

Guide to Scicat Exercises

+ In relation to Publication

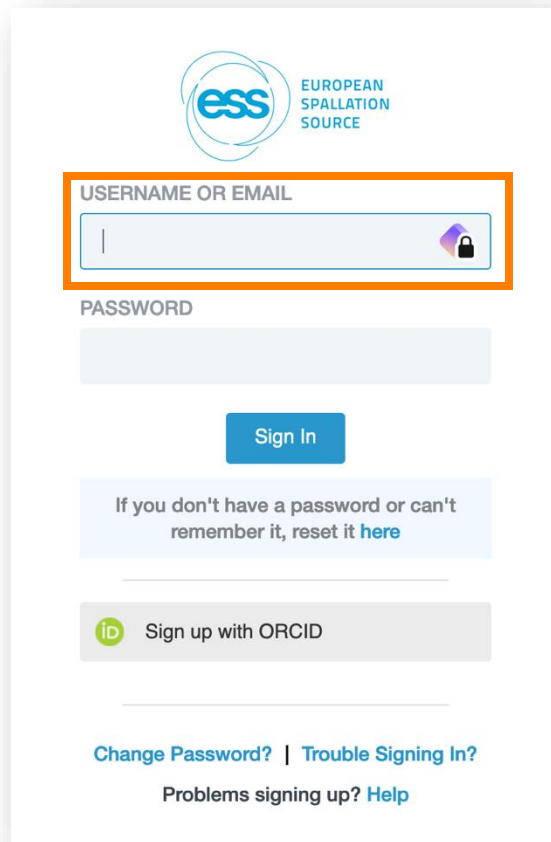
26-08-2025

Prerequisites

ESS Account to log in ESS scicat webpage

<https://staging.scicat.ess.eu>

Same credential you used for the user-office and VISA.



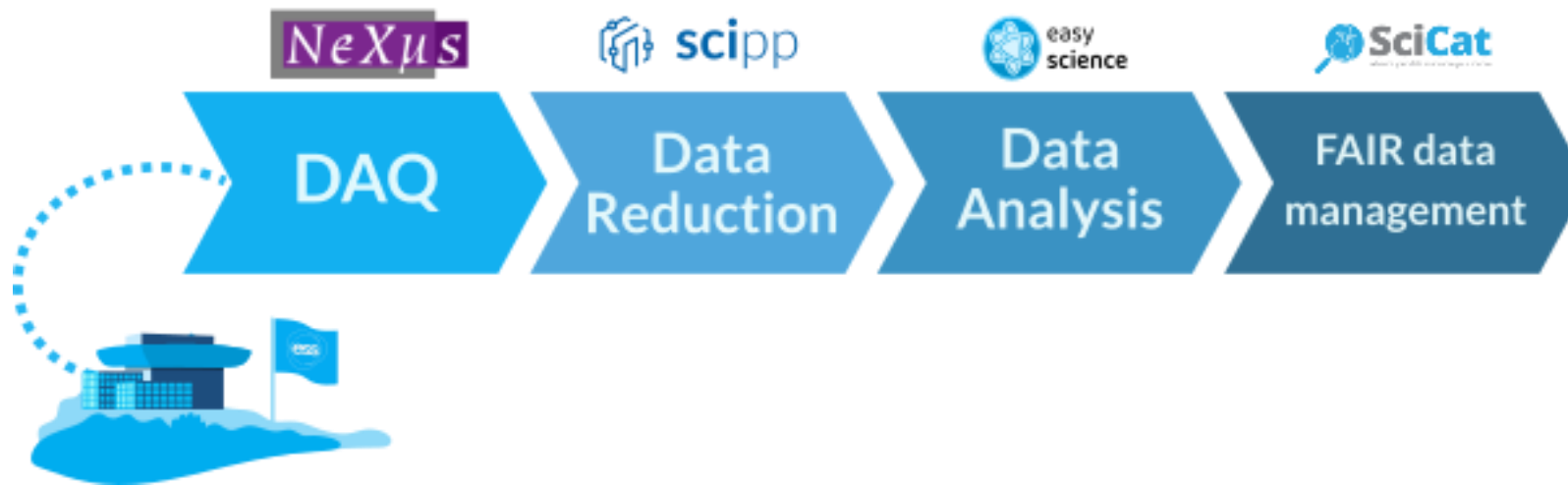
The screenshot shows the ESS login interface. At the top is the ESS logo and the text 'EUROPEAN SPALLATION SOURCE'. Below this is a login form with two main sections: 'USERNAME OR EMAIL' and 'PASSWORD'. The 'USERNAME OR EMAIL' section has a text input field with a cursor and a lock icon, which is highlighted with an orange border. Below it is the 'PASSWORD' section with a password input field. A blue 'Sign In' button is positioned below the password field. Below the button is a light blue box containing the text: 'If you don't have a password or can't remember it, reset it [here](#)'. At the bottom of the form is a grey button with a green 'id' icon and the text 'Sign up with ORCID'. Below the button are two links: 'Change Password?' and 'Trouble Signing In?'. At the very bottom is the text 'Problems signing up? [Help](#)'.

