Package ‘rbhl’

October 14, 2022

Title Interface to the ‘Biodiversity’ ‘Heritage’ Library


Type Package

Version 0.9.2

License MIT + file LICENSE

URL https://github.com/ropensci/rbhl (devel)
     https://docs.ropensci.org/rbhl/ (documentation)

BugReports https://github.com/ropensci/rbhl/issues

LazyData true

Encoding UTF-8

Language en-US

Imports crul, xml2, jsonlite, plyr, tibble

Suggests testthat, vcr

RoxygenNote 7.1.1

X-schema.org-applicationCategory Biodiversity

X-schema.org-keywords BHL, library, API, web-services, conservation, species, taxonomy, OCR, text-mining

X-schema.org-isPartOf https://ropensci.org

NeedsCompilation no

Author Scott Chamberlain [aut, cre] (<https://orcid.org/0000-0003-1444-9135>)

Maintainer Scott Chamberlain <myrmecocystus@gmail.com>

Repository CRAN

Date/Publication 2021-05-13 04:30:03 UTC
R topics documented:

rbhl-package .......................... 2
bhl_authorsearch .......................... 3
bhl_bioherlib .......................... 3
bhl_getauthormetadata ....................... 4
bhl_getcollections .......................... 5
bhl_getinstitutions ......................... 6
bhl_getitemmetadata .......................... 7
bhl_getlanguages .......................... 8
bhl_getpagemetadata ......................... 9
bhl_getpartmetadata ......................... 10
bhl_gettitlemetadata ....................... 10
bhl_namemetadata .......................... 11
bhl_namesearch ........................... 12
bhl_openurl .............................. 13
bhl_pages ................................. 14
bhl_pagesearch ........................... 15
bhl_publicationsearch ...................... 15
bhl_publicationsearchadv ................... 16
bhl_subjectmetadata ....................... 18
bhl_subjectsearch ......................... 19
rbhl-defunct ............................. 19
rbhlmethods .............................. 20

Index .............................. 21

| rbhl-package | rbhl |  |

Description

R interface to the Biodiversity Heritage Library API.

You need an API key to use the Biodiversity Heritage Library API. Get your BHL API key at http://www.biodiversitylibrary.org/getapikey.aspx. Put your API key either in your .Renviron file (or similar) with the name BHL_KEY, or in your .Rprofile file with the name bhl_key, and the we’ll use your API key without you having to enter it every time you run a search.

See rbhl-defunct for defunct functions.

Author(s)

Scott Chamberlain
bhl_authorsearch  Author search

Description

Return a list of authors that match (fully or partially) the specified search string.

Usage

bhl_authorsearch(name = NULL, as = "table", key = NULL, ...)

Arguments

name  full or partial name of the author for which to search (last name listed first, i.e. 'Darwin, Charles') (character)
as  (character) Return a list ("list"), json ("json"), xml ("xml"), or parsed table ("table", default). Note that as="table" can give different data format back depending on the function - for example, sometimes a data.frame and sometimes a character vector.
key  Your BHL API key, either enter, or loads from your .Renviron as BHL_KEY or from .Rprofile as bhl_key.
...  Curl options passed on to crul::HttpClient()

Details

The names searched are those contained in MARC 100a, 110a, 111a, 700a, 710a, and 711a library records.

Examples

```r
## Not run:
bhl_authorsearch(name='dimmock')
bhl_authorsearch(name='Jones')
```

## End(Not run)

bhl_bioherlib  Search BHL across many API methods.

Description

Search BHL across many API methods.
bhl_getauthormetadata

Get metadata about an author

Usage

bhl_bioherlib(
  method = "GetPageMetadata",
  pageid = NULL,
  ocr = FALSE,
  names = FALSE,
  as = "table",
  key = NULL,
  ...
)

Arguments

method The API method to use.
pageid The identifier of an individual page in a scanned book.
ocr return ocr text of the page (logical). Default: FALSE
names return the names that appear on the page (logical). Default: FALSE
as (character) Return a list ("list"), json ("json"), xml ("xml"), or parsed table ("table", default). Note that as="table" can give different data format back depending on the function - for example, sometimes a data.frame and sometimes a character vector.
key Your BHL API key, either enter, or loads from your .Renviron as BHL_KEY or from .Rprofile as bhl_key.
...
Curl options passed on to crul::HttpClient()

