

An Update from public-data.desy.de

Tim Wetzel, Patrick Fuhrmann, Armando Bermudez Martinez, Uwe Jandt, Paul Millar, Sophie Servan, Peter van der Reest, Regina Hinzmann, Johannes Reppin, Christian Voss, Linus Pithan, Anton Barty, ... (DESY)
Julia Kobus, Philipp Jordt, Linus Liedtke, Bridget Murphy (CAU Kiel)

Slides partly with courtesy from Patrick Fuhrmann



This presentation is licensed under
<https://creativecommons.org/licenses/by/4.0/>

FH SciComp Workshop 2025

HELMHOLTZ

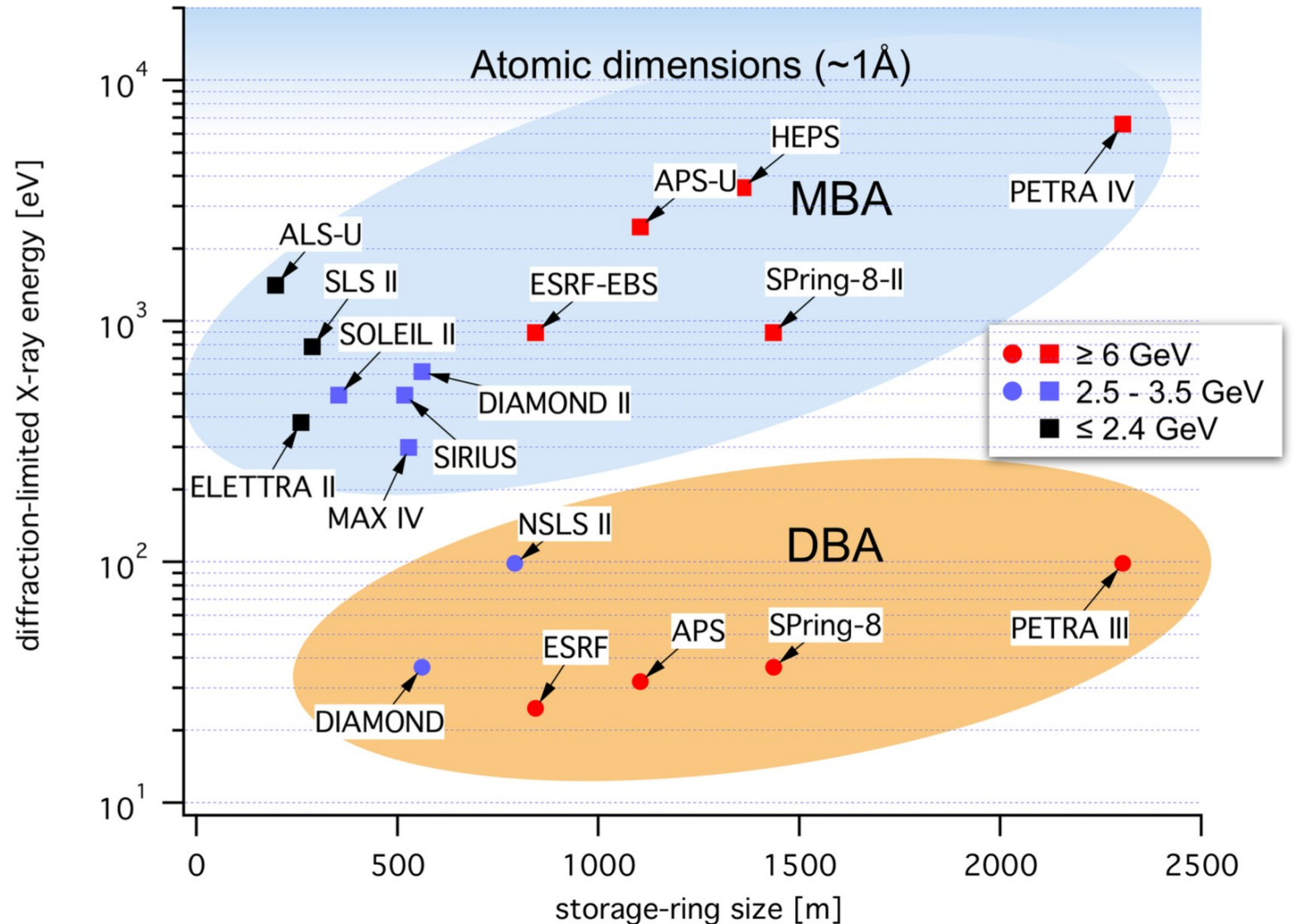


eosc | **BEYOND**

The Synchrotrons

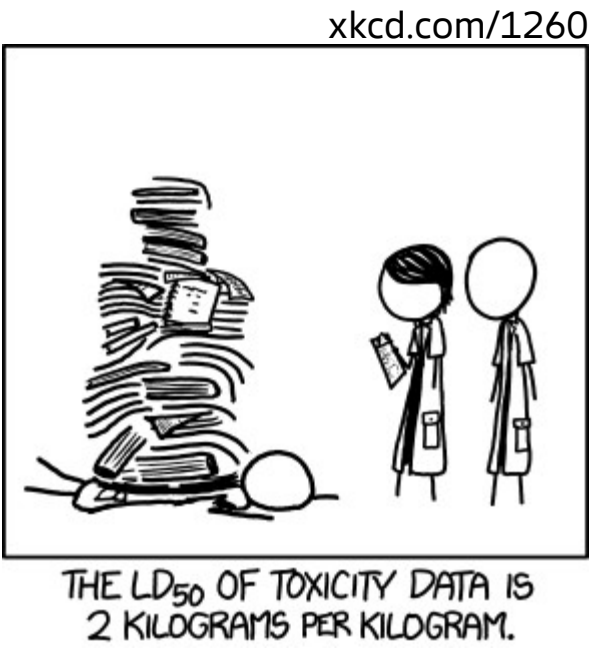
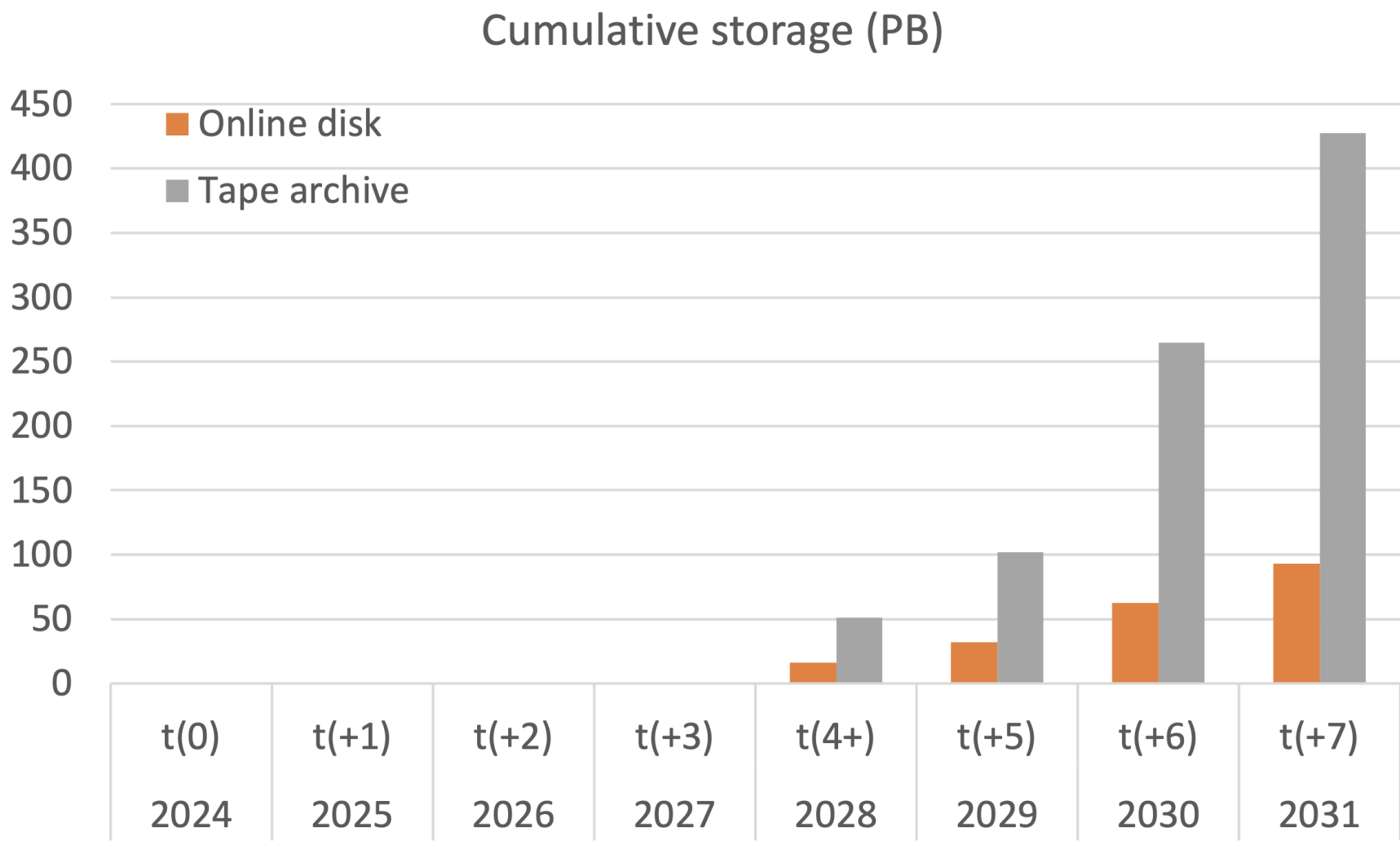
PETRA III/IV and Friends

