

Status of FS SciCat for DOIs

screenshot demo of status

work-in-progress

Regina Hinzmann, IT-InFa

2025-11-24

SciCat Technical meeting

Content overview

1. Status Pilot scicat for beamlines



SciCat at DESY for general DOI minting

Status of pilot in more detail (from week 44)

External service code (*publisher*) is now in gitlab and runs as a container in the cluster.

Part that is done:

- We assume a scicat dataset of a certain structure, as agreed by FS; Linus and Nicolas (Aug. 2025).
- it copies all associated datafiles and attaches them to the original dataset.
- it creates the copy on fsdata.
- A test job was successfully created and submitted that talks to our service.

What's on the plan?

- Trigger custom job from frontend.
- Complete workflow: add email notification with the new url of the user in fsdata (add another action to the job)
- Make code robust.
- Test with other datasets.
- Handover.

`scicat.desy.de`

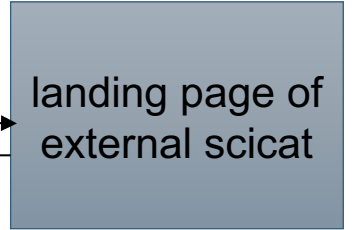


external
service

`fsdata.desy.de`



`doi.desy.de`



DEMO

Status summary

initial dataset from user on
scicat-dev

scicat-dev.desy.de

Cache u SciCatToolsSCSciCat...OpenAPIdict.ccTimer O...erstellenSTOPPUHR onlineSCTHome - XWikiSciCat_ESY ChatSRI 08...is 08-30Configu...ub DocsGitHub...ns inputsystems...GitLabOperato...entationrawBookmarks menu

SciCatDESY

Scicat

HelpAboutRegina Kwee-Hinzmann

Datasets / scicat/Obf11a0b-9e5c-4752-8691-94d78f6e59bd /

DetailsDatafilesRelated DatasetsAttachmentsLifecycle

General Information

Dataset name

Combined Dataset

PID

scicat/Obf11a0b-9e5c-4752-8691-94d78f6e59bd

Creation time

2025-08-27 18:40

Type

derived

Keywords

-

Contact Information

Contact email

hayen@physik.uni-kiel.de

Files Information

Related Documents

Proposal IDs

-

Instrument IDs

-

Sample IDs

-

Input datasets

scicat/BGgly_889861 scicat/528e19dc23

Scientific Metadata

Search

X

initial dataset from user on
scicat-dev

Page 3

Status summary

fsdata.desy.de/explorer#/datasets v4/DatasetsV4Controller_findById_v4

Sign in

Parameters

Name	Description
pid * required string <small>(path)</small>	Id of the dataset to return
include array<string> <small>(query)</small>	<div><div>-- instruments proposals origdatablocks</div></div>

Execute

Clear

Responses

Curl

```
curl -X 'GET' \
'https://fsdata.desy.de/api/v4/datasets/scicat%2F0bf11a0b-9e5c-4752-8691-94d78f6e59bd' \
-H 'accept: application/json' \
-H 'Authorization: Bearer eyJhbGciOiJIUzI1NiIsInR5cCI6ImlzLWZjbnNkZmZjbnNkZmZjbnNkZmZlZDQiLCJ1c2VybmFtZSI6ImNjaW9hdEkiLCJwcm9udGFkbGkiOiJzY2ljYXRhZCZpbkZkZSIsImR1iwiYXV0eQ'
```

Request URL

```
https://fsdata.desy.de/api/v4/datasets/scicat%2F0bf11a0b-9e5c-4752-8691-94d78f6e59bd
```

Server response

Code

Details

404
Undocumented

Error: response status is 404

Response body

```
{
  "message": "dataset: null not found",
  "error": "Not Found",
  "statusCode": 404
}
```

Download

Response headers

```
access-control-allow-origin: *
content-length: 74
content-type: application/json; charset=utf-8
date: Sat, 22 Nov 2025 15:12:12 GMT
etag: W/"4a-jtY6+DzmM0HB8XwcMgOE4XYIrZI"
x-firefox-spdly: h2
x-powered-by: Express
```