Examples

## Not run:
bhl_bioherlib(method='GetPageMetadata', pageid=1328690, ocr=TRUE, names=TRUE)
bhl_bioherlib(method='GetPageMetadata', pageid=1328690, ocr=TRUE, names=TRUE, as="xml")
bhl_bioherlib(method='GetPageMetadata', pageid=1328690, ocr=TRUE, names=TRUE, as="list")

## End(Not run)
Usage

bhl_getauthormetadata(
  id,
  id_type = "bhl",
  pubs = FALSE,
  as = "table",
  key = NULL,
  ...
)

Arguments

id (numeric) the identifier of an individual author
id_type (character) the type of identifier (bhl, biostor, viaf). Default: "bhl"
pubs (logical) TRUE to return the subject's publications. Default: FALSE
as (character) Return a list ("list"), json ("json"), xml ("xml"), or parsed table ("table", default). Note that as="table" can give different data format back depending on the function - for example, sometimes a data.frame and sometimes a character vector.
key Your BHL API key, either enter, or loads from your .Renviron as BHL_KEY or from .Rprofile as bhl_key.

Details

You may choose to include a list of the author's publications. The BHLType element identifies the type of each publication (Title or Part).

Examples

## Not run:
bhl_getauthormetadata(87509)
bhl_getauthormetadata(87509, pubs = TRUE, verbose = TRUE)
bhl_getauthormetadata(87509, as = "json")

## End(Not run)

bhl_getcollections Get a list of collections which are used to group titles and items. A single collection may contain either titles or items, but not both.

Description

Get a list of collections which are used to group titles and items. A single collection may contain either titles or items, but not both.
Usage

bhl_getcollections(as = "table", key = NULL, ...)

Arguments

as (character) Return a list ("list"), json ("json"), xml ("xml"), or parsed table ("table", default). Note that as="table" can give different data format back depending on the function - for example, sometimes a data.frame and sometimes a character vector.

key Your BHL API key, either enter, or loads from your .Renviron as BHL_KEY or from .Rprofile as bhl_key.

... Curl options passed on to `crul::HttpClient()`

Examples

```r
## Not run:
bhl_getcollections()
bhl_getcollections(as = 'list')
bhl_getcollections(as = 'json')
bhl_getcollections(as = 'xml')

## End(Not run)
```

bhl_getinstitutions GetInstitutions data

Description

GetInstitutions data

Usage

bhl_getinstitutions(as = "table", key = NULL, ...)

Arguments

as (character) Return a list ("list"), json ("json"), xml ("xml"), or parsed table ("table", default). Note that as="table" can give different data format back depending on the function - for example, sometimes a data.frame and sometimes a character vector.

key Your BHL API key, either enter, or loads from your .Renviron as BHL_KEY or from .Rprofile as bhl_key.

... Curl options passed on to `crul::HttpClient()`

Details

Get a list of institutions which have contributed to BHL in some way. These institutions may fill roles such as contributors, rights holders, and scanning institutions.
Examples

```r
## Not run:
bhl_getinstitutions()

## End(Not run)
```

---

**bhl_getitemmetadata**  
*Return metadata about an item.*

### Description

You may choose to include a list of the item’s pages.

### Usage

```r
bhl_getitemmetadata(
  itemid = NULL,
  pages = TRUE,
  ocr = FALSE,
  parts = FALSE,
  as = "table",
  key = NULL,
  ...
)
```

### Arguments

- `itemid`  
  item id (character)

- `pages`  
  return the item’s pages (TRUE/FALSE)

- `ocr`  
  (logical) TRUE to return the ocr for the item’s pages. Setting this to TRUE apparently doesn’t return any actual ocr text, but leaving parameter here for now.

- `parts`  
  (logical) TRUE to return the item’s parts. Setting this to TRUE apparently doesn’t return any parts text, but leaving parameter here for now.

