

百度学术题录资源接入说明

说明

百度学术面向所有资源商提供学术资源对接服务，对于已发表的期刊论文，我们免费收录题录数据，收录后在用户检索相关论文时，则优先展现对应期刊官网来源的论文，提升期刊官网流量和品牌曝光度。同时我们也将根据您的需求定期提供基于期刊论文的用户大数据分析报告。

合作流程

1.登录百度学术资源平台（<https://xueshu.baidu.com/usercenter/resource>）提交合作申请。需提交的内容包括机构名称、机构类型、机构网址主域、机构资质等，我们会在一周内完成审核。

2.通过审核后，您即可按照下列“学术资源平台接入文档”进行数据提交和管理。

3.接入文档中如果涉及到 token 等信息，需要您点击下方链接并按照流程申请
<https://openauth.baidu.com/doc/prepare.html>

针对合作如有疑问，可以联系百度学术官方客服：2166226183

版本号	更新时间	更新内容
4.0	2019/8/12	保留可用的上传方式

1.新增数据

1.1 请求接口

https://openapi.baidu.com/rpc/2.0/xueshu-resource/paper/upload?access_token=xxxxxx

1.2 请求参数

使用 `content-type = application/json` 的方式发送请求，在 `url` 中拼接 `access_token` 的值

```
{
  "baidu_account": "机构注册使用的 baidu_account",
  "source_type": 1, // 资源方上传资源类型，必须。字段含义如下:1:pdf; 2:
馆藏; 3:ppv,4:图书 5:资源方提交普通题录
  "data": // 论文的主体数据，不同的类型对应不同的结构，具体说明如下：
}
```

1.2.1 Data 数据结构描述

请求参数	类型	必要	comments
basic	自定义子结构	required	论文相关基本信息
version	自定义子结构	required	论文来源相关的基本信息
ppv	自定义子结构	optional	论文售卖信息，ppv 类型资源必填
savelink	自定义子结构	optional	论文下载信息，pdf 类型资源必填
refer	自定义子结构	optional	论文引文信息，选填

根据 `source_type` 决定需要填充结构: - `refer`—引文信息对于所有的 `source_type` 均为选填字段 - 当 `source_type = 5` 代表资源方提交普通题录信息，只需要填充 `basic` 和 `version` 即可

- 当 `source_type = 1` 代表论文是 pdf（即可免费下载的论文全文）,除了 `basic` 和 `version` 之外， 还需要传递的参数为 `savelink`
- 当 `source_type = 3` 代表论文是 ppv（即付费下载论文）,除了 `basic` 和 `version` 之外,还需要传递的参数为 `ppv`
- 当 `source_type = 2` 代表论文是馆藏（即图书馆馆藏资源）,再普通题录基础上，填写 `version` 里面的 `lib` 字段
- 当 `source_type = 4` 代表资源方提交图书论文，只需要填充 `basic` 和 `version` 即可，但是出版信息只能填写图书子结构相关信息

1.2.2 Basic 结构说明

请求参数	类型	必要	comments
keywords	string	optional	关键字，多个关键字之间用;分隔
abstract	string	optional	论文摘要

authors	自定义子结构	optional	论文作者，可以为多个，例如： “authors”:[{“name”：“张三”，“affiliate”：“清华大学”，“mail”：“xxx@email.com”， “h_index”：2.3},{“name”：“李四”，“affiliate”：“清华大学”，“mail”：“xxx@email.com”， “h_index”：2.3}]
publish	自定义子结构	optional	论文出版信息
title	string	required	论文标题
doi	string	optional	论文 doi 信息
ch_lib_classification	string	optional	中图法分类
extra_data	string	optional	扩展字段，留备用（json encode 之后的 string）
level	string	optional	论文等级，0b0 为默认，0b1 为核心期刊，0b10 为 sci 索引，0b100 为 ei 索引

1.2.2.1 authors 字段结构说明

请求参数	类型	必要	comments
name	string	required	作者名字
affiliate	string	optional	作者机构
mail	string	optional	作者邮箱
h_index	double	optional	作者影响力因子

1.2.2.2 publish 字段结构说明

请求参数	类型	必要	comments
conference	自定义子结构	optional	会议类型论文的会议信息
degree	自定义子结构	optional	学位信息
patent	自定义子结构	optional	专利信息
journal	自定义子结构	optional	期刊信息
book	自定义子结构	optional	图书信息

paper_type	int	required	论文类型，0:未知;1:期刊;2:学位;3:会议;4:图书;5:专利;6:出版物，根据此类型填充上面对应的子结构
paper_published_time	string	optional	出版时间
paper_published_year	int	optional	出版年
publisher	string	optional	出版商
cn	string	optional	国内出版物刊号

