

Package: odyssey (via r-universe)

July 1, 2024

Title Access documents from the Open Archive HAL

Version 0.0.0.9000

Description Access references and download documents from the Open Archive HAL <<https://hal.archives-ouvertes.fr/>> using the search API.

License GPL (>= 3)

URL <https://github.com/nfrerebeau/odyssey>,
<https://nfrerebeau.github.io/odyssey>

BugReports <https://github.com/nfrerebeau/odyssey/issues>

Encoding UTF-8

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.1.2

Imports htr, solrium, utils

Suggests magrittr, knitr, rmarkdown, testthat (>= 3.0.0)

Collate 'AllGenerics.R' 'api.R' 'download.R' 'facet.R' 'group.R'
'odyssey-package.R' 'query.R' 'search.R' 'zzz.R'

VignetteBuilder knitr

Config/testthat/edition 3

Repository <https://nfrerebeau.r-universe.dev>

RemoteUrl <https://github.com/nfrerebeau/odyssey>

RemoteRef main

RemoteSha 84e7d630e7ea68656ed37a98a5b7ad315e816c6f

Contents

hal_api	2
hal_download	3
hal_facet	4

hal_filter	5
hal_group	7
hal_query	8
hal_select	9
hal_sort	10
search	11

Index	14
--------------	-----------

hal_api	<i>HAL API</i>
---------	----------------

Description

Initialize a request to the HAL API.

Usage

```
hal_api()
```

Value

An object of class HALQuery: a [list](#) containing at least the following components:

q (default to "::*").

f1 default to "docid" "label_s").

fq

sort

rows (default to 30).

start (default to 0).

wt (default to "json").

Examples

```
hal_api()
```

hal_download	<i>Download Documents</i>
--------------	---------------------------

Description

Download Documents

Usage

```
hal_download(x, ...)  
  
## S3 method for class 'HALQuery'  
hal_download(  
  x,  
  limit = 30,  
  start = 0,  
  progress = getOption("odyssey.progress"),  
  verbose = getOption("odyssey.verbose"),  
  ...  
)
```

Arguments

x	An object of class HALQuery (typically returned by hal_api()).
...	Currently not used.
limit	An integer giving the maximum number of results. According to HAL policy, it cannot exceed 10000.
start	An integer specifying an absolute offset in the complete sorted list of matches to be used as the beginning of the current page. Only used if <code>cursor</code> is FALSE.
progress	A logical scalar: should a progress bar for for the request be printed? Only used if R is being used interactively.
verbose	A logical scalar: should extra information be reported?

Value

An (invisible) [logical](#) scalar, TRUE for success and FALSE if any failure.

Author(s)

N. Frerebeau

See Also

Other search tools: [search\(\)](#)

Examples

```
## Not run:
library(magrittr)

## Download the 10 most recent archaeological publication
## (if any files)
hal_api() %>%
  hal_query("archéologie") %>%
  hal_filter("ART" %IN% "docType_s") %>%
  hal_sort("producedDate_tdate", decreasing = TRUE) %>%
  hal_download(limit = 10)

## End(Not run)
```

hal_facet

Facet Search

Description

Facet Search

Usage

```
hal_facet(x, ...)

## S3 method for class 'HALQuery'
hal_facet(
  x,
  field,
  limit = 5,
  sort = c("index", "count"),
  prefix = NULL,
  contains = NULL,
  pivot = NULL,
  range = NULL,
  ignore_case = FALSE,
  ...
)
```

Arguments

x	An object of class HALQuery (typically returned by hal_api()).
...	Currently not used.
field	A character string specifying the field to group by.
limit	An integer giving the maximum number of results per group.
sort	A character string specifying the field to be used to sort the results.

prefix	A character string.
contains	A character string.
pivot	A character string.
range	A list containing the following components: "range", "start", "end", "gap".
ignore_case	A logical scalar: should character case be ignored?

Value

An object of class [HALQuery](#).

Author(s)

N. Frerebeau

See Also

Other query tools: [hal_filter\(\)](#), [hal_group\(\)](#), [hal_query\(\)](#), [hal_select\(\)](#), [hal_sort\(\)](#)

Examples

```
## Not run:
library(magrittr)

## Most recent publication by journal
hal_api() %>%
  hal_query("archéologie") %>%
  hal_facet(
    field = "docType_s",
    limit = 100,
    sort = "count",
    range = list(
      range = "producedDateY_i",
      start = 2000,
      end = 2020,
      gap = 1
    )
  ) %>%
  hal_search(limit = 10)

## End(Not run)
```

hal_filter

Filter Results

Description

Filter Results

Usage

```
hal_filter(x, ...)

## S3 method for class 'HALQuery'
hal_filter(x, value, field = NULL, ...)
```

Arguments

x	An object of class HALQuery (typically returned by hal_api()).
...	Currently not used.
value	A character string specifying the value to be used to filter the results.
field	A character string specifying the field to filter along.

Value

An object of class [HALQuery](#).

