Pybman Documentation

Release 2019.10.1

Donatus Herre

Oct 01, 2019

Contents

1 Reference **1**

1.1 pybman.client module . . . 1
1.2 pybman.data module . . . 1
1.3 pybman.extract module . . . 3
1.4 pybman.inspector module . . . 5
1.5 pybman.local module . . . 6
1.6 pybman.query module . . . 6
1.7 pybman.rest module . . . 7
1.8 pybman.utils module . . . 9

2 Indices and tables **11**

Python Module Index **13**

Index **15**

CHAPTER 1

Reference

1.1 pybman.client module

class pybman.client.**Client**(*secret=''*)

Bases: object

get_data(*ctx_id=None*, *ou_id=None*, *pers_id=None*, *lang_id=None*, *jour_name=None*, *misc_query=None*)

update_data(*idx*, *item_data*, *comment*)

1.2 pybman.data module

class pybman.data.**DataSet**(*data_id*, *data=None*, *raw=None*)

Bases: object

get_cone_persons()

extract persons' CoNE IDs and associated names

get_contexts()

extract IDs of records an associated IDs of collection

get_creators()

extract data of creators from records

get_creators_data()

extract persons' CoNE IDs and associated records

get_creators_from_records()

extract creators lists from records

get_genre_data()

extract genres and associated records

get_genre_relationships()
extract genres and associated source genres

get_genres()
extract genres and associated IDs of records

get_item(*item_id*)
extract record with given record ID

get_items_from_year(*year*)
extract IDs of records with given publication year

get_items_from_year_data(*year*)
extract records with given publication year

get_items_released()
extract IDs of records with state 'released' and associated records

get_items_submitted()
extract records with state 'submitted' and associated IDs

get_items_with_external_url()
extract records with external url

get_items_with_genre(*genre*)
extract IDs of records with state 'released' and associated records

get_items_with_identifier_uri()
extract records with identifier uri

get_items_with_source_genre(*source_genre*)
extract records with given source genre

get_journals()
extract publication sources from records with JOURNAL as source genre

get_journals_data()
extract records with JOURNAL as source genre

get_languages()
extract publication languages and associated record IDs

get_languages_data()
extract publication languages and associated records

get_organizations()
extract organizations of creators and associated IDs of records

get_places()
extract publication places and associated IDs of records

get_publishers()
extract publishers and associated IDs of records

get_series()
extract publication source from records with SERIES as source genre

get_source_from_items_with_source_genre(*source_genre*)
extract data of record's source with given source genre

get_source_genres()
extract genres of publication sources and associated record IDs

get_source_genres_data()
extract genres of publication sources and associated records

get_sources_identifiers()
extract identifers of publication sources and associated record IDs

get_sources_titles()
extract titles of publication sources and associated record IDs

get_titles()
extract titles and associated IDs of records

get_titles_from_source()
extract titles of sources and associated IDs of records

get_years()
extract publication years and associated record IDs

get_years_data()
extract publication years and associated records

1.3 pybman.extract module

data extraction routines

pybman.extract.**creators_from_item**(*item*)
extract creators of item

pybman.extract.**creators_from_source**(*source*)
extract sources' titles and genres from item

pybman.extract.**ctx_from_item**(*item*)
extract identifier of item

pybman.extract.**ctx_idx_from_item**(*item*)
extract identifier of item

pybman.extract.**data**(*item*)
extract data of item

pybman.extract.**date_accepted_from_item**(*item*)
extract date of acceptance from item

pybman.extract.**date_created_from_item**(*item*)
extract date of creation from item

pybman.extract.**date_from_item**(*item*)
extract date from item

pybman.extract.**date_modified_from_item**(*item*)
extract date of modification from item

pybman.extract.**date_pubonline_from_item**(*item*)
extract publication date (online) from item

pybman.extract.**date_pubprint_from_item**(*item*)
extract publication date (print) from item

pybman.extract.**date_submitted_from_item**(*item*)
extract date of submission from item

`pybman.extract.`**`field_from_creator`**(*field*, *creator*)
extract field from creator of item