- `as`  
  (character) Return a list ("list"), json ("json"), xml ("xml"), or parsed table ("table", default). Note that as="table" can give different data format back depending on the function - for example, sometimes a data.frame and sometimes a character vector.

- `key`  
  Your BHL API key, either enter, or loads from your `.Renviron` as `BHL_KEY` or from `.Rprofile` as `bhl_key`.

- `...`  
  Curl options passed on to `curl::HttpClient()`
Examples

```r
## Not run:
bhl_getitemmetadata('16800', TRUE)
bhl_getitemmetadata('16800', TRUE, as='xml')
bhl_getitemmetadata('16800', TRUE, as='json')
bhl_getitemmetadata('16800', TRUE, as='list')
bhl_getitemmetadata(20419, pages=FALSE, parts=TRUE)

## End(Not run)
```

### bhl_getlanguages

*bhl_getlanguages*  
Get a list of languages in which materials in BHL have been written.

#### Description

Get a list of languages in which materials in BHL have been written.

#### Usage

```r
bhl_getlanguages(as = "table", key = NULL, ...)
```

#### Arguments

- `as` (character) Return a list ("list"), json ("json"), xml ("xml"), or parsed table ("table", default). Note that `as="table"` can give different data format back depending on the function - for example, sometimes a data.frame and sometimes a character vector.
- `key` Your BHL API key, either enter, or loads from your `.Renviron` as `BHL_KEY` or from `.Rprofile` as `bhl_key`.
- `...` Curl options passed on to `crl::HttpClient()`

#### Examples

```r
## Not run:
bhl_getlanguages()
bhl_getlanguages('json')
bhl_getlanguages('xml')
bhl_getlanguages('list')

## End(Not run)
```
bhl_getpagemetadata  

Return metadata about a page.

Description

You may choose to include the OCR text and a list of names found on the page.

Usage

bhl_getpagemetadata(
  page = NULL,
  ocr = FALSE,
  names = FALSE,
  as = "table",
  key = NULL,
  ...
)

Arguments

page    page number to get
ocr     return ocr text of the page (TRUE/FALSE)
names   return the names that appear on the page (TRUE/FALSE)
as      (character) Return a list ("list"), json ("json"), xml ("xml"), or parsed table ("table", default). Note that as="table" can give different data format back depending on the function - for example, sometimes a data.frame and sometimes a character vector.
key     Your BHL API key, either enter, or loads from your .Renviron as BHL_KEY or from .Rprofile as bhl_key.
...     Curl options passed on to `curl::HttpClient()`

Examples

## Not run:
bhl_getpagemetadata(page=1328690, ocr=TRUE)
bhl_getpagemetadata(page=1328690, ocr=TRUE, as='json')
bhl_getpagemetadata(page=1328690, ocr=TRUE, as='xml')
bhl_getpagemetadata(page=1328690, ocr=TRUE, as='list')

## End(Not run)
bhl_getpartmetadata  
*Return a list of an item’s pages.*

**Description**

Return a list of an item’s pages.

**Usage**

```r
bhl_getpartmetadata(partid, key = NULL, ...)
```

**Arguments**

- `partid`  
  The identifier of an individual part (article, chapter, etc) (numeric)
- `key`  
  Your BHL API key, either enter, or loads from your .Renviron as BHL_KEY or from .Rprofile as bhl_key.
- `...`  
  Curl options passed on to `crl::HttpClient()`

**Examples**

```r
## Not run:
bhl_getpartmetadata(10409)
## End(Not run)
```

bhl_gettitlemetadata  
*Get title metadata*

**Description**

Return metadata about a title. You may choose to include a list of the items (books) associated with the title.

**Usage**

```r
bhl_gettitlemetadata(titleid = NA, items = FALSE, as = "list", key = NULL, ...)
```

**Arguments**

- `titleid`  
  the identifier of an individual title (numeric)
- `items`  
  (logical) TRUE of FALSE (default) to include items
- `as`  
  (character) Return a list ("list"), json ("json"), xml ("xml"), or parsed table ("table", default). Note that as="table" can give different data format back depending on the function - for example, sometimes a data.frame and sometimes a character vector.
bhl_namemetadata

Get basic title, item, and page metadata for each page on which the specified name appears.