Author(s)

N. Frerebeau

See Also

Other query tools: [hal_facet\(\)](#), [hal_group\(\)](#), [hal_query\(\)](#), [hal_select\(\)](#), [hal_sort\(\)](#)

Examples

```
library(magrittr)

## Simple filter
hal1 <- hal_api() %>% hal_filter("file", "submitType_s")
hal1$fq

## Advanced filter
hal2 <- hal_api() %>% hal_filter(c("THESE", "HDR"), "docType_s")
hal2$fq

## Multiple filters
hal3 <- hal_api() %>%
  hal_filter("[NOW-1MONTHS/DAY TO NOW/HOUR]", "submittedDate_tdate") %>%
  hal_filter("-notice", "submitType_s")
hal3$fq

## Range filters
hal4 <- hal_api() %>% hal_filter("[2000 TO 2013]", "submittedDateY_i")
hal4$fq

hal5 <- hal_api() %>% hal_filter("[Aa TO Ab]", "city_s")
hal5$fq
```

hal_group	<i>Group Results</i>
-----------	----------------------

Description

Group Results

Usage

```
hal_group(x, ...)
```

```
## S3 method for class 'HALQuery'
```

```
hal_group(x, field, limit = 1, sort = "score", decreasing = FALSE, ...)
```

Arguments

x	An object of class HALQuery (typically returned by hal_api()).
...	Currently not used.
field	A character string specifying the field to group by.
limit	An integer giving the maximum number of results per group.
sort	A character string specifying the field to be used to sort the results.
decreasing	A logical scalar: should the sort be increasing or decreasing?

Value

An object of class [HALQuery](#).

Author(s)

N. Frerebeau

See Also

Other query tools: [hal_facet\(\)](#), [hal_filter\(\)](#), [hal_query\(\)](#), [hal_select\(\)](#), [hal_sort\(\)](#)

Examples

```
## Not run:
library(magrittr)

## Most recent publication by journal
hal_api() %>%
  hal_query("archéologie") %>%
  hal_select("producedDate_tdate") %>%
  hal_filter("ART" %IN% "docType_s") %>%
  hal_sort("producedDate_tdate", decreasing = TRUE) %>%
  hal_group(field = "journalTitle_s", limit = 1,
```

```

        sort = "producedDate_tdate", decreasing = TRUE) %>%
hal_search(limit = 10)

## End(Not run)

```

hal_query

Build Query

Description

Build Query

Usage

```

hal_query(x, ...)

## S3 method for class 'HALQuery'
hal_query(x, value, field = NULL, ...)

x %OR% y

x %AND% y

x %NOT% y

x %BY% y

x %IN% y

x %TO% y

```

Arguments

x, y	A character vector or an object of class HALQuery (typically returned by hal_api()).
...	Currently not used.
value	A vector giving the topics to be searched for (see details).
field	A character string specifying the field to search within. If NULL (the default), search within "text".

Value

An object of class [HALQuery](#).

Author(s)

N. Frerebeau

See Also

Other query tools: [hal_facet\(\)](#), [hal_filter\(\)](#), [hal_group\(\)](#), [hal_select\(\)](#), [hal_sort\(\)](#)

Examples

```
library(magrittr)

## Select fields
hal1 <- hal_api() %>% hal_query("asie")
hal1$q

term <- list("japon", "france")
hal2 <- hal_api() %>% hal_query(term, field = "title_t")
hal2$q

term <- list("Journal", c("Histoire", "History"))
hal3 <- hal_api() %>% hal_query(term, field = "title_t")
hal3$q

## Update query
hal4 <- hal_query(hal1, "agricol?")
hal4$q

## Operators
term <- list("Paris", "France", "history") %NOT% list("Texas", "history")
hal5 <- hal_api() %>% hal_query(term %IN% "text")
hal5$q

term <- c("aluminium", "fer") %BY% 3 %IN% "title_t"
hal6 <- hal_api() %>% hal_query(term)
hal6$q

term <- list("ecology" %IN% "title_t", "cell" %IN% "text")
hal7 <- hal_api() %>% hal_query(term)
hal7$q
```

hal_select

Select Fields

Description

Select Fields

Usage

```
hal_select(x, ...)

## S3 method for class 'HALQuery'
hal_select(x, ...)
```

Arguments

- x An object of class HALQuery (typically returned by [hal_api\(\)](#)).
- ... One or more [character](#) string separated by commas giving the fields to be returned.

Value

An object of class [HALQuery](#).

Author(s)

N. Frerebeau

See Also

Other query tools: [hal_facet\(\)](#), [hal_filter\(\)](#), [hal_group\(\)](#), [hal_query\(\)](#), [hal_sort\(\)](#)

Examples

```
library(magrittr)

## Select fields
hal1 <- hal_api() %>% hal_select("label_s")
hal1$f1

hal2 <- hal_api() %>% hal_select("halId_s", "uri_s")
hal2$f1

## Update query
hal3 <- hal_select(hal1, "docType_s")
hal3$f1
```

hal_sort

Sort Results

Description

Sort Results

Usage

```
hal_sort(x, field, decreasing = FALSE, ...)
```

```
hal_sort(x, field, decreasing = FALSE, ...)
```

Arguments

x	An object of class <code>HALQuery</code> (typically returned by <code>hal_api()</code>).
field	A character string specifying the field to be used to sort the results.
decreasing	A logical scalar: should the sort be increasing or decreasing?
...	Currently not used.