`pybman.extract.`**`field_from_metadata`**(*field*, *item*, *value=True*)
extract field from metadata of item

`pybman.extract.`**`field_from_pubinfo`**(*field*, *item*)
extract field from publishing info of item

`pybman.extract.`**`field_in_level`**(*field*, *level*)
true if field is in level, false otherwise

`pybman.extract.`**`genre_from_item`**(*item*)
extract genre of item

`pybman.extract.`**`identifers_from_item`**(*item*)
extract values and types of identifiers

`pybman.extract.`**`idx_from_item`**(*item*)
extract identifier of item

`pybman.extract.`**`items`**(*records*)
iterate items

`pybman.extract.`**`iter_fields`**(*field*, *level*)
iterate over field of level

`pybman.extract.`**`languages_from_item`**(*item*)
extract languages of item

`pybman.extract.`**`list_from_level`**(*field*, *level*)
extract list from field in level of item

`pybman.extract.`**`metadata`**(*item*)
extract meta data of item

`pybman.extract.`**`organizations_from_item`**(*item*)
extract organizations from item

`pybman.extract.`**`organizations_identifier_from_creator`**(*creator*)
extract organizations' identifiers from item

`pybman.extract.`**`organizations_name_from_creator`**(*creator*)
extract organizations' names from item

`pybman.extract.`**`persons_affiliation_from_creator`**(*creator*)
exract persons' affiliations from creator of item

`pybman.extract.`**`persons_from_item`**(*item*)
extract persons from item

`pybman.extract.`**`persons_id_from_creator`**(*creator*)
exract persons' ID and type of ID from creator of item

`pybman.extract.`**`persons_identifier_from_creator`**(*creator*)
exract persons' identifier from creator of item

`pybman.extract.`**`persons_name_from_creator`**(*creator*)
exract persons' name from creator of item

`pybman.extract.`**`persons_organizations_from_creator`**(*creator*)
exract persons' organizations from creator of item

pybman.extract.**place_from_item**(*item*)
extract publishing place of item

pybman.extract.**pubinfo_from_item**(*item*)
extract publishing info of item

pybman.extract.**publisher_from_item**(*item*)
extract publisher from item

pybman.extract.**role_from_creator**(*creator*)
extract role from creator of item

pybman.extract.**source_titles**(*collection*)
extract source titles from records

pybman.extract.**source_titles_genres**(*collection*)
extract source titles and genres from records

pybman.extract.**sources**(*item*)
iterate over sources of item

pybman.extract.**sources_identifiers_from_item**(*item*)
extract values and types of identifiers from item's sources

pybman.extract.**sources_persons_affiliations_from_item**(*item*)
extract persons affilations from source of item

pybman.extract.**sources_persons_from_item**(*item*)
extract sources' creators from item if person

pybman.extract.**sources_persons_id_from_item**(*item*)
extract persons ID from source of item

pybman.extract.**sources_titles_from_item**(*item*)
extract sources' titles from item

pybman.extract.**sources_titles_genres_from_item**(*item*)
extract sources' titles and genres from item

pybman.extract.**title_from_item**(*item*)
extract title of item

pybman.extract.**titles_from_records**(*records*)
extract titles from records

pybman.extract.**type_from_creator**(*creator*)
extract type of creator

pybman.extract.**value_from_level**(*field*, *level*)
extract value from field in level of item

1.4 pybman.inspector module

class pybman.inspector.**Inspector**(*client*, *records*)
Bases: object

change_genre(*new_genre*, *old_genre*)

change_pers_name(*old_family_name=None*, *new_family_name=None*, *old_given_name=None*, *new_given_name=None*)

change_source_genre(*new_genre*, *old_genre*)

check_publication_date()

check_publication_titles(*clean=False*)

check_publication_uri()

check_publication_url()

check_publishers(*clean=False*)

check_publishers_omission(*clean=False*)

check_publishing_places(*clean=False*)

check_publishing_places_omission(*clean=False*)

check_source_titles(*clean=False*)

clean_publishers()

clean_publishing_places()

clean_source_titles()

clean_titles()

update_genre(*new_genre*, *old_genre*)

update_source_genre(*new_genre*, *old_genre*)