出版信息主要是根据 paper_type 字段值填充对应类型的子结构，conference/degree/patent/journal/book,选择对应的子结构填写即可。

1.2.2.2.1 journal 结构

当 paper_type = 1 的时候，也就是论文类型为期刊(journal)

请求参数	类型	必要	comments
journal_issue	string	required	期刊期号
journal_volume_no	string	optional	期刊卷号
journal_page_no	string	optional	论文在期刊的页数
journal_name	string	optional	期刊名称
eissn	string	optional	电子出版物唯一编号
issn	string	optional	出版物唯一编号
is_prepublish	int	required	是否为预发布版本，1 代表为预发布版本，0 代表正式版本

1.2.2.2.2 degree 结构

当 paper_type = 2 的时候，也就是论文类型为学位, 填写 degree 子结构.

请求参数	类型	必要	comments
degree_name	string	required	学位名称
school	string	optional	所在学校
tutor	string 数组	optional	导师，可以为多个，例如“tutor”:["张三","李四"]
school	string	optional	所在学校
major	string	optional	专业
research_area	string	optional	研究方向
oral_defense_time	string	optional	答辩时间

fund	string 数组	optional	赞助基金，可以为多个 “fund”:[“xx”,“xxx”]
------	-----------	----------	-----------------------------------

1.2.2.2.3 conference 结构

当 paper_type = 3 的时候，也就是论文类型为会议, 填写 conference 子结构。

请求参数	类型	必要	comments
conference_name	string	required	会议名称
proceedings	string	optional	会议论文集
meeting_place	string	optional	会议地点
fund	string 数组	optional	赞助基金，可以为多个 “fund”:[“xx”,“xxx”]

1.2.2.2.4 book 结构

当 paper_type = 4 的时候，也就是论文类型为图书,填写 book 子结构。

请求参数	类型	必要	comments
book_name	string	required	图书名称
eisbn	string	optional	电子出版物唯一编号
isbn	string	optional	出版物唯一编号

1.2.2.2.5 patent 结构

当 paper_type = 5 的时候，也就是论文类型为专利, 填写 patent 子结构。

请求参数	类型	必要	comments
assignees	自定义子结构	optional	专利申请人，可以为多个 “assignees”:[{“name”:“张三”,“address”:“北京”},{“name”:“李四”,“address”:“北京”}]
pubid	string	optional	专利发表编号（专利发表编号和申请编号至少选一个填写）
apptime	string	optional	申请时间
appid	string	optional	申请编号
sovereignty	string 数组	optional	专利主权项，可以为多个 “sovereignty”:[“xxx”,“xxx”]
agent	string 数组	optional	专利代理人，可以为多个 “agent”:[“xxx”,“xxx”]

agency	string 数组	optional	专利代理机构, 可以为多个 “agency”:[“xxx”,“xxx”]
area_id	string	optional	国省代码
main_classification	string	optional	专利主分类号
patent_classification	string	optional	专利分类号
patent_priority	string 数组	optional	专利优先权 “patent_priority”:[“xxx”,“xxx”]
patent_status	string	optional	专利申请状态, 申请中/已发布
page_nums	int	optional	专利页数

1.2.2.2.5.1 assignees 子结构定义如下:

请求参数	类型	必要	comments
name	string	required	申请人名字
address	string	optional	申请人地址

1.2.3 version 结构说明如下:

请求参数	类型	必要	comments
url_anchor	string	optional	前端展示的 anchor 信息
lib	string	optional	馆藏简称, 如果 source_type=2,该字段必填
scholar_url	string	required	论文 url 地址, 带协议头

1.2.4 savelink 结构

当 source_type = 1 也就是说论文是 pdf,除了 basic 和 version 之外, 还需要传递的参数为 savelink。

请求参数	类型	必要	comments
save_url	string	required	pdf 地址
save_anchor	string	optional	pdf 链接的锚点
save_size	int	optional	pdf 文章的大小
save_type	int	required	是否免费下载, 1 表示免费; 2 表示登录下载

1.2.5 PPV 结构

当 source_type = 3 也就是说论文是 ppv, 除了 basic 和 version 之外,还需要传递的参数为 ppv。

请求参数	类型	必要	comments
individual_price	int	required	个人购买价格, 单位为“分”

individual_currency	string	required	个人购买单位,人民币美元等
institution_price	int	optional	机构购买价格, 单位为“分”
institution_currency	string	optional	机构购买单位,人民币美元等
sale_source	string	required	售卖方
article_id	string	required	文章 id

