

Status of FS SciCat for DOIs

screenshot demo of status

Regina Hinzmann, IT-InFa
2025-11-24
SciCat Technical meeting

work-in-progress

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.

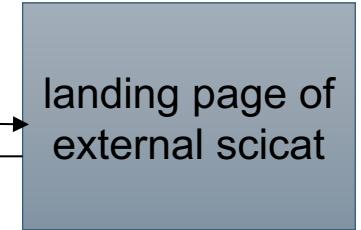
`scicat.desy.de`



`fsdata.desy.de`



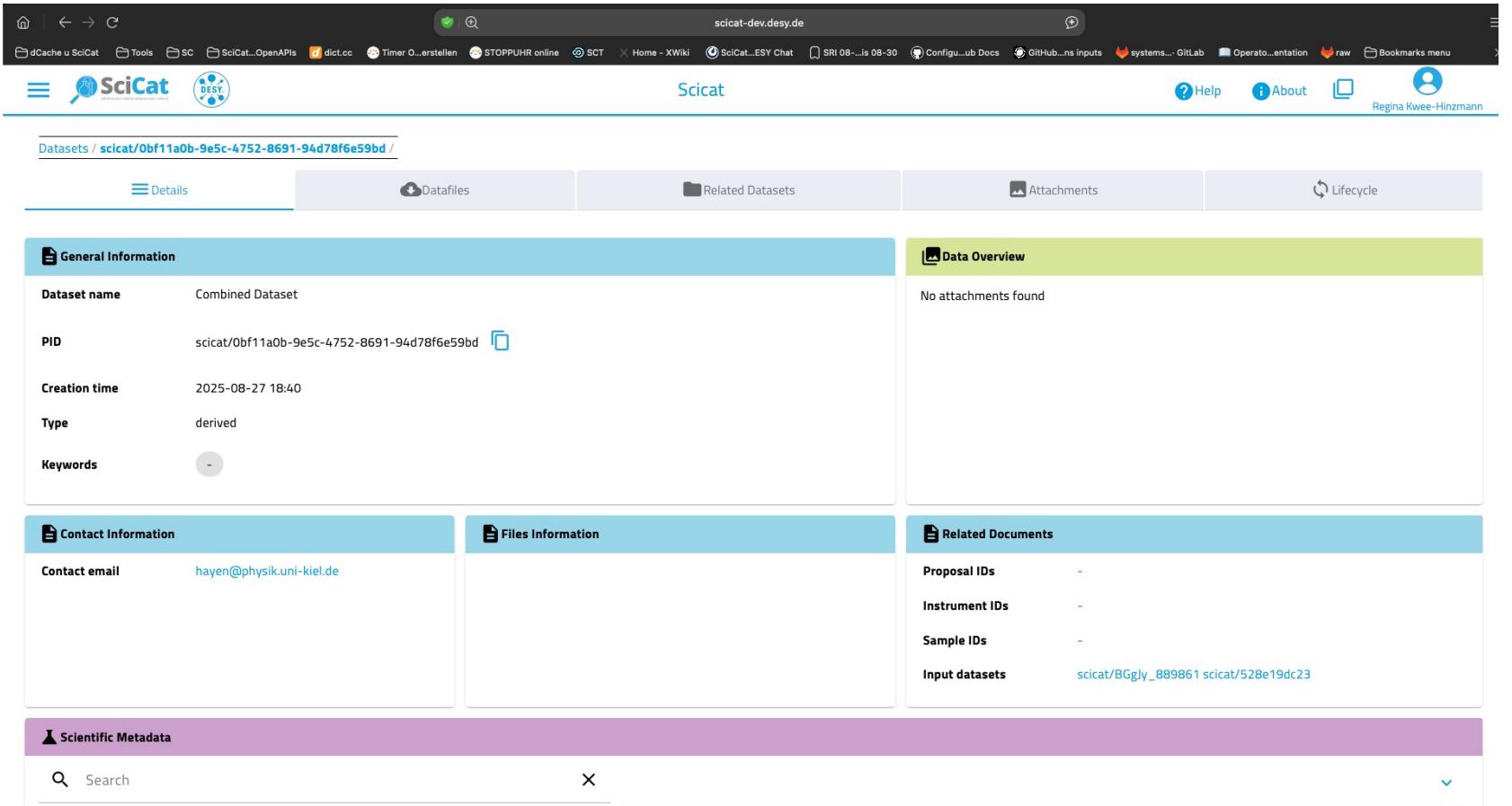
`doi.desy.de`



DEMO

Status summary

initial dataset from user on
scicat-dev



The screenshot shows a web browser window for the URL scicat-dev.desy.de/datasets/scicat/0bf11a0b-9e5c-4752-8691-94d78f6e59bd/. The page is titled "Scicat" and features a navigation bar with links to dCache, SciCat, Tools, SC, SciCat_OpenAPIs, dict.cc, Timer, STOPPUHR online, SCT, Home - XWiki, SciCat_ESY Chat, SRI 08-is 08-30, Configurable Docs, GitHub, systems, GitLab, Operator, raw, and Bookmarks menu. The user is logged in as "Regina Kwee-Hinzmann".