1.5 pybman.local module

class pybman.local.**LocalData**(*base_dir='./data/'*, *ous_dir='ous'*, *ctx_dir='ctx'*, *pers_dir='pers'*, *create=False*)

Bases: object

find_data_path(*pattern*)

generate_data_path(*data_id*)

get_data(*pattern*)

store_data(*idx*, *dict_data*)

1.6 pybman.query module

class pybman.query.**AllQuery**

Bases: object

get_files_query()

get_locators_query()

class pybman.query.**ConeQuery**

Bases: object

class pybman.query.**ContextQuery**

Bases: object

get_item_query(*ctx_id*)

get_released_item_query(*ctx_id*)

class pybman.query.**JournalQuery**
Bases: object

get_item_query(*jour_name*)

get_released_item_query(*jour_name*)

class pybman.query.**LangQuery**
Bases: object

get_item_query(*lang_id*)

get_released_item_query(*lang_id*)

class pybman.query.**OrgUnitQuery**
Bases: object

get_item_query(*ou_id*)

get_item_released_query(*ou_id*)

class pybman.query.**PersQuery**
Bases: object

get_item_query(*cone_id*)

get_item_released_query(*cone_id*)

class pybman.query.**Query**
Bases: object

1.7 pybman.rest module

class pybman.rest.**BaseController**(*url='https://pure.mpg.de/'*)
Bases: object

class pybman.rest.**ConeController**
Bases: *pybman.rest.BaseController*

class pybman.rest.**ContextRestController**
Bases: *pybman.rest.RestController*

get(*ctx_id='ctx_924547'*)

get_all()

class pybman.rest.**FeedRestController**
Bases: *pybman.rest.RestController*

get_recent_oa()

get_recent_releases()

get_recent_releases_for_ou()

get_search_as_feed()

class pybman.rest.**ItemRestController**(*secret=''*)
Bases: *pybman.rest.LoginRestController*

count_items(*query=None*)

release_item(*item_id*, *data*, *comment*)

revise_item()

scroll_items(*scroll_id*, *counter*)

search_items(*query=None*)

submit_item()

update_item(*item_id*, *data*)

withdraw_item()

class pybman.rest.**JournalConeController**
Bases: *pybman.rest.ConeController*

get_entities()

get_entity(*journal_id='954928372938'*)

search_entity(*cone_id*)

class pybman.rest.**LanguageConeController**
Bases: *pybman.rest.ConeController*

full_init()

get_entities()

get_entity(*language_id='deu'*)

init()

search_entity(*cone_id*)

class pybman.rest.**LoginRestController**(*auth=True*, *cred=''*)
Bases: *pybman.rest.RestController*

get_user()

login()

logout()

class pybman.rest.**OrgUnitRestController**
Bases: *pybman.rest.RestController*

child_organizations(*ou_id='ou_907574'*)

get(*ou_id='ou_907574'*)

get_all()

top_level()

class pybman.rest.**PersonConeController**
Bases: *pybman.rest.ConeController*

full_init()

get_entities()

get_entity(*person_id='persons32341'*)

init()

ous_graph()

search_entity(*cone_id*)

class pybman.rest.**RestController**
Bases: *pybman.rest.BaseController*

get_docs()
retrieve api docs of rest interface

1.8 pybman.utils module

pybman.utils.**add_value**(*d*, *dkey=''*, *dvalue=''*)

pybman.utils.**check_url**(*url*)

pybman.utils.**clean_string**(*string*)

pybman.utils.**get_request**(*url*, *params=None*, *headers=None*, *json_response=True*)

pybman.utils.**post_request**(*url*, *params=None*, *headers=None*, *data=None*, *json_res=True*)

pybman.utils.**put_request**(*url*, *header*, *data*)

pybman.utils.**read_csv_with_header**(*path*)

pybman.utils.**read_json**(*path*)

pybman.utils.**read_plain_clean**(*path*)

pybman.utils.**resolve_path**(*path*)

pybman.utils.**url_exists**(*url*)

pybman.utils.**url_exists2**(*url*)

pybman.utils.**write_csv**(*path*, *results*)

pybman.utils.**write_json**(*path*, *data*)

pybman.utils.**write_list**(*path*, *results*)