DEMO

Status summary

job talks to publisher

```
scicat-publisher — python - fastapi dev app/main.py — 172x44
...x0 NVMM_INC=/Users/regina/nvm/versions/node/v20.17.0/include/node
~/work/scicat-publisher — python - fastapi dev app/main.py
-- -bash
/Users/regina/work/scicat-publisher — -bash

1eHAI0jE3NzQwOdG0MjB9.Pg0g-AglfgUSjSL_Q1LAUYASgrDtaoQU207w2gxtoeE\r\n\r\n'
reply: 'HTTP/1.1 200 OK\r\n'
header: Access-Control-Allow-Origin: *
header: Content-Length: 4276
header: Content-Type: application/json; charset=utf-8
header: Date: Sat, 22 Nov 2025 15:12:46 GMT
header: Etag: W/"10b4-8wEo9BJmmPgWbhBHcSvZaB0dzSw"
header: X-Powered-By: Express
https://fsdata.desy.de:443 "GET /api/v3/datasets/scicat%2F0bf11a0b-9e5c-4752-8691-94d78f6e59bd/origdatablocks HTTP/1.1" 200 4276
2025-11-22 16:12:46,298 - urllib3.connectionpool - DEBUG - https://fsdata.desy.de:443 "GET /api/v3/datasets/scicat%2F0bf11a0b-9e5c-4752-8691-94d78f6e59bd/origdatablocks HTT
P/1.1" 200 4276
2025-11-22 16:12:46,298 - app.main - DEBUG - Got orig_datablocks: [OrigDatablock(created_by='scicatAdmin', updated_by='scicatAdmin', created_at=datetime.datetime(2025, 10,
9, 9, 2, 46, 702000, tzinfo=TzInfo(UTC)), updated_at=datetime.datetime(2025, 10, 9, 9, 2, 46, 702000, tzinfo=TzInfo(UTC)), owner_group='it-scicat', access_groups=['proposal
Ingestor'], instrument_group=None, is_published=False, id='dcc274e5-61d8-44cf-94ee-f1c007d88400', dataset_id='scicat/0bf11a0b-9e5c-4752-8691-94d78f6e59bd', size=0, chk_alg=
None, data_file_list=[DataFile(path='dppc_20degc_00210.fio', size=3968, time=datetime.datetime(2025, 9, 18, 8, 22, 3, 188000, tzinfo=TzInfo(UTC)), chk='string', uid='p08use
r', gid='irc', perm='-rw-r--r-', type='raw'), DataFile(path='dppc_20degc_00210/eiger4m/dppc_20degc_00210_master.h5', size=36065828, time=datetime.datetime(2025, 9, 18, 8, 2
2, 3, 188000, tzinfo=TzInfo(UTC)), chk='string', uid='p08user', gid='irc', perm='-rw-r--r-', type='raw'), DataFile(path='dppc_20degc_00210/eiger4m/dppc_20degc_00210_data_00
0001.h5', size=3603524, time=datetime.datetime(2025, 9, 18, 8, 22, 3, 188000, tzinfo=TzInfo(UTC)), chk='string', uid='p08user', gid='irc', perm='-rw-r--r-', type='raw'), Da
taFile(path='/some/nice/fiofile.fio', size=0, time=datetime.datetime(2025, 9, 18, 13, 9, 11, 360000, tzinfo=TzInfo(UTC)), chk='100', uid='p08user', gid='irc', perm='-rw-r--
r--', type='derived'), DataFile(path='he/is/back/halleluja/', size=1120, time=datetime.datetime(2025, 9, 23, 9, 0, 23, 193000, tzinfo=TzInfo(UTC)), chk=None, uid=None, gid
=None, perm=None, type='raw'), DataFile(path='dppc_20degc_00210.fio', size=3968, time=datetime.datetime(2025, 9, 18, 8, 22, 3, 188000, tzinfo=TzInfo(UTC)), chk=None, uid=No
ne, gid=None, perm=None, type=None), DataFile(path='dppc_20degc_00210.fio', size=3968, time=datetime.datetime(2025, 9, 18, 8, 22, 3, 188000, tzinfo=TzInfo(UTC)), chk=None,
