



SKADI @ DMSC: SciCat and data curation

Massimiliano (Max) Novelli

Senior data curation scientist

SIMS team (Scientific Information Management Systems)

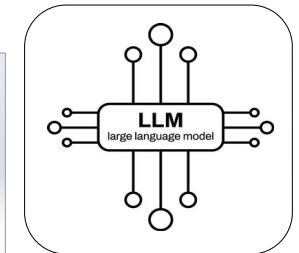
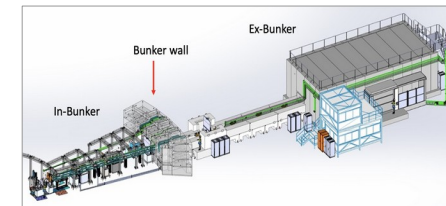
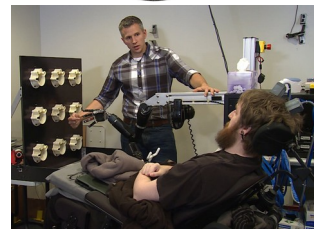
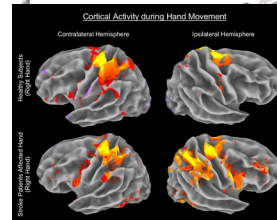
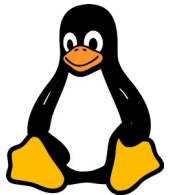
max.novelli@ess.eu

About Me



Senior Data Curation Scientist SIMS team

Scientific Information Management Systems



About the Team



SIMS

Software Information and Management Systems

Janos Babic

0009-0000-0854-5054



Jekabs Karklins

0000-0003-1355-1488



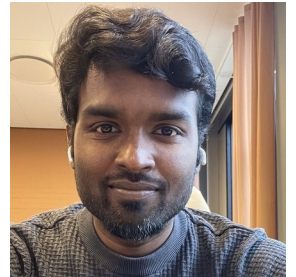
Fredrik Bolmsten



Junjie Quan



Yoganandan Pandiyan

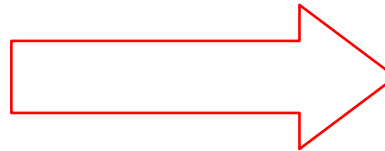
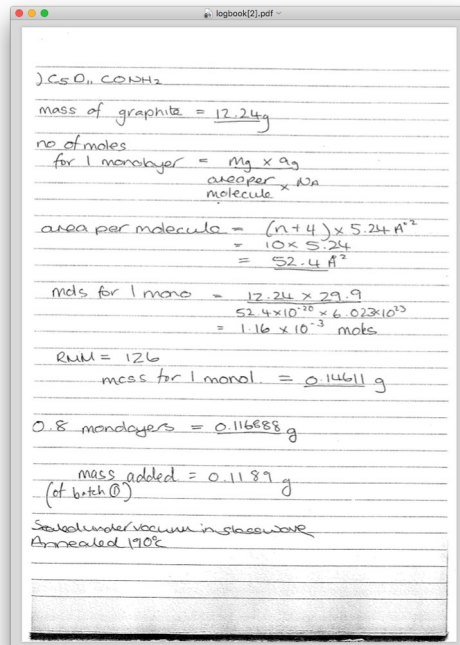


Massimiliano (Max) Novelli

0000-0003-2818-0368



My experiment → My Paper



6

The data is clystar clear



I know what each column is

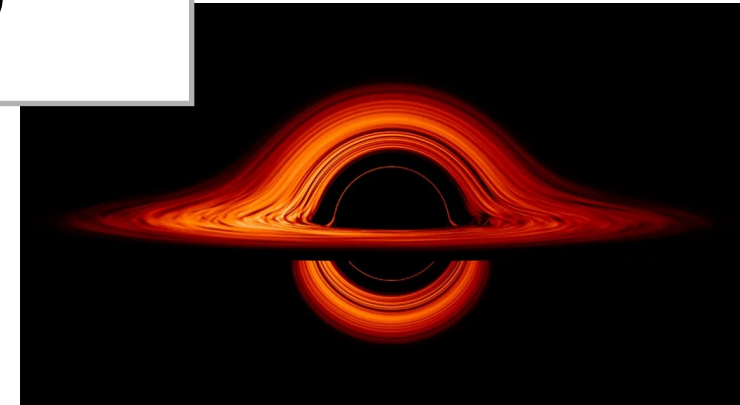
Zplon	Krayd	Xemtu	Bargl	Vutix
4832	wokle	129.55	drex	9021
flarn	7753	zintu	0.348	mokta
221.09	glorb	8876	nexul	1574
trivz	43.7	sporn	6051	quixy



No Reusability



Data Loss



Missed Opportunity to Increase Value

Unusable collection of data



FAIR means...