Value

An object of class `HALQuery`.

Author(s)

N. Frerebeau

See Also

Other query tools: `hal_facet()`, `hal_filter()`, `hal_group()`, `hal_query()`, `hal_select()`

Examples

```
library(magrittr)

## Select fields
hal1 <- hal_api() %>% hal_sort("docid")
hal1$sort

## Update query
hal2 <- hal_sort(hal1, "producedDate_tdate", decreasing = TRUE)
hal2$sort
```

search

Search

Description

Search

Usage

```
hal_count(x, ...)

hal_search(x, ...)

## S3 method for class 'HALQuery'
hal_count(x, path = "search", ...)

## S3 method for class 'HALQuery'
```

```

hal_search(
  x,
  path = "search",
  limit = 30,
  start = 0,
  cursor = FALSE,
  parse = "df",
  concat = ",",
  progress = getOption("odyssey.progress"),
  verbose = getOption("odyssey.verbose"),
  ...
)

```

Arguments

x	An object of class HALQuery (typically returned by <code>hal_api()</code>).
...	Currently not used.
path	A character string specifying the url path (see details).
limit	An integer giving the maximum number of results. According to HAL policy, it cannot exceed 10000.
start	An integer specifying an absolute offset in the complete sorted list of matches to be used as the beginning of the current page. Only used if <code>cursor</code> is FALSE.
cursor	A logical scalar: should a cursor be used for the pagination of results? If TRUE, the <code>sort</code> parameter of the query will set to "docid asc".
parse	A character string specifying the type of the results. It must be one of "df" (default) or "list" (see <code>solrium::SolrClient()</code>).
concat	A character string specifying the character to concatenate elements of longer than length 1.
progress	A logical scalar: should a progress bar for for the request be printed? Only used if R is being used interactively.
verbose	A logical scalar: should extra information be reported?

Value

- `hal_count()` returns the total number of results.
- `hal_search()` returns a [data.frame](#) or [list](#) (according to type).

Author(s)

N. Frerebeau

References

- [Apache Solr documentation.](#)
- [HAL search documentation.](#)
- [HAL reference frame documentation.](#)

See Also

Other search tools: [hal_download\(\)](#)

Examples

```
## Not run:
library(magrittr)

## Simple search
topic <- list("archéologie", "Celtés", "France") # Combined with AND
hal_api() %>%
  hal_query(topic) %>%
  hal_search(limit = 10)

## Get a list of archaeological journals
topic <- c("archéologie", "archaeology", "archäologie") # Combined with OR
hal_api() %>%
  hal_query(topic) %>%
  hal_select("title_s", "issn_s") %>%
  hal_filter(" " %T0% "*" %IN% "issn_s") %>%
  hal_sort("title_s") %>%
  hal_search(path = "ref/journal")

## Get the most recent archaeological publication (in French) by journal
hal_api() %>%
  hal_query("archéologie") %>%
  hal_select("producedDate_tdate") %>%
  hal_filter("ART" %IN% "docType_s") %>%
  hal_sort("producedDate_tdate", decreasing = TRUE) %>%
  hal_group(
    field = "journalTitle_s",
    sort = "producedDate_tdate",
    decreasing = TRUE
  ) %>%
  hal_search(limit = 10)

## Get a list of archaeological laboratories
## (only joint laboratories of the CNRS and a French university)
topic <- list("archéologie" %IN% "text", "UMR" %IN% "code_t")
hal_api() %>%
  hal_query(topic) %>%
  hal_select("name_s", "acronym_s", "code_s") %>%
  hal_filter("VALID" %IN% "valid_s") %>%
  hal_sort("acronym_s", decreasing = TRUE) %>%
  hal_search(path = "ref/structure", limit = 15)

## End(Not run)
```

Index

* API

hal_api, 2

* query tools

hal_facet, 4

hal_filter, 5

hal_group, 7

hal_query, 8

hal_select, 9

hal_sort, 10

* search tools

hal_download, 3

search, 11

%AND% (hal_query), 8

%BY% (hal_query), 8

%IN% (hal_query), 8

%NOT% (hal_query), 8

%OR% (hal_query), 8

%TO% (hal_query), 8

character, 4–8, 10–12

data.frame, 12

hal_api, 2

hal_api(), 3, 4, 6–8, 10–12

hal_count (search), 11

hal_download, 3, 13

hal_facet, 4, 6, 7, 9–11

hal_filter, 5, 5, 7, 9–11

hal_group, 5, 6, 7, 9–11

hal_query, 5–7, 8, 10, 11

hal_search (search), 11

hal_select, 5–7, 9, 9, 11

hal_sort, 5–7, 9, 10, 10

HALQuery, 5–8, 10, 11

integer, 3, 4, 7, 12

list, 2, 5, 12

logical, 3, 5, 7, 11, 12

search, 3, 11

solrium::SolrClient(), 12

vector, 8