Username is your ORCID ID

Scicat Exercise – FAIR data

At the end of simulation, data reduction and data analysis classes.

Expected data pipeline at an experiment



Scicat Exercise – FAIR data

At the end of simulation, data reduction and data analysis classes.

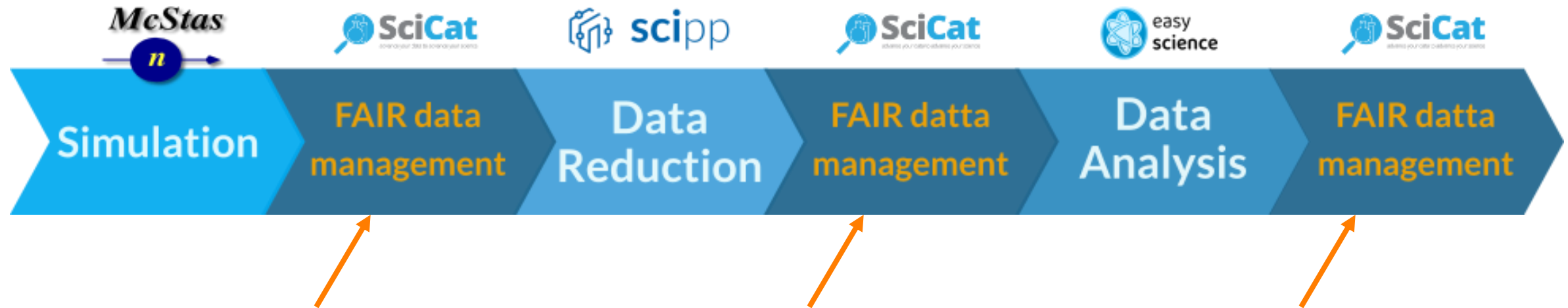
Summer school data pipeline



Scicat Exercise – FAIR data

At the end of simulation, data reduction and data analysis classes.

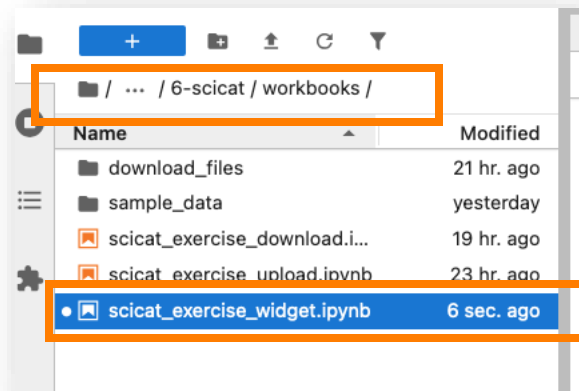
Summer school data pipeline



Scicat Exercise – FAIR data

How to use scicat widget

Open 6-scicat/workbooks/scicat_exercise_widget.ipynb



Make your own dictionary of scientific metadata

```
import scipp as sc

example_metadata = {
    "sample-name": "Iron",
    "sample-width": sc.scalar(0.01, unit="m"),
    "sample-height": 0.2,
    "invalid-metadata": sc.array(dims=["x"], values=[1, 2, 3], unit="m"),
}
```

Instantiate
Scicat Upload Widget

```
from scicat_widgets import upload_widget
upload_wg = upload_widget(metadata_registry=example_metadata)
```

Scicat Widget - Follow the tabs from left to right to upload dataset.