The main content area displays the details of a dataset. The "Details" tab is active, showing the following information:

- Dataset name:** Combined Dataset
- PID:** scicat/0bf11a0b-9e5c-4752-8691-94d78f6e59bd
- Creation time:** 2025-08-27 18:40
- Type:** derived
- Keywords:** -

Other tabs include "Datafiles", "Related Datasets", "Attachments", and "Lifecycle".

The "Data Overview" section indicates "No attachments found".

The "Contact Information" section shows the contact email hayen@physik.uni-kiel.de.

The "Files Information" section is currently empty.

The "Related Documents" section lists:

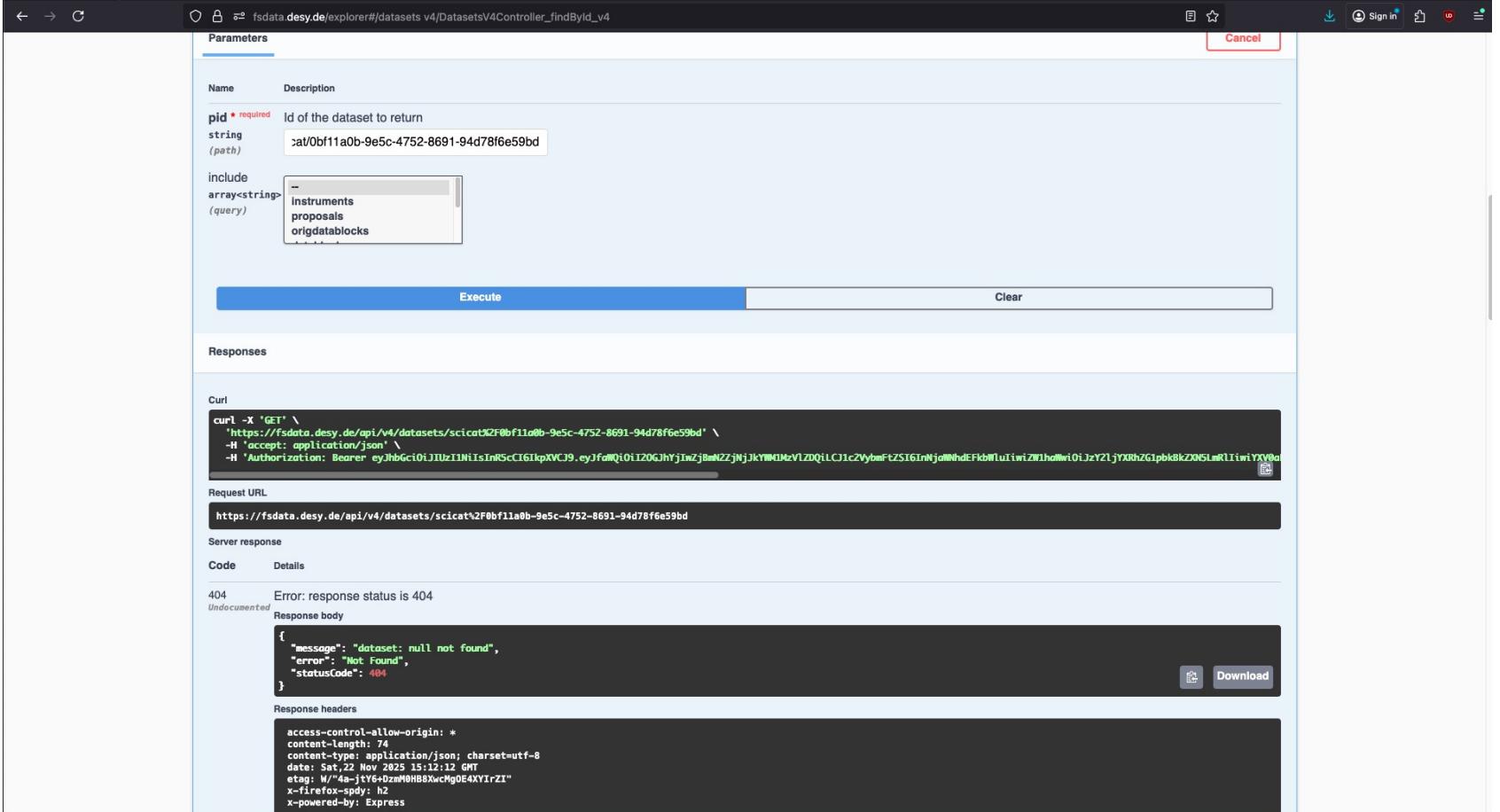
- Proposal IDs:** -
- Instrument IDs:** -
- Sample IDs:** -
- Input datasets:** [scicat/BGgly_889861](https://scicat-dev.desy.de/datasets/scicat/BGgly_889861) [scicat/528e19dc23](https://scicat-dev.desy.de/datasets/scicat/528e19dc23)