uid=None, gid=None, perm=None, type=None), DataFile(path='/some/nice/fiofile.fio', size=0, time=datetime.datetime(2025, 9, 18, 13, 9, 11, 360000, tzinfo=TzInfo(UTC)), chk=N
one, uid=None, gid=None, perm=None, type=None), DataFile(path='/some/nice/fiofile.fio', size=0, time=datetime.datetime(2025, 9, 18, 13, 9, 11, 360000, tzinfo=TzInfo(UTC)),
chk=None, uid=None, gid=None, perm=None, type=None), DataFile(path='dppc_20degc_00210.fio', size=3968, time=datetime.datetime(2025, 9, 18, 8, 22, 3, 188000, tzinfo=TzInfo(U
TC)), chk=None, uid=None, gid=None, perm=None, type=None), DataFile(path='dppc_20degc_00210.fio', size=3968, time=datetime.datetime(2025, 9, 18, 8, 22, 3, 188000, tzinfo=Tz
Info(UTC)), chk=None, uid=None, gid=None, perm=None, type=None), DataFile(path='/some/nice/fiofile.fio', size=0, time=datetime.datetime(2025, 9, 18, 13, 9, 11, 360000, tzinfo=
TzInfo(UTC)), chk=None, uid=None, gid=None, perm=None, type=None), DataFile(path='/some/nice/fiofile.fio', size=0, time=datetime.datetime(2025, 9, 18, 13, 9, 11, 360000, tzinfo=
TzInfo(UTC)), chk=None, uid=None, gid=None, perm=None, type=None), DataFile(path='dppc_20degc_00210.fio', size=3968, time=datetime.datetime(2025, 9, 18, 8, 22, 3, 188000, tzinfo=
TzInfo(UTC)), chk=None, uid=None, gid=None, perm=None, type=None), DataFile(path='dppc_20degc_00210.fio', size=3968, time=datetime.datetime(2025, 9, 18, 8, 22, 3, 188000, tzinfo=
TzInfo(UTC)), chk=None, uid=None, gid=None, perm=None, type=None), DataFile(path='/some/nice/fiofile.fio', size=0, time=datetime.datetime(2025, 9, 18, 13, 9, 11, 360000, tzinfo=
TzInfo(UTC)), chk=None, uid=None, gid=None, perm=None, type=None), DataFile(path='dppc_20degc_00210.fio', size=3968, time=datetime.datetime(2025, 9, 18, 8, 22, 3, 188000, tzinfo=
TzInfo(UTC)), chk=None, uid=None, gid=None, perm=None, type=None), DataFile(path='/some/nice/fiofile.fio', size=0, time=datetime.datetime(2025, 9, 18, 13, 9, 11, 360000, tzinfo=
TzInfo(UTC)), chk=None, uid=None, gid=None, perm=None, type=None), DataFile(path='/this/is/a/new/path/to/fiofile1.fio', size=17000, time=datetime.datetime(2025, 10, 6, 10, 16, 7, 604000, tzinfo=TzInfo(UTC)), chk=None, uid=None, gid=None, perm=None, type=None), DataFile(path='/this/is/a/new/p
ath/to/fiofile2.fio', size=17000, time=datetime.datetime(2025, 10, 6, 10, 16, 7, 604000, tzinfo=TzInfo(UTC)), chk=None, uid=None, gid=None, perm=None, type=None), DataFile(
path='/this/is/a/new/path/to/fiofile3.fio', size=17000, time=datetime.datetime(2025, 10, 6, 10, 16, 7, 604000, tzinfo=TzInfo(UTC)), chk=None, uid=None, gid=None, perm=None,
type=None)]]
2025-11-22 16:12:46,298 - app.main - DEBUG - Already ODBs attached. Passing
INFO 127.0.0.1:60445 - "POST /datasetpid/ HTTP/1.1" 200
```


DEMO

Status summary

publisher deals with the request: gets dataset from scicat-dev, modifies it and creates a copy on fsdata. Sends email notification to user.