1.2.6 refer 结构

请求参数	类型	必要	comments
refertext	自定义结构	optional	引文

自定义结构 **refertext** 如下, 主体部分和 Basic 结构基本一致, 具体如下:

请求参数	类型	必要	comments
keywords	string	optional	关键字, 多个关键字之间用;分隔
abstract	string	optional	论文摘要
authors	自定义子结构	optional	论文作者, 可以为多个, 例如: “authors”:[{“name”:“张三”,“affiliate”:“清华大学”,“mail”:“xxx@email.com”, “h_index”:2.3},{“name”:“李四”,“affiliate”:“清华大学”,“mail”:“xxx@email.com”,“h_index”:2.3}]
publish	自定义子结构	optional	论文出版信息
title	string	required	论文标题
doi	string	optional	论文 doi 信息

1.2.7 Data 实例说明:

```
{
  "basic":{
    "doi":"10.1007/s00419-008-0233-7",
    "title":"24 周不同量钙磷饮食对尿毒症大鼠甲状旁腺结构、功能及死亡率的影响",
    "abstract":"高磷饮食是慢性肾衰竭(CRF)患者发生继发性甲状旁腺功能亢进(SHPT)及死亡率增加的重要原因。饮食钙水平可干扰肠道磷吸收与肾脏磷排泄。目前临床指南中多推荐给予 CRF 患者补充钙剂治疗,这对 CRF 患者高磷饮食相关 SHPT 有无影响?另外,虽然目前临床指南多推荐 CRF 患者进行低磷饮食以抑制 SHPT,但由此导致的饮食钙不足对低磷饮食所致 SHPT 的抑制有无重要影响?本研究目的在于:(1)研究补充饮食钙对高磷饮食所致 SHPT 及相关死亡率的影响及机制;(2)研究限制饮食钙水平对低磷饮食所致 SHPT 抑制的影响及机制。",
    "publish":{
      "paper_published_time":"2012 年 1 月",
```

```

    "paper_published_year":2009,
    "publisher":"北京**出版社",
    "conference":{
        "conference_name":"xxx 会议",
        "proceedings":"xx 会议论文集",
        "meeting_place":"xxx",
        "fund":[
            "***基金",
            "xxx 基金"
        ]
    },
    "paper_type":3,
    "cn":"*****"
},
"authors":[
    {
        "name":"袁鲁晓",
        "affiliate":"天津医科大学总医院",
        "mail":"luxiaoyuan@tjm.com",
        "h_index":4.3
    },
    {
        "name":"xxx",
        "affiliate":"天津医科大学总医院",
        "mail":"xx@tjm.com",
        "h_index":2.3
    }
],
"keywords":"尿毒症;继发性甲状旁腺功能亢进;钙磷饮食;细胞外调节磷酸激
酶;钙敏感受体;维生素 D 受体;功能机制",
"level":"0b0",
"ch_lib_classification":"R748",
"extra_data":""
},
"ppv":{
    "sale_source":"wanfang",
    "institution_currency":"USD",
    "institution_price":38,
    "individual_price":38,
    "individual_currency":"USD"
},
"version":{
    "source":"cnki.com.cn",
    "url_anchor":"知网",
    "partner_id":"cnki_1",
    "scholar_url":"http://cdmd.cnki.com.cn/Article/CDMD-10062-10124
97520.htm"
},
"savelink":{

```

```
"save_size":0,
"save_url":"http://www.doc88.com/p-1078030394955.html",
"save_anchor":"道客巴巴"
"save_type":1
}
"refer":[
{
  "refertext":{
    "doi":"10.1007/s00419-008-0233-7",
    "abstract":"xxxxxxxxxx",
    "title":"24 周不同量钙磷饮食对尿毒症大鼠甲状旁腺结构、功能及
死亡率的影响",
    "publish":{
      "paper_published_time":"2012 年 1 月",
      "paper_published_year":2009,
      "publisher":"北京**出版社",
      "conference":{
        "conference_name":"xxx 会议",
        "proceedings":"xx 会议论文集",
        "meeting_place":"xxx",
        "fund":[
          "***基金",
          "xxx 基金"
        ]
      }
    },
    "paper_type":3,
  },
  "authors":[
    {
      "name":"袁鲁晓",
      "affiliate":"天津医科大学总医院",
      "mail":"luxiaoyuan@tjm.com",
      "h_index":4.3
    },
    {
      "name":"xxx",
      "affiliate":"天津医科大学总医院",
      "mail":"xx@tjm.com",
      "h_index":2.3
    }
  ],
  "keywords":"尿毒症;继发性甲状旁腺功能亢进;钙磷饮食",
},
{
  "refertext":{
    "title":"中国房地产上市公司治理结构、多元化与企业绩效研究",
    "authors": [{"name": "高帅"}]
  }
}
```

```
    }
  ]
}
```