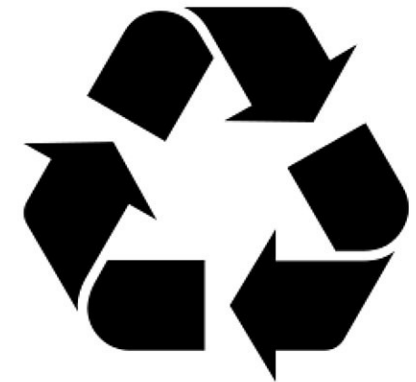
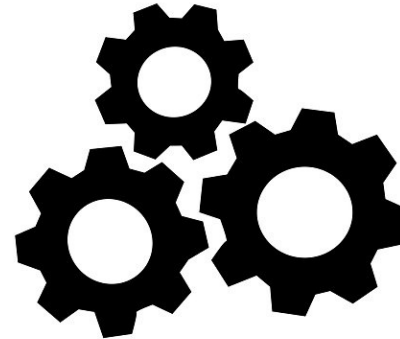


F
indable

A
ccessible

I
nteroperable

R
eusable



<https://www.go-fair.org/fair-principles/>

Data Curation



The Journey to FAIRness

Column 1
Connector 3
Channel B
Blue cable
Left light
Unknown
Resolution



Units
Which will be Metadata
Which aggregation
Other questions?

Sample Temperature
Temperature Controller 1
Sample position X axis
Light detector
Tomography light source
Slit position Y
Sample position resolution

SciCat Project



SciCat

advance your data to advance your science

<https://scicatproject.org>

<https://github.com/scicatproject>

Open Source Project
International collaboration

Industry Partner

DECTRIS
CLOUD



Fredrik Eriksson



Carlo Minotti



Carla Takahashi



Laura Shemilt



Björn Pedersen



Dylan McReynolds



Regina Hinzmann



Patrick Madela



Regina Hinzmann, Nicolas Hayen



Oliver Knodel



Paolo Mutti

SciCat: Proposal



Proposals / 681588 /

Details

Datasets

Proposal Information

Proposal Title

UO to SciCat 2026-06-17 (experiment and sample test)

Abstract

UO to SciCat 2026-06-17 (experiment and sample test)

Proposal Id

681588

Proposal Type

Proposal

Number of Datasets

0

People

Main proposer

Junjie Quan

Principal Investigator

Junjie Quan

Additional Information

Name

UOS Proposal Status

PI First Name

PI Last Name

PI Email

PI ORCID

Number of Cols

Number of Data Access Users

Call

Call Id

Call Start Date

Call End Date

Number of Instruments

Col 1 First Name

Col 1 Last Name

Details

Datasets

Related Proposals

Logbook

Items per page: 10 1 – 10 of 41489

Persistent ID	Dataset Name	Run Nu...	Source Folder	Size	Creation Time	Dataset ...	Thumbnail	Proposal...	
20.500.12269/13...	loki	82759	/ess/data/coda/2...	0 B	2025-09-02 09:41	raw		999999	
20.500.12269/69...	nmx	82760	/ess/data/coda/2...	0 B	2025-09-02 09:43	raw		999999	
20.500.12269/bf...	odin	82761	/ess/data/coda/2...	0 B	2025-09-02 09:46	raw		999999	
20.500.12269/14...	tbl	82762	/ess/data/coda/2...	0 B	2025-09-02 09:52	raw		999999	
20.500.12269/6d...	bifrost	82763	/ess/data/coda/2...	0 B	2025-09-02 09:56	raw		999999	
20.500.12269/07...	estia	82765	/ess/data/coda/2...	0 B	2025-09-02 10:16	raw		999999	
20.500.12269/62...	dream	82764	/ess/data/coda/2...	0 B	2025-09-02 10:17	raw		999999	
20.500.12269/4a...	loki	82766	/ess/data/coda/2...	0 B	2025-09-02 10:32	raw		999999	
20.500.12269/81...	nmx	82767	/ess/data/coda/2...	0 B	2025-09-02 10:45	raw		999999	
20.500.12269/57...	odin	82768	/ess/data/coda/2...	0 B	2025-09-02 11:00	raw		999999	