Credentials | Public Personal Info | Prepare Upload | Upload

Go to [your profile](#) to find your SciCat token. (You might need to log in first.)

Scicat Address ☐ Enable Editing

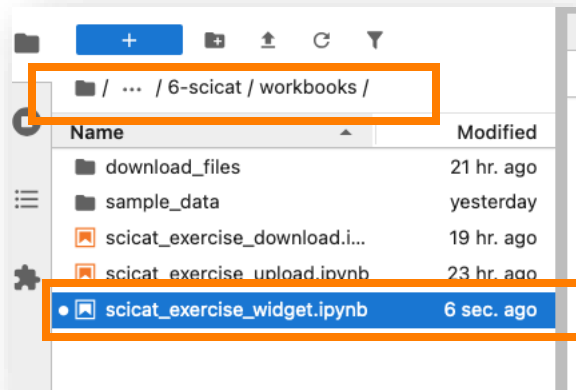
Scicat Token

► Output (Log)

Scicat Exercise – FAIR data

How to use scicat widget

Open 6-scicat/workbooks/scicat_exercise_widget.ipynb



Make your own dictionary of scientific metadata

```
import scipp as sc

example_metadata = {
    "sample-name": "Iron",
    "sample-width": sc.scalar(0.01, unit="m"),
    "sample-height": 0.2,
    "invalid-metadata": sc.array(dims=["x"], values=[1, 2, 3], unit="m"),
}
```

Instantiate
Scicat Upload Widget

```
from scicat_widgets import upload_widget
upload_wg = upload_widget(metadata_registry=example_metadata)
```

Scicat Widget - Follow the tabs from left to right to upload dataset.

Credentials

Public Personal Info

Prepare Upload

Upload

Go to [your profile](#) to find your SciCat token. (You might need to log in first.)

Scicat Address ☐ Enable Editing

Scicat Token

Output (Log)


Click this link to go to your scicat profile

Scicat Exercise – FAIR data

How to use scicat widget - Credentials

User /

User Information



Name


Email

Id

Groups

SciCat Token

Auth strategy

 Copy token to clipboard

```
from scicat_widgets import upload_widget  
upload_wg = upload_widget(metadata_registry=example_metadata)
```

SciCat Widget - Follow the tabs from left to right to upload dataset.

Credentials Public Personal Info Prepare Upload Upload

Go to [your profile](#) to find your SciCat token. (You might need to log in first.)

SciCat Address ☐ Enable Editing

SciCat Token

► Output (Log)

- Token expires every hour due to security
- Even that, do NOT share the token
- Do NOT save it within the code

Scicat Exercise – FAIR data

How to use scicat widget – Prefill Personal Info for dataset

Scicat Widget - Follow the tabs from left to right to upload dataset.

Credentials

Public Personal Info

Prepare Upload

Upload

This information will automatically fill some fields in the next section, **Prepare Upload**.
Click **Save** button to store it in `~/.cache/scicat_widgets/public_personal_info.yaml`.

Name

Email

Can be your work email

ORCID ID

Should be a valid ORCID ID

Save

► Output (Log)

Scicat Widget - Follow the tabs from left to right to upload dataset.

Credentials

Public Personal Info

Prepare Upload

Upload

Fill the mandatory fields below to create a dataset.
Some fields are pre-filled and ineditable.

🚫 **Public Personal Info is NOT in sync with the dataset.**
Click **Fill from Public Info** button to update the dataset fields.

Fill from Public Info

	type	proposal_id	source_folder
Prefilled Fields	derived	213256	/ess/data/workshop/2025/213256/upload/sunyoungyoo,



Scicat Exercise – FAIR data

How to use scicat widget – Fill dataset metadata

name	Name of the dataset
description	
keywords	DMSC Summer School 2025 DMSC Summer School 2025
technique	
used_software	
owner	Sunyoung Yoo
owner_email	sunyoung.yoo@ess.eu
orcid_of_owner	https://orcid.org/0000-0002-6720-8944
owner_group	213256
Parent Dataset PID	Enter the PID of the parent dataset (if any)
File Path	Enter the path to the file to upload
Attachment Path	Enter the path to the file to upload

Name of the dataset should be compact, something you would use as a variable name in programming. i.e. short enough to be convenient, and long enough to be descriptive.