1.3 接口返回

```
{
  "msg": "处理结果提示信息",
  "status": 0, // 0 表示提交信息成功, 非 0 表示不成功
  "data": []
}
```

1.4demo 实例

```
# -*- coding: utf-8 -*-
```

```
import requests
import json
import logging
```

```
NEW_UPLOAD_TEST_URL = "https://openapi.baidu.com/rpc/2.0/xueshu-resource/paper/upload" # noqa
```

```
OPENAPI_TEST_URL = "http://openapi.baidu.com/oauth/2.0/token?grant_type=client_credentials" # noqa
```

```
TEST_CLIENT_ID = "97ZW6jYcN68TozvSsFHP80Uks6xmzwTQ"
TEST_CLIENT_SECRET = "2s3FGZvBXx32dZRgMKpXl8Lo7IoGZAzf"
```

```
new_paper_params = {
    "baidu_account": "18163316537",
    "source_type": 1,
    "data": {
        "basic": {
            "doi": "10.1007/s00419-008-0233-7",
            "title": "24 周不同量钙磷饮食对尿毒症大鼠甲状旁腺结构、功能及死亡率的影响",
            "abstract": "高磷饮食是慢性肾衰竭(CRF)患者发生继发性甲状旁腺功能亢进(SHPT)及死亡率增加的重要原因。饮食钙水平可干扰肠道磷吸收与肾脏磷排泄。目前临床指南中多推荐给予 CRF 患者补充钙剂治疗,这对 CRF 患者高磷饮食相关 SHPT 有无影响?另外,虽然目前临床指南多推荐 CRF 患者进行低磷饮食以抑制 SHPT,但由此导致的饮食钙不足对低磷饮食所致 SHPT 的抑制有无重要影响?本研究目的在于:(1)研究补充饮食钙对高磷饮食所致 SHPT 及相关死亡率的影响及机制;(2)研究限制饮食钙水平对低磷饮食所致 SHPT 抑制的影响及机制。",
            "publish": {
                "paper_published_time": "2012 年 1 月",
                "paper_published_year": 2009,
                "publisher": "北京**出版社",
```

```
    "conference":{
      "conference_name":"xxx 会议",
      "proceedings":"xx 会议论文集",
      "meeting_place":"xxx",
      "fund":[
        "****基金",
        "xxx 基金"
      ]
    },
    "paper_type":3,
    "cn":"*****"
  },
  "authors":[
    {
      "name":"袁鲁晓",
      "affiliate":"天津医科大学总医院",
      "mail":"luxiaoyuan@tjm.com",
      "h_index":4.3
    },
    {
      "name":"xxx",
      "affiliate":"天津医科大学总医院",
      "mail":"xx@tjm.com",
      "h_index":2.3
    }
  ],
  "keywords":"尿毒症;继发性甲状旁腺功能亢进;钙磷饮食;细胞外调节磷酸激酶;钙敏感受体;维生素 D 受体;功能机制",
  "level":"0b0",
  "ch_lib_classification":"R748",
  "extra_data":""
},
"ppv":{
  "sale_source":"wanfang",
  "institution_currency":"USD",
  "institution_price":38,
  "individual_price":38,
  "individual_currency":"USD"
},
"version":{
  "source":"cnki.com.cn",
  "url_anchor":"知网",
  "partner_id":"cnki_1",
  "scholar_url":"http://cdmd.cnki.com.cn/Article/CDMD-10062-1012497520.htm"
},
"savelink":{
  "save_size":0,
  "save_url":"http://www.doc88.com/p-1078030394955.html",
```

```

        "save_anchor": "道客巴巴"
        "save_type": 1
    }
}

def get_access_token():
    payload = {
        "client_id": TEST_CLIENT_ID,
        "client_secret": TEST_CLIENT_SECRET
    }
    try:
        r = requests.get(OPENAPI_TEST_URL, params=payload)
        return json.loads(r.content)['access_token']
    except Exception as e:
        logging.error("get access_token failed:{}".format(repr(e)))

def new_upload_papaer():
    token = get_access_token()

    # 注意在 url 中拼接 access_token 的值
    url = NEW_UPLOAD_TEST_URL + "?access_token=" + token

    try:
        response = requests.post(url, json=new_paper_params)
        return json.loads(response.content)
    except Exception as e:
        logging.error("upload failed:{}".format(repr(e)))

if __name__ == '__main__':
    res = new_upload_papaer()
    print res

```

2. 更新数据

如果需要更新已上传的资源，则更新信息后，直接重新上传即可