Items per page: 10 1 – 10 of 41489

SciCat: Datasets



Datasets /

Add to Selection

Filters



My data

Public data

Pid

Location

Find filter

- ☐ ymir (26945)
- ☐ ESS:CODA:ODIN (12178)
- ☐ ESS:CODA:NMX (12118)
- ☐ ESS:CODA:DREAM (12075)
- ☐ ESS:CODA:BIFROST (12019)
- ☐ Unknown (11874)
- ☐ ESS:CODA:ESTIA (10084)
- ☐ ESS:CODA:LoKI (6428)
- ☐ ESS:CODA:LOKI (5861)
- ☐ ESS:CODA:Test Beamline Multiblade (5632)

[Show more...](#)

Group

Find filter

- ☐ 999999 (41490)
- ☐ 977695 (32038)
- ☐ 616254 (13749)
- ☐ 268227 (11019)
- ☐ 038243 (10966)

Dataset name, Run number...

Items per page:

100

1 – 100 of 144530



☐	Persistent ID	Dataset Name	Run Num...	Source Folder	Size	Creation Time	Dataset T...	Thumbnail	Proposal ...	
☐	20.500.12269/...	freia	12650	/ess/data/coda/9...	0 B	2026-07-13 08:41	raw		999999	
☐	20.500.12269/...	estia	12649	/ess/data/coda/9...	0 B	2026-07-13 08:27	raw		999999	
☐	20.500.12269/...	dream	12648	/ess/data/coda/9...	0 B	2026-07-13 08:15	raw		999999	
☐	20.500.12269/...	bifrost	12647	/ess/data/coda/9...	0 B	2026-07-13 08:06	raw		999999	
☐	20.500.12269/...	beer	12646	/ess/data/coda/9...	0 B	2026-07-13 07:58	raw		999999	
☐	20.500.12269/...	tbl	12645	/ess/data/coda/9...	0 B	2026-07-13 07:43	raw		999999	
☐	20.500.12269/...	odin	12644	/ess/data/coda/9...	0 B	2026-07-13 07:31	raw		999999	
☐	20.500.12269/...	nmx	12643	/ess/data/coda/9...	0 B	2026-07-13 07:25	raw		999999	
☐	20.500.12269/...	magic	12642	/ess/data/coda/9...	0 B	2026-07-13 07:18	raw		999999	
☐	20.500.12269/...	loki	12641	/ess/data/coda/9...	0 B	2026-07-13 07:05	raw		999999	
☐	20.500.12269/...	freia	12640	/ess/data/coda/9...	0 B	2026-07-13 06:52	raw		999999	
☐	20.500.12269/...	estia	12639	/ess/data/coda/9...	0 B	2026-07-13 06:44	raw		999999	
☐	20.500.12269/...	dream	12638	/ess/data/coda/9...	0 B	2026-07-13 06:34	raw		999999	
☐	20.500.12269/...	bifrost	12637	/ess/data/coda/9...	0 B	2026-07-13 06:23	raw		999999	
☐	20.500.12269/...	beer	12636	/ess/data/coda/9...	0 B	2026-07-13 06:14	raw		999999	
☐	20.500.12269/...	tbl	12635	/ess/data/coda/9...	0 B	2026-07-13 06:04	raw		999999	

SciCat: Dataset



Details

Datafiles

Related Datasets

Logbook

Attachments

Lifecycle

Admin

Scientific Metadata Table

Name	Value	Unit
Start Time <i>start_time</i>	2026-07-13 08:06	--
End Time <i>end_time</i>	2026-07-13 08:15	--
Run Number <i>run_number</i>	12648	--
ESS Data Collection Job Id <i>job_id</i>	99999901-2648-5d03-89dc-...	--
ESS Ingestor Run Id <i>ingestor_run_id</i>	14f2052b-64ca-4db9-94e2-2...	--
Acquisition Team Members <i>acquisition_team_members</i>	Anders Pettersson, André Cost...	--
Simulated Instrument Name <i>simulated_instrument_name</i>	DREAM	--
Simulated Instrument Id <i>simulated_instrument_id</i>	554e7fe2-a16e-11ed-bafb-f3...	--
Simulated Instrument Url <i>simulated_instrument_url</i>	https://staging.scicat.ess.eu/in...	--
Ingestor metadata schema ID <i>ingestor_metadata_schema_id</i>	8cc81c52-9b96-11ef-ac75-b...	--

Summary

Proposal ID

999999

Run Number

12648

Dataset Name

dream

Dataset Type

raw

Principal Investigator

George O'Neill

Start Time

2026-07-13 08:06

End Time

2026-07-13 08:15

Instrument ID

20.500.12269/00fe23a2-f276-4e2e-9005-a89a9c6ae9fe

Persistent ID

20.500.12269/99999901-2648-5d03-89dc-07e9111f18ba

Embargo Period End

2029-08-13 08:15

Attachments

No attachments found

1 datafile.

