When `as = "tibble"`,the return value is a tibble with paths as elements and abbreviations as names.

Examples

```
samples()
samples_default_columns()

uuid <- "d3525d35f6d5ee3dc3186613b0ab1762"
sample_detail(uuid)
uuid <- "3e7dc14313262af577f686dcb09f5119"
sample_derived(uuid, "Sample")
uuid <- "8d7ce3aab7013e416263d23b0c048900"
sample_metadata(uuid)
```

`uuid_provenance` *HuBMAP Provenance*

Description

HuBMAP Provenance

Usage

```
uuid_provenance(uuid)
```

Arguments

`uuid`	character(1) corresponding to the HuBMAP Dataset or Sample UUID string. This is expected to be a 32-digit hex number.

Details

Additional details are provided on the HuBMAP consortium webpage, https://software.docs.hubmapconsortium.org/apis

Value

`uuid_provenance()` takes a Dataset or Sample UUID, and returns the provenance details as a list of characters, from the most recent ancestor to the furthest ancestor. The entity type of ancestors include Dataset, Sample, and Donor. Note that there is no ancestor for Donor entity such that the returned result will be an empty list.

Examples

```
uuid <- "acf816bed56dd6ba55da82960d937ec4"
uuid_provenance(uuid)
```

Index