Description

Get basic title, item, and page metadata for each page on which the specified name appears.

Usage

bhl_namemetadata(
    namebankid = NULL,
    name = NULL,
    id_type = NULL,
    as = "list",
    key = "NULL",
    ...
)

Arguments

namebankid (numeric) (not used if 'name' specified) NameBank identifier for a name
name (character) (not used if 'namebankid' specified) a complete name string
id_type (character) the type of identifier (namebank, eol, gni, ion, col, gbif, itis, ipni, worms). Not used if name is specified
as (character) Return a list ("list"), json ("json"), xml ("xml"), or parsed table ("table", default). Note that as="table" can give different data format back depending on the function - for example, sometimes a data.frame and sometimes a character vector.
key Your BHL API key, either enter, or loads from your .Renviron as BHL_KEY or from .Rprofile as bhl_key.
... Curl options passed on to crul::HttpClient()
bhl_namesearch

Search for a particular name.

Description

Names both with and without NameBank identifiers are returned.

Usage

bhl_namesearch(name = NULL, as = "table", key = NULL, ...)

Arguments

- name: species name (character)
- as: (character) Return a list ("list"), json ("json"), xml ("xml"), or parsed table ("table", default). Note that `as="table"` can give different data format back depending on the function - for example, sometimes a data.frame and sometimes a character vector.
- key: Your BHL API key, either enter, or loads from your .Renviron as BHL_KEY or from .Rprofile as bhl_key.
- ...: Curl options passed on to `cru::HttpClient()`

Examples

## Not run:
```
bhl_namesearch("poa annua")
bhl_namesearch(name='helianthus annuus')
bhl_namesearch(name='helianthus annuus', as='xml')
bhl_namesearch(name='helianthus annuus', as='json')
```

## End(Not run)
bhl_openurl

Not sure how this differs from their other API...

Description

Not sure how this differs from their other API...

Usage

bhl_openurl(
  genre = NULL,
  title = NULL,
  aufirst = NULL,
  aulast = NULL,
  date = NULL,
  spage = NULL,
  issue = NULL,
  version = 0.1,
  as = "list",
  key = NULL,
  ...
)

Arguments

gene Book genre

title Book title

aufirst First author

aulast Last author

date Date of publication

spage Start page

issue Issue number

version One of 0.1 or 1.0

as (character) Return a list ("list"), json ("json"), xml ("xml"), or parsed table ("table", default). Note that as="table" can give different data format back depending on the function - for example, sometimes a data.frame and sometimes a character vector.

key Your BHL API key, either enter, or loads from your .Renviron as BHL_KEY or from .Rprofile as bhl_key.

... Curl options passed on to curl::HttpClient()
## bhl_pages

### Get many OCR-generated pages given a single item id

#### Description

Get many OCR-generated pages given a single item id

#### Usage

```r
bhl_pages(itemid, key = NULL, ...)
```

#### Arguments

- `itemid`: the item id (character). Required
- `key`: Your BHL API key, either enter, or loads from your `.Renviron` as `BHL_KEY` or from `.Rprofile` as `bhl_key`
- `...`: Curl options passed on to `crl::HttpClient()`

#### Examples

```r
## Not run:
x <- bhl_publicationsearch('Selborne', year=1825, collection=4, language='eng')
bhl_pages(itemid = x$ItemID[1])

## End(Not run)
```
bhl_pagesearch

Search an item for pages containing the specified text

Description

Search an item for pages containing the specified text

Usage

bhl_pagesearch(id, text, as = "table", key = NULL, ...)

Arguments

id
(integer) BHL identifier of the item to be searched

text
(character) the text for which to search

as
(character) Return a list ("list"), json ("json"), xml ("xml"), or parsed table ("table", default). Note that as="table" can give different data format back depending on the function - for example, sometimes a data.frame and sometimes a character vector.

key
Your BHL API key, either enter, or loads from your .Renviron as BHL_KEY or from .Rprofile as bhl_key.