File Path is the path to the file that has the underlying data of this dataset.

Attachment Path is the path to the file that has extra, higher level information of the dataset. i.e. plot in png file.



Scicat Exercise – FAIR data

How to use scicat widget – Add file and attachment file to the dataset

name

File Path

Enter the path to the file to upload

Add File

Attachment Path

Enter the path to the file to upload

Add Attachment

used_software

owner

Sunyoung Yoo

owner_email

sunyoung.yoo@ess.eu

File Path

Enter the path to the file to upload

Add File

/ess/data/workshop/2025/213256/derived/dmsc-school/6-scicat/workbooks/sample_data/dmsc_summer_school_test_data_file_1.dat

X

File Path

Enter the path to the file to upload

Attachment Path

Enter the path to the file to upload

Buttons are activated only if the path is valid.

Once added, it will show the absolute full path and you can remove by clicking `X` button

Scicat Exercise – FAIR data

How to use scicat widget – Fill dataset **scientific** metadata

```
: import scipp as sc

example_metadata = {
    "sample-name": "Iron",
    "sample-width": sc.scalar(0.01, unit="m"),
    "sample-height": 0.2,
    "invalid-metadata": sc.array(dims=["x"], values=[1, 2, 3], unit="m"),
}

: from scicat_widgets import upload_widget

upload_wg = upload_widget(metadata_registry=example_metadata)
```

You can define scientific metadata registry and pass it to the widget constructor.

Scientific Metadata Add a scientific metadata to the dataset.
Tip: Pass **metadata_registry** to the widget constructor.

Add Metadata ☒ From Registry ☐ Arbitrary Input

Key ☒ sample-name ☐ sample-width ☐ sample-height Value Unit

Reset

Create scientific metadata either from registry or from a free form (Arbitrary Input)

You have to click `Add Metadata` to add it to the dataset.

Add Metadata ☐ From Registry ☒ Arbitrary Input

Key Value Unit

sample-name: Iron [] **X**

Click `X` button to remove the metadata from the dataset

Scicat Exercise – FAIR data

How to use scicat widget – Review the final dataset and upload

CredentialsPublic Personal InfoPrepare UploadUpload

Here you can see the dataset you are going to upload.

Click **Upload** button to upload the dataset to SciCat.

Dataset to Upload:

Review the dataset here

DerivedDataset

Name	Type	Value	Description
* creation_time	datetime	2025-08-26 09:55:31+0000	Time when dataset became fully availab.
* input_datasets	list[PID]	None	Array of input dataset identifiers used i...
* source_folder	RemotePath	RemotePath('/ess/data/workshop/2025/213256/upload/su...	Absolute file path on file server containi.
description	str		Free text explanation of contents of dat..
name	str		A name for the dataset, given by the cr...
🔒 pid	PID	None	Persistent identifier of the dataset.
proposal_id	str	213256	The ID of the proposal to which the dat..

► Advanced fields

► Files: 0 (0 B)

▼ Scientific Metadata

Name	Value

Don't forget to open the folded fields too!

Upload

Once you finish reviewing and confirm the upload, the existing dataset form in the previous tab will be reset.

Scicat Exercise – FAIR data

How to use scicat widget – Review the final dataset and upload

Scicat Widget - Follow the tabs from left to right to upload dataset.

It will complain if it fails!

Credentials Public Personal Info Prepare Upload Upload

Failed to upload dataset. Error Message:

Error in operation 'create_dataset_model': 500 Internal Server Error: {"errors":{"datasetName":{"name":"ValidatorError","message":"Path `datasetName` is required."},"properties":{"message":"Path `datasetName` is required."},"type":"required","path":"datasetName","value":"","kind":"required","path":"datasetName","value":""},"_message":"DatasetClass validation failed","name":"ValidationError","message":"DatasetClass validation failed: datasetName: Path `datasetName` is required."}

Click **Confirm** to close this message.