Download Selected

Download All

Items per page:

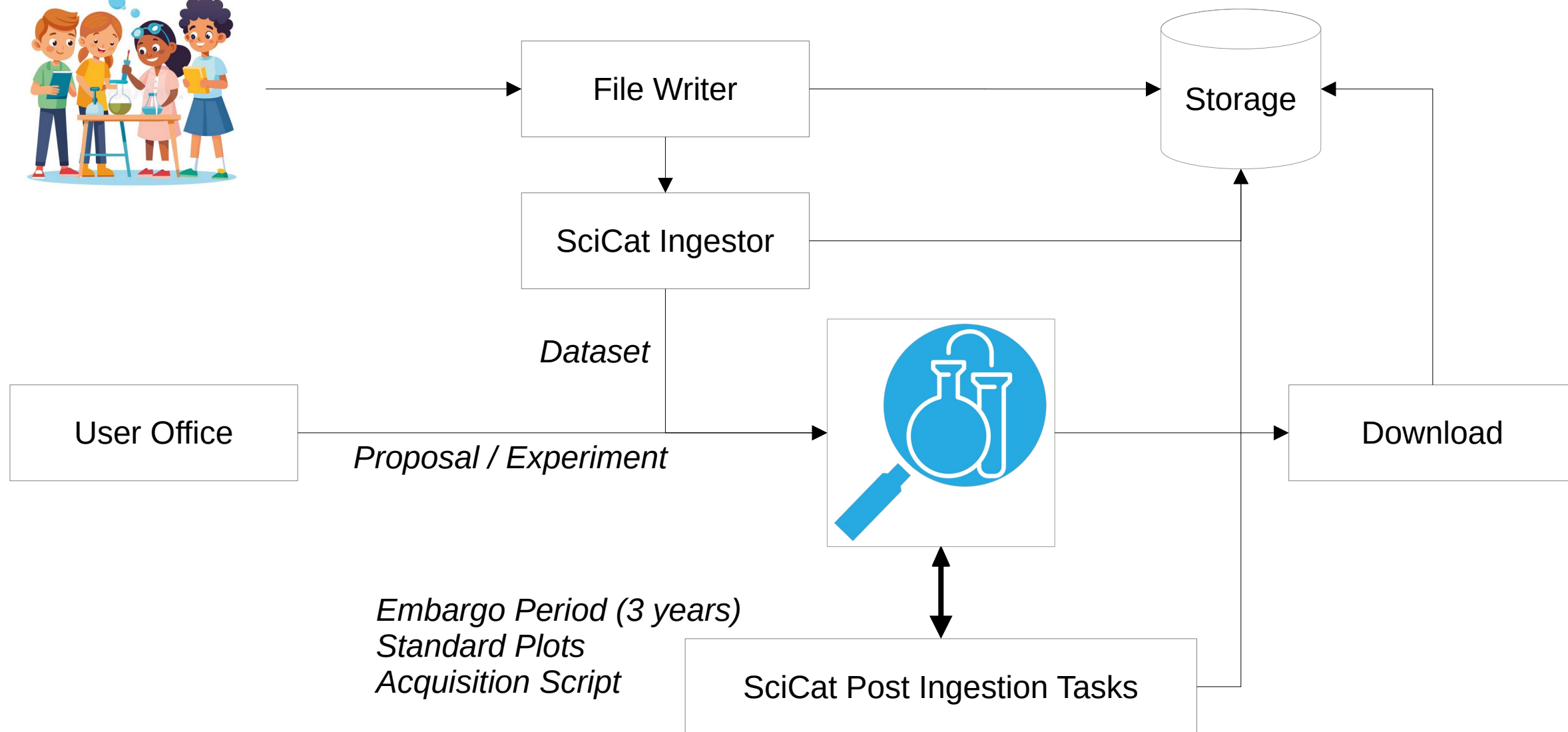
25

1 - 1 of 1



	Path	Size	Time
☐	/ess/raw/coda/999999/raw/coda_dream_999999_00012648.hdf	1 GiB	2026-07-13 08:15