...  Curl options passed on to curl::HttpClient()

Examples

## Not run:
bhl_pagesearch(22004, "dog")
bhl_pagesearch(22004, "dog", as = "json")

## End(Not run)

bhl_publicationsearch  Publication search

Description

Publication search
bhl_publicationsearchadv

Description

Publication search advanced
Usage

```r
bhl_publicationsearchadv(
  title = NULL,
  titleop = NULL,
  authorname = NULL,
  year = NULL,
  subject = NULL,
  language = NULL,
  collection = NULL,
  notes = NULL,
  notesop = NULL,
  text = NULL,
  textop = NULL,
  page = NULL,
  as = "table",
  key = NULL,
  ...
)
```

Arguments

- `title` (character) a title for which to search
- `titleop` (character) 'all' to search for all specified words in the title fields; 'phrase' to search for the exact phrase specified
- `authorname` (character) an author name for which to search
- `year` (integer) a four-digit publication year for which to search
- `subject` (character) a subject for which to search
- `language` (character) a language code; search will only return publications in the specified language
- `collection` (character) a collection id; search will only return publications from the specified collection
- `notes` (character) one or more words for which to search in the publication notes
- `notesop` (character) 'all' to search for all specified words in the notes field; 'phrase' to search for the exact phrase specified
- `text` (character) one or more words for which to search in the text of the publications
- `textop` (character) 'all' to search for all specified words in the text field; 'phrase' to search for the exact phrase specified
- `page` (integer) 1 for the first page of results, 2 for the second, and so on
- `as` (character) Return a list ("list"), json ("json"), xml ("xml"), or parsed table ("table", default). Note that as="table" can give different data format back depending on the function - for example, sometimes a data.frame and sometimes a character vector.
- `key` Your BHL API key, either enter, or loads from your .Renviron as BHL_KEY or from .Rprofile as bhl_key.
- `...` Curl options passed on to `crul::HttpClient()`
bhl_subjectmetadata

Get metadata about a subject

Description

Get metadata about a subject

Usage

bhl_subjectmetadata(subject, pubs = FALSE, as = "table", key = NULL, ...)

Arguments

subject (character) the subject for which to return metadata
pubs (logical) TRUE to return the subject’s publications. Default: FALSE
as (character) Return a list ("list"), json ("json"), xml ("xml"), or parsed table ("table", default). Note that as="table" can give different data format back depending on the function - for example, sometimes a data.frame and sometimes a character vector.
key Your BHL API key, either enter, or loads from your .Renviron as BHL_KEY or from .Rprofile as bhl_key.
... Curl options passed on to curl::HttpClient()

Details

You may choose to include a list of the subject’s publications. The BHLType element identifies the type of each publication (Title or Part).

Examples

## Not run:
bhl_subjectmetadata(subject = "water")
x <- bhl_subjectmetadata(subject = "water", pubs = TRUE)
head(x$Publications[[1]])

## End(Not run)
bhl_subjectsearch

Return a list of subjects that match (fully or partially) the specified search string.

Description

Return a list of subjects that match (fully or partially) the specified search string.

Usage

bhl_subjectsearch(subject, as = "table", key = NULL, ...)

Arguments

subject  
the full or partial subject for which to search (character)

as  
(character) Return a list ("list"), json ("json"), xml ("xml"), or parsed table ("table", default). Note that as="table" can give different data format back depending on the function - for example, sometimes a data.frame and sometimes a character vector.

key  
Your BHL API key, either enter, or loads from your .Renviron as BHL_KEY or from .Rprofile as bhl_key.

...  
Curl options passed on to crul::HttpClient()