The "Scientific Metadata" section contains a search bar with the placeholder "Search" and a clear button "X".

DEMO

Status summary

fsdata.desy.de is empty



The screenshot shows the fsdata.desy.de API explorer interface. The URL in the address bar is `fsdata.desy.de/explorer#/datasets v4/DatasetsV4Controller_findById_v4`. The main area is titled "Parameters" and contains two entries:

- pid** required **Description**: Id of the dataset to return
- string** (path) `>at/0bf11a0b-9e5c-4752-8691-94d78f6e59bd`

Below the parameters is a "Responses" section. The "Curl" section shows a command to make a GET request to the specified URL. The "Request URL" is `https://fsdata.desy.de/api/v4/datasets/scicat%2F0bf11a0b-9e5c-4752-8691-94d78f6e59bd`. The "Server response" section shows a 404 error with the following details:

- Code**: 404
- Details**: Error: response status is 404
- Response body**:

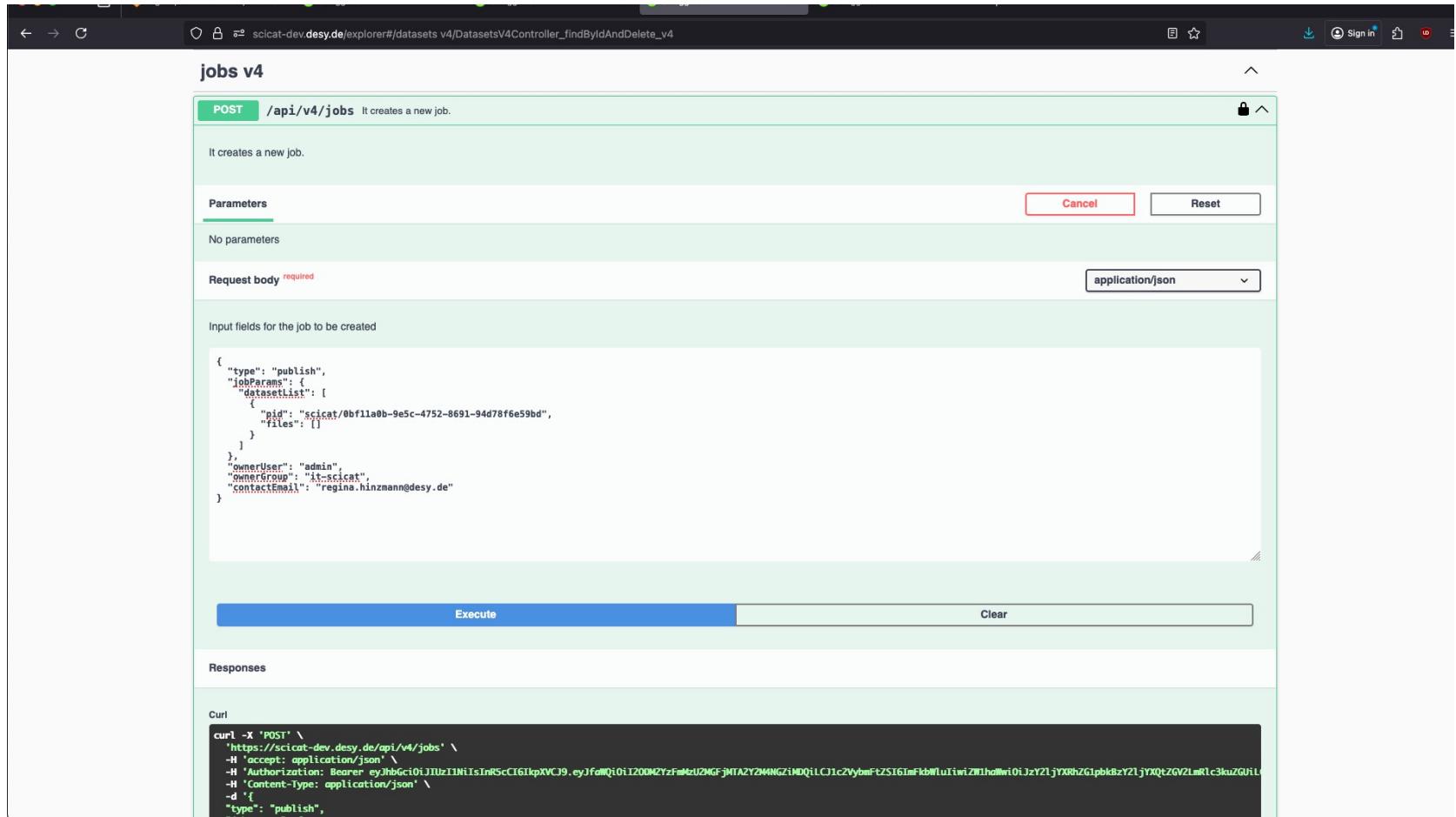
```
{  "message": "dataset: null not found",  "error": "Not Found",  "statusCode": 404}
```
- Response headers**:

```
access-control-allow-origin: *  content-length: 74  content-type: application/json; charset=utf-8  date: Sat, 22 Nov 2025 15:12:12 GMT  etag: M/4a-jtY6+0zmM0HB8XwcMgOE4XYIrZI  x-firebase-spdy: h2  x-powered-by: Express
```