CHAPTER 2

Indices and tables

- genindex
- modindex
- search

Python Module Index

p

pybman.client, 1
pybman.data, 1
pybman.extract, 3
pybman.inspector, 5
pybman.local, 6
pybman.query, 6
pybman.rest, 7
pybman.utils, 9

Index

A

add_value() (*in module pybman.utils*), 9
AllQuery (*class in pybman.query*), 6

B

BaseController (*class in pybman.rest*), 7

C

change_genre() (*pybman.inspector.Inspector method*), 5
change_pers_name() (*pybman.inspector.Inspector method*), 5
change_source_genre() (*pybman.inspector.Inspector method*), 5
check_publication_date() (*pybman.inspector.Inspector method*), 5
check_publication_titles() (*pybman.inspector.Inspector method*), 6
check_publication_uri() (*pybman.inspector.Inspector method*), 6
check_publication_url() (*pybman.inspector.Inspector method*), 6
check_publishers() (*pybman.inspector.Inspector method*), 6
check_publishers_omission() (*pybman.inspector.Inspector method*), 6
check_publishing_places() (*pybman.inspector.Inspector method*), 6
check_publishing_places_omission() (*pybman.inspector.Inspector method*), 6
check_source_titles() (*pybman.inspector.Inspector method*), 6
check_url() (*in module pybman.utils*), 9
child_organizations() (*pybman.rest.OrgUnitRestController method*), 8
clean_publishers() (*pybman.inspector.Inspector method*), 6
clean_publishing_places() (*pybman.inspector.Inspector method*), 6
clean_source_titles() (*pybman.inspector.Inspector method*), 6
clean_string() (*in module pybman.utils*), 9
clean_titles() (*pybman.inspector.Inspector method*), 6
Client (*class in pybman.client*), 1
ConeController (*class in pybman.rest*), 7
ConeQuery (*class in pybman.query*), 6
ContextQuery (*class in pybman.query*), 6
ContextRestController (*class in pybman.rest*), 7
count_items() (*pybman.rest.ItemRestController method*), 7
creators_from_item() (*in module pybman.extract*), 3
creators_from_source() (*in module pybman.extract*), 3
ctx_from_item() (*in module pybman.extract*), 3
ctx_idx_from_item() (*in module pybman.extract*), 3

D

data() (*in module pybman.extract*), 3
DataSet (*class in pybman.data*), 1
date_accepted_from_item() (*in module pybman.extract*), 3
date_created_from_item() (*in module pybman.extract*), 3
date_from_item() (*in module pybman.extract*), 3
date_modified_from_item() (*in module pybman.extract*), 3
date_pubonline_from_item() (*in module pybman.extract*), 3
date_pubprint_from_item() (*in module pybman.extract*), 3
date_submitted_from_item() (*in module pybman.extract*), 3

F

FeedRestController (*class in pybman.rest*), 7
field_from_creator() (*in module pybman.extract*), 3
field_from_metadata() (*in module pybman.extract*), 4
field_from_pubinfo() (*in module pybman.extract*), 4
field_in_level() (*in module pybman.extract*), 4
find_data_path() (*pybman.local.LocalData method*), 6
full_init() (*pybman.rest.LanguageConeController method*), 8
full_init() (*pybman.rest.PersonConeController method*), 8