Scicat Exercise – FAIR data

How to use scicat widget – Check it in the scicat webpage.

Scicat Widget - Follow the tabs from left to right to upload dataset.

CredentialsPublic Personal InfoPrepare UploadUpload

Dataset uploaded successfully!

Link to the dataset in scicat

Dataset URL: <https://staging.scicat.ess.eu/datasets/20.500.12269%2Fba25b73d-bb05-4a28-866f-3ad0a9caa79e>

Prepare Upload tab was reset for a new dataset.

Click **Confirm** to close this message.

Confirm

General Information

Dataset Name	Simulation Dataset
Description	Persistent Identifier (PID)
Pid	20.500.12269/ba25b73d-bb05-4a28-866f-3ad0a9caa79e
Type	derived
Creation time	2025-08-26 12:04
Keywords	DMSC Summer School 2025



Scicat Exercise – FAIR data

If you hate using GUI – Scitacean Python API

/ ... / 6-scat / workbooks /	
Name	Modified
download_files	22 hr. ago
sample_data	yesterday
scicat_exercise_download.ipynb	20 hr. ago
scicat_exercise_upload.ipynb	yesterday
scicat_exercise_widget.ipynb	47 min. ago

These two notebooks have example code snippet of using scitacean.

*However, you have to fill a lot of arguments yourself, including the fields that are filled automatically if you use widget.

```
[*]: import getpass  
  
TOKEN = getpass.getpass("Enter your token: ")  
Enter your token: |
```

`getpass` allows you to store the token in memory **relatively** securely than saving it as plain text.

Scicat Exercise – FAIR data

If you hate using GUI – Scitacean Python API

1. Create a client object using right url, token and file transfer object

```
[ ]: from scitacean import Client, RemotePath
      from scitacean.transfer.copy import CopyFileTransfer

      source_folder = RemotePath("/PATH/TO/THE/SOURCE/FOLDER/TO/DATASET")
      client = Client.from_token(
          url="https://staging.scicat.ess.eu/api/v3",
          token=TOKEN,
          file_transfer=CopyFileTransfer(source_folder=source_folder),
      )
```

2. Create a dataset object

```
[ ]: from scitacean import Dataset
      from scitacean import DatasetType

      MY_NAME = "YOUR_NAME"
      MY_EMAIL = "YOUR_EMAIL"
      PROPOSAL_ID = "YOUR_PROPOSAL_ID"

      raw_dataset = Dataset(
          type=DatasetType.DERIVED,
          contact_email="sunyoung.yoo@ess.eu",
          investigator="MY_NAME",
          owner=MY_NAME,
          owner_email=MY_EMAIL,
          used_software=["scipp", "pymuhrec"],
          data_format="tiff",
          is_published=False,
          owner_group=PROPOSAL_ID,
          access_groups=[PROPOSAL_ID],
          instrument_id=None,
          techniques=[],
          keywords=["DMSC Summer School 2025", "Powder"],
          license="unknown",
          proposal_id=PROPOSAL_ID,
          source_folder=source_folder.posix,
          name="Summer School Reduced Dataset",
          description="Awesome reduced dataset from the DMSC Summer School 2025",
      )
      raw_dataset
```

3. Add file and upload!

```
[ ]: raw_dataset.add_local_files("/PATH/TO/THE/FILE")
      raw_dataset
```

Upload the dataset!

```
[ ]: client.upload_new_dataset_now(dataset=raw_dataset)
```