DEMO

Status summary

submit job from scicat-dev



The screenshot shows the 'scicat-dev.desy.de/explorer#/datasets v4/DatasetsV4Controller_findByIdAndDelete_v4' interface. The main section is titled 'jobs v4' with a 'POST /api/v4/jobs' endpoint. The description states 'It creates a new job.' Under 'Parameters', it says 'No parameters'. The 'Request body' is marked as 'required' and contains the following JSON:

```
{
  "type": "publish",
  "jobParams": {
    "datasetList": [
      {
        "pid": "scicat/0bf11a0b-9e5c-4752-8691-94d78f6e59bd",
        "files": []
      }
    ],
    "ownerUser": "admin",
    "ownerGroup": "it-scicat",
    "contactEmail": "regina.hinzmann@desy.de"
  }
}
```

Below the request body, there are 'Execute' and 'Clear' buttons. The 'Responses' section includes a 'Curl' command:

```
curl -X 'POST' \
  'https://scicat-dev.desy.de/api/v4/jobs' \
  -H 'accept: application/json' \
  -H 'Authorization: Bearer eyJhbGciOiJIUzI1NiIsInR5cCI6IkpXVCJ9.eyJfdWQiOiJ2Z0M2YzFmMzU2NGFjMTAY2M4NGZiMDQjLCJ1c2Vybmt2S1G1mFkbWluIiwidhdhIjoiJzY2l1jYX8hZG1pbkbzY2l1jYXQzZG12LmRlc3kuZGUiL' \
  -H 'Content-Type: application/json' \
  -d '{
    "type": "publish",
    "jobParams": {
      "datasetList": [
        {
          "pid": "scicat/0bf11a0b-9e5c-4752-8691-94d78f6e59bd",
          "files": []
        }
      ],
      "ownerUser": "admin",
      "ownerGroup": "it-scicat",
      "contactEmail": "regina.hinzmann@desy.de"
    }
  }'
```

DEMO

Status summary

job talks to publisher

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.