Examples

## Not run:
bhl_subjectsearch('diptera')
bhl_subjectsearch('diptera', "json")

## End(Not run)

rbhl-defunct

Defunct functions in rbhl

Description

- bhl_getpartendnote(): BHL removed this API method.
- bhl_gettitleendnote(): BHL removed this API method.
- bhl_booksearch(): method removed in API v3, see bhl_publicationsearch()
- bhl_partsearch(): method removed in API v3, see bhl_publicationsearch()
- bhl_titlesearchsimple(): method removed in API v3, see bhl_publicationsearch()
- bhl_getauthorparts(): method removed in API v3, see bhl_getauthormetadata()
- bhl_getauthortitles(): method removed in API v3, see bhl_getauthormetadata()
- `bhl_getitembyidentifier()`: method removed in API v3, see `bhl_getitemmetadata()`
- `bhl_getitempages()`: method removed in API v3, see `bhl_getitemmetadata()`
- `bhl_getitemparts()`: method removed in API v3, see `bhl_getitemmetadata()`
- `bhl_getpagenames()`: method removed in API v3, see `bhl_getpagemetadata()`
- `bhl_getpageocrtext()`: method removed in API v3, see `bhl_getpagemetadata()`
- `bhl_getpartbyidentifier()`: method removed in API v3, see `bhl_getpartmetadata()`
- `bhl_getpartnames()`: method removed in API v3, see `bhl_getpartmetadata()`
- `bhl_getsubjectparts()`: method removed in API v3, see `bhl_subjectmetadata()`
- `bhl_getsubjecttitles()`: method removed in API v3, see `bhl_subjectmetadata()`
- `bhl_getpartbibTEX()`: method removed in API v3
- `bhl_gettitlebibTEX()`: method removed in API v3
- `bhl_gettitlebyidentifier()`: API method merged in API v3, see `bhl_gettitlemetadata()`
- `bhl_gettitleitems()`: API method merged in API v3, see `bhl_gettitlemetadata()`
- `bhl_getunpublisheditems()`: method removed in API v3
- `bhl_getunpublishedparts()`: method removed in API v3
- `bhl_getunpublishedtitles()`: method removed in API v3
- `bhl_namecount()`: method removed in API v3
- `bhl_namelist()`: method removed in API v3
- `bhl_namegetdetail()`: method renamed - see `bhl_namemetadata()`

---

**rbhlmethods**

*Data.frame of all the BHL API methods from the BHL website.*

**Description**

Data.frame of all the BHL API methods from the BHL website.
Index

* datasets
  rbhlmethods, 20
* package
  rbhl-package, 2
* publication-search
  bhl_publicationsearch, 15
  bhl_publicationsearchadv, 16
  bhl_authorsearch, 3
  bhl_bioherlib, 3
  bhl_booksearch(), 19
  bhl_getauthormetadata, 4
  bhl_getauthormetadata(), 19
  bhl_getauthorparts(), 19
  bhl_getauthortitles(), 19
  bhl_getcollections, 5
  bhl_getinstitutions, 6
  bhl_getitembyidentifier(), 20
  bhl_getitemmetadata, 7
  bhl_getitemmetadata(), 20
  bhl_getitempages(), 20
  bhl_getitemparts(), 20
  bhl_getlanguages, 8
  bhl_getpagemetadata, 9
  bhl_getpagemetadata(), 20
  bhl_getpagenames(), 20
  bhl_getpageocrtext(), 20
  bhl_getpartbibtext(), 20
  bhl_getpartbyidentifier(), 20
  bhl_getpartmetadata, 10
  bhl_getpartmetadata(), 20
  bhl_getpartnames(), 20
  bhl_getsubjectparts(), 20
  bhl_getsubjecttitles(), 20
  bhl_gettitlebibTex(), 20
  bhl_gettitlebyidentifier(), 20
  bhl_gettitleitems(), 20
  bhl_gettitlemetadata, 10
  bhl_gettitlemetadata(), 20
  bhl_getunpublisheditems(), 20
  bhl_getunpublishedparts(), 20
  bhl_getunpublishedtitles(), 20
  bhl_namecount(), 20
  bhl_namelist(), 20
  bhl_namemetadata, 11
  bhl_namemetadata(), 20
  bhl_namesearch, 12
  bhl_openurl, 13
  bhl_pages, 14
  bhl_pagesearch, 15
  bhl_partsearch(), 19
  bhl_publicationsearch, 15, 18
  bhl_publicationsearch(), 19
  bhl_publicationsearchadv, 16, 16
  bhl_subjectmetadata, 18
  bhl_subjectmetadata(), 20
  bhl_subjectsearch, 19
  bhl_titlesearchsimple(), 19
  crul::HttpClient(), 3–19
  rbhl (rbhl-package), 2
  rbhl-defunct, 2, 19
  rbhl-package, 2
  rbhlmethods, 20