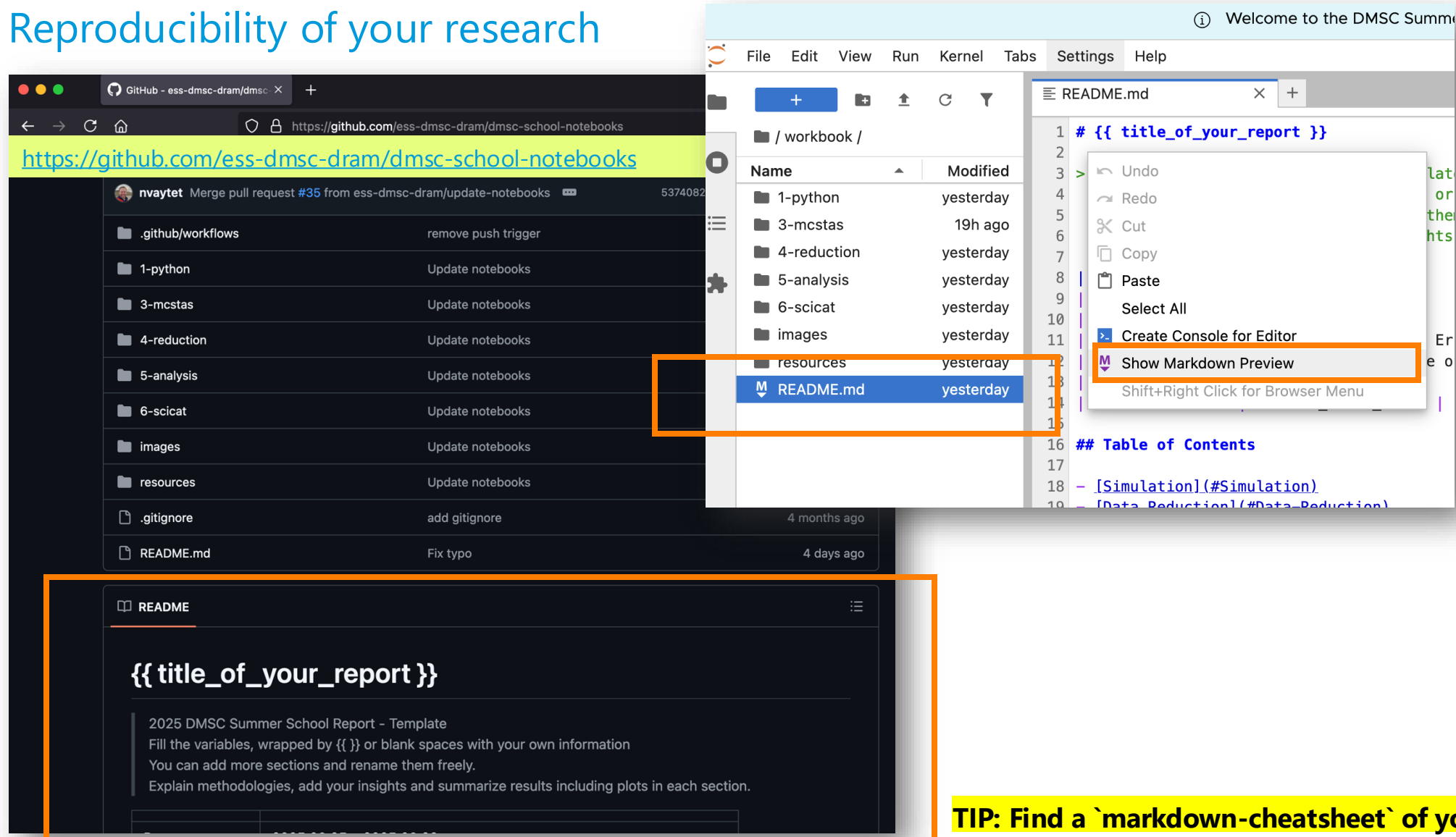
Scicat Exercise – FAIR data

Scicat Exercise, Task 0– Create a dataset with these metadata

Field Name	Field Value
Name (Dataset Name)	scicat_widget_exercise
Description	Scicat widget exercise task 0
Used Software	Python, scitacean
File Path	solutions/6-scatat/workbooks/sample_data/dmsc_summer_school_test_data_file_1.dat

Scicat Exercise – README

Reproducibility of your research



The image displays a GitHub repository and a Jupyter Notebook interface. The GitHub page shows a pull request for 'dmsc-school-notebooks' with a file list including folders 1-python through 6-scicat, images, resources, and a README.md file. The Jupyter Notebook interface shows the same file structure in the left sidebar, with README.md selected. A context menu is open over README.md, showing options like 'Create Console for Editor' and 'Show Markdown Preview'. The main editor area shows the README.md content in Markdown format, including a title placeholder and a table of contents.