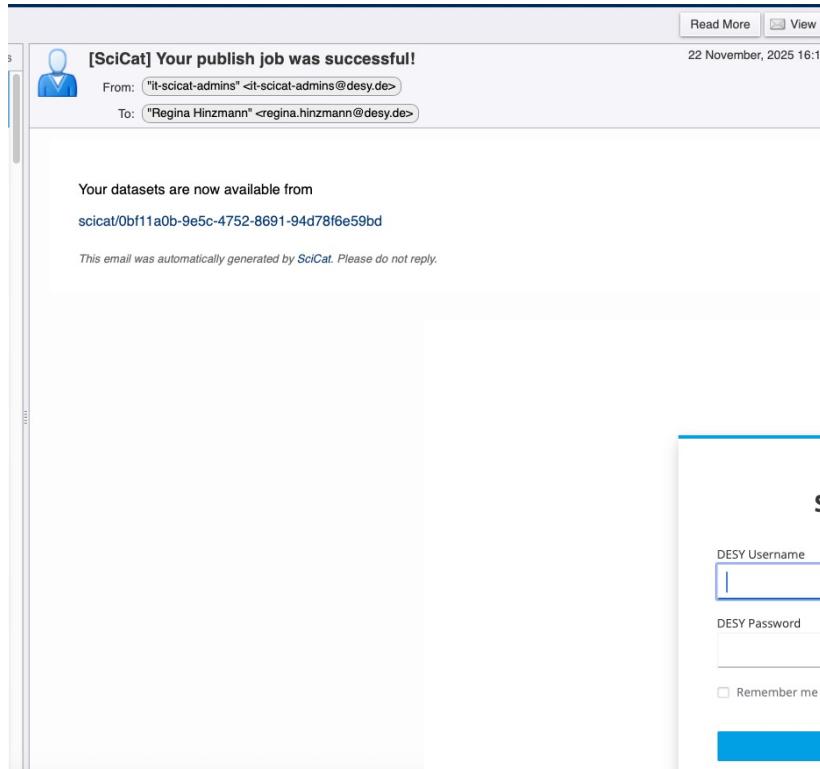
```
  ...x0 NVM_INC=/Users/regina/.nvm/versions/node/v20.17.0/include/node
  .../work/scicat-publisher -- python -m fastapi dev app/main.py
  ... ~ -- -bash
  ... /Users/regina/work/scicat-publisher -- bash

  "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": "eyJhbGciOiJIUzI1NiIsInR5cCI6IkpXVCJ9eyJfaWQioiI20TE0NTE1MDgyMTg0YTk2ZThhZjk0ZTYiLCJ1c2VybmtZSI6ImFkbWluIiwiZW1haWwiOiJpdC1zY2ljYXQtYWRtaW5zQRlc3kUZGUiLCJhdXroU3RyYXR1ZkI0iJsb2NhbCisI19fdiI6MCwiaWQiOii20TE0NTE1MDgyMTg0YTk2ZThhZjk0ZTYiLCJpYXQiOjE3NjM2MjQwNDksImV4cCI6MTc2MzY2MDA0X0. TNKPkWLDaj9RuatOjs79T0aDfEhsuY1w233NyeJHU",
  "FSDATA_TOKEN": "eyJhbGciOiJIUzI1NiIsInR5cCI6IkpXVCJ9eyJfaWQioiI20GJhYjIwZjBmN2ZjNjJKYWM1MzVlZDQiLCJ1c2VybmtZSI6InNjaWNhdEFkbWluIiwiZW1haWwiOiJzY2ljYXRsZG1pbkBkZXN5LmR1IiwiYXV0aFN0cmF0Zwd5IjoibG9jYWw1LCJfX3Yi0jAsImlkIjoiNjhiYWIyMGYwZjdmYzYyZGFjNTM1ZWQ0IiwiWF0IjoxNzYwOTU2Nzg0LCJleHaiOjE3NjIxNjYzODR9. KP9HG0PZ-xMLbFeCPaqrKQwE2V8Wa-0sYabYxQFO04"
},
"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

DESY Username

DESY Password

Remember me

Sign In

External users & external collaborations

 Helmholtz AAI DEV