SciCat: Data Path



SciCat: Data Curation



Details

Datafiles

Related Datasets

Logbook

Attachments

Lifecycle

Admin

Scientific Metadata Table

Name	Value	Unit
Start Time <i>start_time</i>	2026-07-13 08:06	--
End Time <i>end_time</i>	2026-07-13 08:15	--
Run Number <i>run_number</i>	12648	--
ESS Data Collection Job Id <i>job_id</i>	99999901-2648-5d03-89dc-...	--
ESS Ingestor Run Id <i>ingestor_run_id</i>	14f2052b-64ca-4db9-94e2-2...	--
Acquisition Team Members <i>acquisition_team_members</i>	Anders Pettersson, André Cost...	--
Simulated Instrument Name <i>simulated_instrument_name</i>	DREAM	--
Simulated Instrument Id <i>simulated_instrument_id</i>	554e7fe2-a16e-11ed-bafb-f3...	--
Simulated Instrument Url <i>simulated_instrument_url</i>	https://staging.scicat.ess.eu/in...	--
Ingestor metadata schema ID <i>ingestor_metadata_schema_id</i>	8cc81c52-9b96-11ef-ac75-b...	--

Summary

Proposal ID

999999

Run Number

12648

Dataset Name

dream

Dataset Type

raw

Principal Investigator

George O'Neill

Start Time

2026-07-13 08:06

End Time

2026-07-13 08:15

Instrument ID

20.500.12269/00fe23a2-f276-4e2e-9005-a89a9c6ae9fe

Persistent ID

20.500.12269/99999901-2648-5d03-89dc-07e9111f18ba

Embargo Period End

2029-08-13 08:15

Attachments

No attachments found

Which metadata we want in SciCat?

SciCat: Data Curation: LOKi example



Status	Property Name	Description	Examples	Optional	Notes	Instrument Team Vetting	Data Curation Vetting
	Dataset PID	Dataset Persistent Identifier. Assigned automatically by the acquisition system	LoKi NeXus /entry/instrument/				
	Proposal ID	Id of the proposal the dataset has been acquired under. This id is generated by the User Office software It also referred to as Experiment Identifier by some members of the instrument teams					
	Dataset Name	Name of the dataset, also referred as title. Currently on Nicos GUI is named "experiment"					
	Principal Investigator	Main proposer of the proposal (PI) Also referred as <i>Proposer</i>					
	Start Time	Date and Time when the data acquisition was started					
	End Time	Date and Time when the data acquisition was stop					
	Instrument Name	Name of the instrument the dataset has been acquired on					
	Run Number	Instrument run number as provided by control system					
	Acquisition Team Members	Name of the people present at the time of the data acquisition, including people not part of the proposal team and ess staff.					

<https://confluence.ess.eu/spaces/SWAP/pages/936378802/LOKi+metadata+-+v2>

SciCat: (CODA) LOKi dataset



⋮	Temperature Sensor 1 Avg <i>temperature_sensor_1_avg</i>	0	C
⋮	Temperature Sensor 2 Avg <i>temperature_sensor_2_avg</i>	0	C
⋮	Temperature Sensor 3 Avg <i>temperature_sensor_3_avg</i>	0	C
⋮	Temperature Sensor 4 Avg <i>temperature_sensor_4_avg</i>	0	C
⋮	Temperature Sensor 5 Avg <i>temperature_sensor_5_avg</i>	0	
⋮	Temperature Sensor 6 Avg <i>temperature_sensor_6_avg</i>	0	
⋮	Temperature Sensor 7 Avg <i>temperature_sensor_7_avg</i>	0	
⋮	Temperature Sensor 8 Avg <i>temperature_sensor_8_avg</i>	0	
⋮	Sample Position X Avg <i>sample_position_x_avg</i>	0	

Summary

Proposal ID	999999
Run Number	12651
Dataset Name	loki
Dataset Type	raw
Principal Investigator	George O'Neill
Start Time	2026-07-13 08:41
End Time	2026-07-13 08:53
Instrument ID	20.500.12269/00fe23a2-f276-4e2e-9005-a89a9c6ae9fe
Persistent ID	20.500.12269/99999901-2651-5b69-8537-a7e9611f18ba
Embargo Period End	2029-08-13 08:53

Thank you!!!

Questions?

Max Novelli
max.novelli@ess.eu