README.md Content:

```
1 # {{ title_of_your_report }}
2
3 > Undo
4 > Redo
5 ✂ Cut
6 📄 Copy
7 📄 Paste
8 Select All
9 ▶ Create Console for Editor
10 📄 Show Markdown Preview
11 Shift+Right Click for Browser Menu
12
13 ## Table of Contents
14
15 - [Simulation](#Simulation)
16 - [Data Reduction](#Data-Reduction)
```

README Content:

{{ title_of_your_report }}

2025 DMSC Summer School Report - Template

Fill the variables, wrapped by {{ }} or blank spaces with your own information

You can add more sections and rename them freely.

Explain methodologies, add your insights and summarize results including plots in each section.

TIP: Find a `markdown-cheatsheet` of your flavor



Scicat Exercise – README

Reproducibility of your research

Tip: Make a table of dataset in the README

Simulation Result	
Dataset Name	Scicat Dataset Persistent Identifier (PID)
{{ SIMULATION_DATASET_NAME }}	[{{ SIMULATION_DATASET_PID }}]({{ LINK_TO_THE_SIMULATION_DATASET_IN_SCICAT }}


Scicat Exercise – README

Reproducibility of your research

Simulation Result

Dataset Name	Scicat Dataset Persistent Identifier (PID)
{{ SIMULATION_DATASET_NAME }}	[{{ SIMULATION_DATASET_PID }}] ({{ LINK_TO_THE_SIMULATION_DATASET_IN_SCICAT }}

Instead of:

Name
 simulation_result.h5



Scicat Exercise – README

Reproducibility of your research

Simulation Result

Dataset Name	Scicat Dataset Persistent Identifier (PID)
{{ SIMULATION_DATASET_NAME }}	[{{ SIMULATION_DATASET_PID }}] ({{ LINK_TO_THE_SIMULATION_DATASET_IN_SCICAT }}

Instead of:

Name
 simulation_result.h5
 actual_simulation_result.h5




Scicat Exercise – README

Reproducibility of your research

Simulation Result

Dataset Name	Scicat Dataset Persistent Identifier (PID)
{{ SIMULATION_DATASET_NAME }}	[{{ SIMULATION_DATASET_PID }}] ({{ LINK_TO_THE_SIMULATION_DATASET_IN_SCICAT }}

Instead of:

Name
 simulation_result.h5
 actual_simulation_result.h5
 real_simulation_result.h5





Scicat Exercise – README

Reproducibility of your research

Simulation Result

Dataset Name	Scicat Dataset Persistent Identifier (PID)
{{ SIMULATION_DATASET_NAME }}	[{{ SIMULATION_DATASET_PID }}]({{ LINK_TO_THE_SIMULATION_DATASET_IN_SCICAT }})

Instead of:

Name
 simulation_result.h5
 actual_simulation_result.h5
 real_simulation_result.h5
 final_simulation_result.h5






Scicat Exercise – README

Reproducibility of your research

Simulation Result

Dataset Name	Scicat Dataset Persistent Identifier (PID)
{{ SIMULATION_DATASET_NAME }}	[{{ SIMULATION_DATASET_PID }}] ({{ LINK_TO_THE_SIMULATION_DATASET_IN_SCICAT }}

Instead of:

Name
 simulation_result.h5
 actual_simulation_result.h5
 real_simulation_result.h5
 final_simulation_result.h5
 final_final_simulation_result.h5







Scicat Exercise – README

Reproducibility of your research

Simulation Result

Dataset Name	Scicat Dataset Persistent Identifier (PID)
{{ SIMULATION_DATASET_NAME }}	[{{ SIMULATION_DATASET_PID }}] ({{ LINK_TO_THE_SIMULATION_DATASET_IN_SCICAT }}

Instead of:

Name
 simulation_result.h5
 actual_simulation_result.h5
 real_simulation_result.h5
 final_simulation_result.h5
 final_final_simulation_result.h5
 final_real_final_simulation_result.h5



You can now
use scicat widget and
write README.md!

Continue to your daily exercise.

26-08-2025