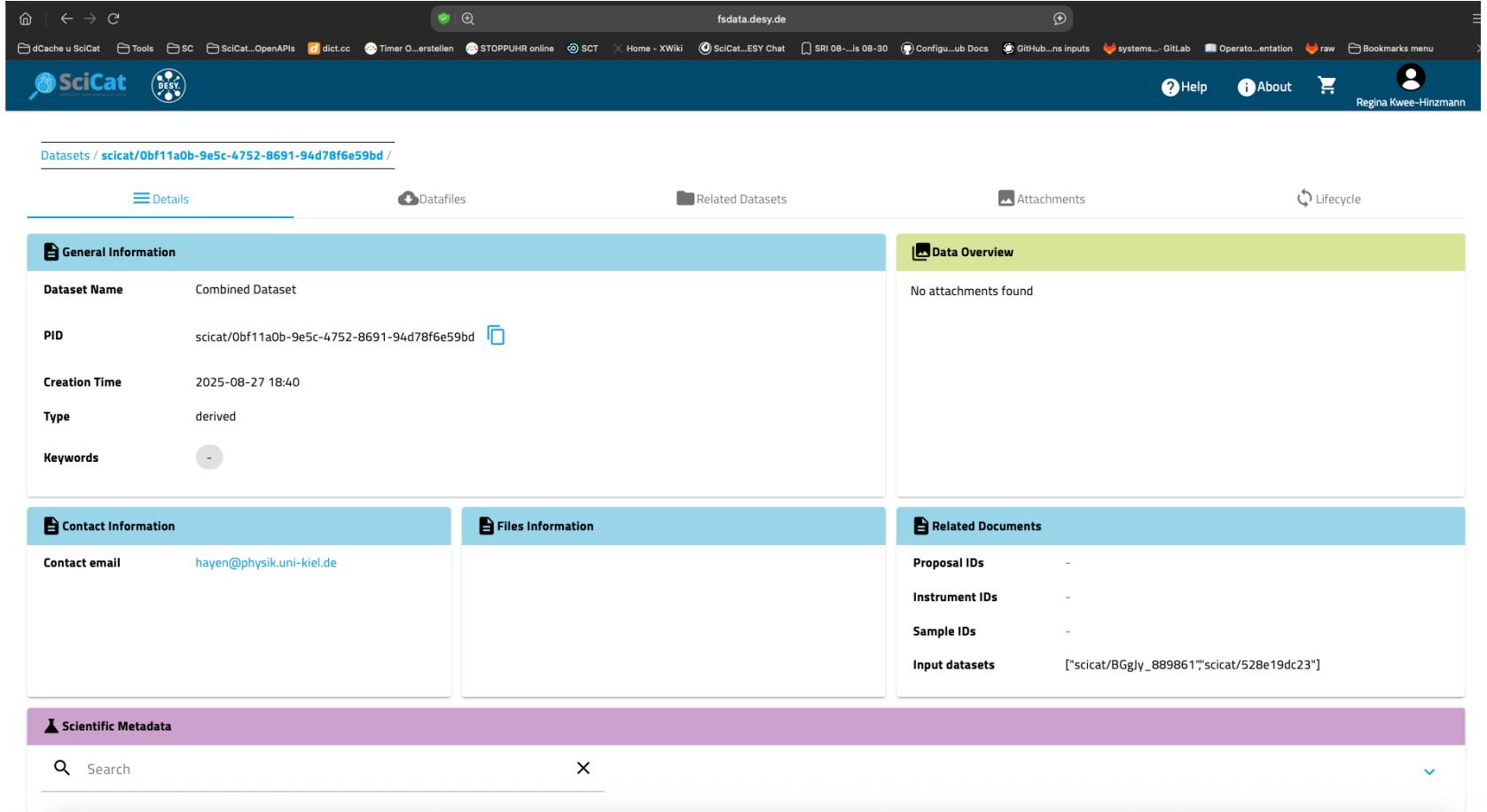
HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

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

DEMO

Status summary

...the dataset on fsdata



The screenshot shows a web browser displaying the fsdata.desy.de dataset page for a combined dataset. The URL in the address bar is [Datasets / scicat/0bf11a0b-9e5c-4752-8691-94d78f6e59bd /](https://fsdata.desy.de/datasets/scicat/0bf11a0b-9e5c-4752-8691-94d78f6e59bd/). The page is organized into several sections:

- General Information**: Dataset Name: Combined Dataset, PID: scicat/0bf11a0b-9e5c-4752-8691-94d78f6e59bd, Creation Time: 2025-08-27 18:40, Type: derived, Keywords: -
- Data Overview**: No attachments found
- Contact Information**: Contact email: hayen@physik.uni-kiel.de
- Files Information**: (empty)
- Related Documents**: Proposal IDs: -, Instrument IDs: -, Sample IDs: -, Input datasets: ["scicat/BGgly_889861", "scicat/528e19dc23"]
- Scientific Metadata**: A search bar with placeholder "Search" and a clear button "X".

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

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