G

generate_data_path() (*pybman.local.LocalData method*), 6
genre_from_item() (*in module pybman.extract*), 4
get() (*pybman.rest.ContextRestController method*), 7
get() (*pybman.rest.OrgUnitRestController method*), 8
get_all() (*pybman.rest.ContextRestController method*), 7
get_all() (*pybman.rest.OrgUnitRestController method*), 8
get_cone_persons() (*pybman.data.DataSet method*), 1
get_contexts() (*pybman.data.DataSet method*), 1
get_creators() (*pybman.data.DataSet method*), 1
get_creators_data() (*pybman.data.DataSet method*), 1
get_creators_from_records() (*pybman.data.DataSet method*), 1
get_data() (*pybman.client.Client method*), 1
get_data() (*pybman.local.LocalData method*), 6
get_docs() (*pybman.rest.RestController method*), 8
get_entities() (*pybman.rest.JournalConeController method*), 8
get_entities() (*pybman.rest.LanguageConeController method*), 8
get_entities() (*pybman.rest.PersonConeController method*), 8
get_entity() (*pybman.rest.JournalConeController method*), 8
get_entity() (*pybman.rest.LanguageConeController method*), 8
get_entity() (*pybman.rest.PersonConeController method*), 8
get_files_query() (*pybman.query.AllQuery method*), 6
get_genre_data() (*pybman.data.DataSet method*), 1
get_genre_relationships() (*pybman.data.DataSet method*), 1
get_genres() (*pybman.data.DataSet method*), 2
get_item() (*pybman.data.DataSet method*), 2
get_item_query() (*pybman.query.ContextQuery method*), 6
get_item_query() (*pybman.query.JournalQuery method*), 7
get_item_query() (*pybman.query.LangQuery method*), 7
get_item_query() (*pybman.query.OrgUnitQuery method*), 7
get_item_query() (*pybman.query.PersQuery method*), 7
get_item_released_query() (*pybman.query.OrgUnitQuery method*), 7
get_item_released_query() (*pybman.query.PersQuery method*), 7
get_items_from_year() (*pybman.data.DataSet method*), 2
get_items_from_year_data() (*pybman.data.DataSet method*), 2
get_items_released() (*pybman.data.DataSet method*), 2
get_items_submitted() (*pybman.data.DataSet method*), 2
get_items_with_external_url() (*pybman.data.DataSet method*), 2
get_items_with_genre() (*pybman.data.DataSet method*), 2
get_items_with_identifier_uri() (*pybman.data.DataSet method*), 2
get_items_with_source_genre() (*pybman.data.DataSet method*), 2
get_journals() (*pybman.data.DataSet method*), 2
get_journals_data() (*pybman.data.DataSet method*), 2
get_languages() (*pybman.data.DataSet method*), 2
get_languages_data() (*pybman.data.DataSet method*), 2
get_locators_query() (*pybman.query.AllQuery method*), 6
get_organizations() (*pybman.data.DataSet method*), 2
get_places() (*pybman.data.DataSet method*), 2
get_publishers() (*pybman.data.DataSet method*), 2
get_recent_oa() (*pybman.rest.FeedRestController method*), 7
get_recent_releases() (*pyb-*

man.rest.FeedRestController method), 7
get_recent_releases_for_ou() (pybman.rest.FeedRestController method), 7
get_released_item_query() (pybman.query.ContextQuery method), 6
get_released_item_query() (pybman.query.JournalQuery method), 7
get_released_item_query() (pybman.query.LangQuery method), 7
get_request() (in module pybman.utils), 9
get_search_as_feed() (pybman.rest.FeedRestController method), 7
get_series() (pybman.data.DataSet method), 2
get_source_from_items_with_source_genre() (pybman.data.DataSet method), 2
get_source_genres() (pybman.data.DataSet method), 2
get_source_genres_data() (pybman.data.DataSet method), 2
get_sources_identifiers() (pybman.data.DataSet method), 3
get_sources_titles() (pybman.data.DataSet method), 3
get_titles() (pybman.data.DataSet method), 3
get_titles_from_source() (pybman.data.DataSet method), 3
get_user() (pybman.rest.LoginRestController method), 8
get_years() (pybman.data.DataSet method), 3
get_years_data() (pybman.data.DataSet method), 3

I

identifers_from_item() (in module pybman.extract), 4
idx_from_item() (in module pybman.extract), 4
init() (pybman.rest.LanguageConeController method), 8
init() (pybman.rest.PersonConeController method), 8
Inspector (class in pybman.inspector), 5
ItemRestController (class in pybman.rest), 7
items() (in module pybman.extract), 4
iter_fields() (in module pybman.extract), 4

J

JournalConeController (class in pybman.rest), 8
JournalQuery (class in pybman.query), 6