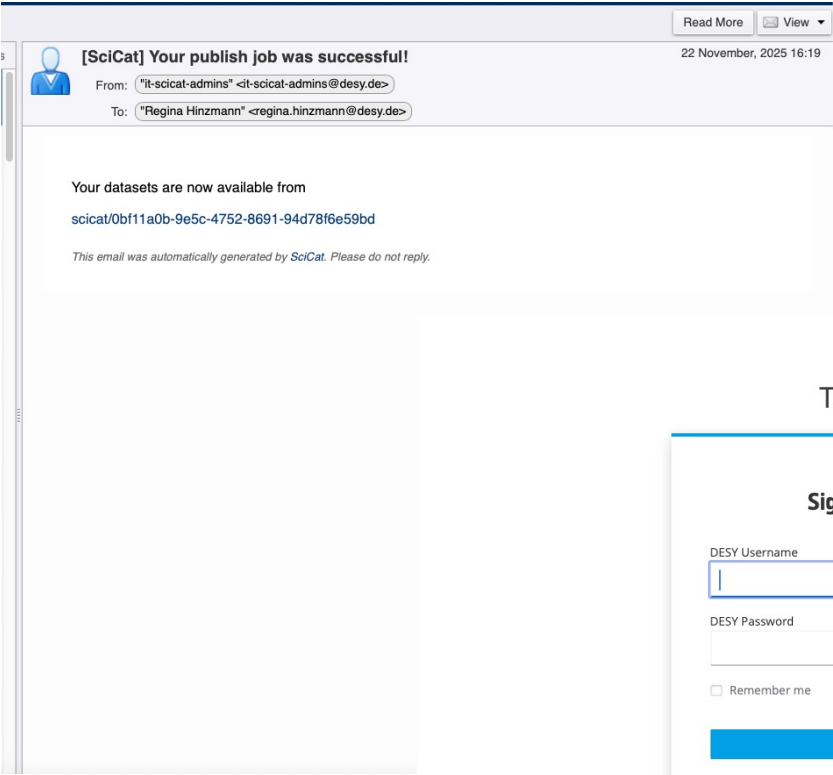
```
scicat-backend-next — node • npm run start:dev SHELL=jopt/homebrew/bin/bash XPC_FLAGS=0x0 NVM_INC=/Users/regina/.nvm/versions/node/v20.17.0/include/node — 172x44
...x0 NVM_INC=/Users/regina/.nvm/versions/node/v20.17.0/include/node  ~/work/scicat-publisher — python • fastapi dev app/main.py  -- -bash  /Users/regina/work/scicat-publisher — -bash

"MAX_FILE_UPLOAD_SIZE": "16mb",
"FUNCTIONAL_ACCOUNTS_FILE": "functionalAccounts.json",
"JOB_CONFIGURATION_FILE": "jobConfig.yaml",
"JOB_DEFAULT_STATUS_CODE": "jobSubmitted",
"JOB_DEFAULT_STATUS_MESSAGE": "Job submitted.",
"PUBLISHER_API_KEY": "273051a57408487de13def44c2ee291c64b296676e3b8fd9a208dd57420923ea",
"SCICAT_TOKEN": "eyJhbGciOiJIUzI1NiIsInR5cCI6IkpXVCJ9.eyJfaWQiOiI2OTU0NTU0MDgyMTg0YTk2ZThhZjZkZTYiLCJ1c2VybmFtZSI6ImFkbWluIiwiaWZ1haWwiOiJpdC1zY21jYXQtYWRtaW5zQGRIc3kuZGUlLCJhdXRoU3RyYXRlZ3kiOiJsb2NhbiIsIml9fdiI6MCwiaWQiOiI2OTU0NTU0MDgyMTg0YTk2ZThhZjZkZTYiLCJpYXQiOiE3NjM2MjQwNDksImV4cCI6MTc2MzY2MDA0OX0.TNkKPwLDaj9Ruad0Js79T0aDFEhsuYlW233NyfeJHBU",
"FSDATA_TOKEN": "eyJhbGciOiJIUzI1NiIsInR5cCI6IkpXVCJ9.eyJfaWQiOiI2OTU0NTU0MDgyMTg0YTk2ZThhZjZkZTYiLCJ1c2VybmFtZSI6ImFkbWluIiwiaWZ1haWwiOiJpdC1zY21jYXQtYWRtaW5zQGRIc3kuZGUlLCJhdXRoU3RyYXRlZ3kiOiJsb2NhbiIsIml9fdiI6MCwiaWQiOiI2OTU0NTU0MDgyMTg0YTk2ZThhZjZkZTYiLCJpYXQiOiE3NjM2MjQwNDksImV4cCI6MTc2MzY2MDA0OX0.TNkKPwLDaj9Ruad0Js79T0aDFEhsuYlW233NyfeJHBU",
  },
  "job": {
    "createdBy": "admin",
    "updatedBy": "admin",
    "createdAt": "2025-11-22T15:12:46.077Z",
    "updatedAt": "2025-11-22T15:12:46.077Z",
    "ownerGroup": "it-scicat",
    "accessGroups": [],
    "isPublished": false,
    "_id": "b44b6a6a-70e1-4ef1-a403-d5aaaa70de27",
    "id": "b44b6a6a-70e1-4ef1-a403-d5aaaa70de27",
    "ownerUser": "admin",
    "type": "publish",
    "statusCode": "jobSubmitted",
    "statusMessage": "Job submitted.",
    "jobParams": {
      "datasetList": [
        {
          "pid": "scicat/0bf11a0b-9e5c-4752-8691-94d78f6e59bd",
          "files": []
        }
      ]
    },
    "contactEmail": "regina.hinzmann@desy.de",
    "configVersion": "v1.0 2024-03-01 6f3f38",
    "jobResultObject": {},
    "__v": 0
  }
}
[Nest] 104 - 11/22/2025, 4:12:46 PM LOG [EmailJobAction] (Job b44b6a6a-70e1-4ef1-a403-d5aaaa70de27) Performing EmailJobAction
[Nest] 104 - 11/22/2025, 4:12:46 PM LOG [Utils.sendMail] Sending email to: regina.hinzmann@desy.de
```

DEMO

Status summary

email contains link to



TEST REALM

English ▾

Sign in to your account

DES Y Username

DES Y Password

☐ Remember me

Sign In

External users & external collaborations

Helmholtz AAI DEV

HELMHOLTZ

RESEARCH FOR
GRAND CHALLENGES

Support | Data Privacy Policy | Imprint | Contact
© 2023, Deutsches Elektronen-Synchrotron DES Y

DEMO

Status summary

...the dataset on fsdata

fsdata.desy.de

SciCat

Help About

Regina Kwee-Hinzmann

Datasets / scicat/0bf11a0b-9e5c-4752-8691-94d78f6e59bd /

Details Datafiles Related Datasets Attachments Lifecycle

General Information

Dataset Name

Combined Dataset

PID

scicat/0bf11a0b-9e5c-4752-8691-94d78f6e59bd

Creation Time

2025-08-27 18:40

Type

derived

Keywords

-

Contact Information

Contact email

hayen@physik.uni-kiel.de

Files Information

Related Documents

Proposal IDs

-

Instrument IDs

-

Sample IDs

-

Input datasets

["scicat/BGgJly_889861","scicat/528e19dc23"]

Scientific Metadata

Search

X

DAΦNE

DESY

Page 9

Outlook

Status summary

Our tasks

- enable job submission from scicat instead of scicat-dev
- adjust sender email. DONE
- attach jobsubmission to frontend button once available (see below)
- catch error of missing type? → Usually is fulfilled by initial dataset (see below).

Other people's input:

- wait for real dataset from FS to exercise full chain (requested)
- wait for Max to implement buttons from detailed dataset view.

Handover as planned: this year (date to be announced).