L

LangQuery (class in pybman.query), 7
LanguageConeController (class in pybman.rest), 8
languages_from_item() (in module pybman.extract), 4
list_from_level() (in module pybman.extract), 4
LocalData (class in pybman.local), 6
login() (pybman.rest.LoginRestController method), 8
LoginRestController (class in pybman.rest), 8
logout() (pybman.rest.LoginRestController method), 8

M

metadata() (in module pybman.extract), 4

O

organizations_from_item() (in module pybman.extract), 4
organizations_identifier_from_creator() (in module pybman.extract), 4
organizations_name_from_creator() (in module pybman.extract), 4
OrgUnitQuery (class in pybman.query), 7
OrgUnitRestController (class in pybman.rest), 8
ous_graph() (pybman.rest.PersonConeController method), 8

P

PersonConeController (class in pybman.rest), 8
persons_affiliation_from_creator() (in module pybman.extract), 4
persons_from_item() (in module pybman.extract), 4
persons_id_from_creator() (in module pybman.extract), 4
persons_identifier_from_creator() (in module pybman.extract), 4
persons_name_from_creator() (in module pybman.extract), 4
persons_organizations_from_creator() (in module pybman.extract), 4
PersQuery (class in pybman.query), 7
place_from_item() (in module pybman.extract), 4
post_request() (in module pybman.utils), 9
pubinfo_from_item() (in module pybman.extract), 5
publisher_from_item() (in module pybman.extract), 5
put_request() (in module pybman.utils), 9
pybman.client (module), 1
pybman.data (module), 1
pybman.extract (module), 3
pybman.inspector (module), 5
pybman.local (module), 6
pybman.query (module), 6
pybman.rest (module), 7
pybman.utils (module), 9

Q

Query (*class in pybman.query*), 7

R

read_csv_with_header() (*in module pybman.utils*), 9
read_json() (*in module pybman.utils*), 9
read_plain_clean() (*in module pybman.utils*), 9
release_item() (*pybman.rest.ItemRestController method*), 7
resolve_path() (*in module pybman.utils*), 9
RestController (*class in pybman.rest*), 8
revise_item() (*pybman.rest.ItemRestController method*), 7
role_from_creator() (*in module pybman.extract*), 5

S

scroll_items() (*pybman.rest.ItemRestController method*), 7
search_entity() (*pybman.rest.JournalConeController method*), 8
search_entity() (*pybman.rest.LanguageConeController method*), 8
search_entity() (*pybman.rest.PersonConeController method*), 8
search_items() (*pybman.rest.ItemRestController method*), 8
source_titles() (*in module pybman.extract*), 5
source_titles_genres() (*in module pybman.extract*), 5
sources() (*in module pybman.extract*), 5
sources_identifiers_from_item() (*in module pybman.extract*), 5
sources_persons_affiliations_from_item() (*in module pybman.extract*), 5
sources_persons_from_item() (*in module pybman.extract*), 5
sources_persons_id_from_item() (*in module pybman.extract*), 5
sources_titles_from_item() (*in module pybman.extract*), 5
sources_titles_genres_from_item() (*in module pybman.extract*), 5
store_data() (*pybman.local.LocalData method*), 6
submit_item() (*pybman.rest.ItemRestController method*), 8

T

title_from_item() (*in module pybman.extract*), 5
titles_from_records() (*in module pybman.extract*), 5
top_level() (*pybman.rest.OrgUnitRestController method*), 8
type_from_creator() (*in module pybman.extract*), 5

U

update_data() (*pybman.client.Client method*), 1
update_genre() (*pybman.inspector.Inspector method*), 6
update_item() (*pybman.rest.ItemRestController method*), 8
update_source_genre() (*pybman.inspector.Inspector method*), 6
url_exists() (*in module pybman.utils*), 9
url_exists2() (*in module pybman.utils*), 9

V

value_from_level() (*in module pybman.extract*), 5

W

withdraw_item() (*pybman.rest.ItemRestController method*), 8
write_csv() (*in module pybman.utils*), 9
write_json() (*in module pybman.utils*), 9
write_list() (*in module pybman.utils*), 9