Diffraction-limited photon energy for some synchrotron radiation sources and their future upgraders. PETRA IV will be the first source worldwide to reach the diffraction limit for hard X-ray energies.



The Consequence

Storage space for PETRA IV



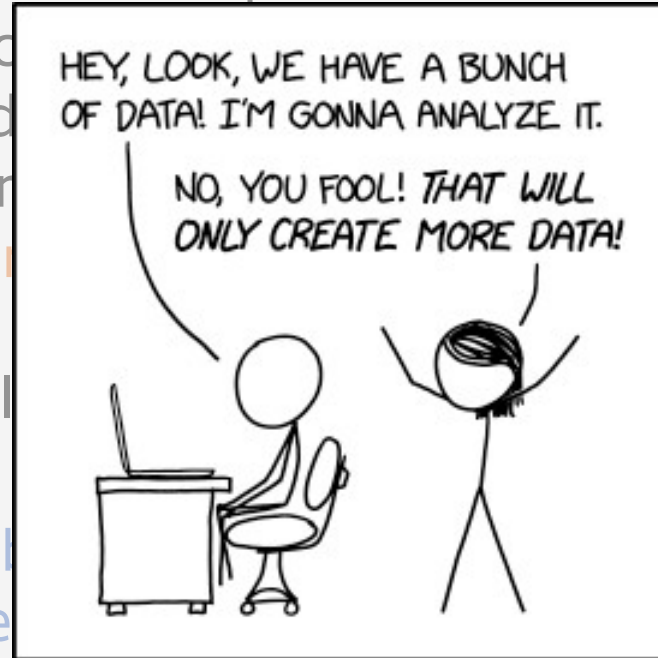
N.B.:
Timeline is
shifted by one
year!



The Data

Where and by whom is it produced?

- Light sources are huge infrastructures, providing extremely brilliant light, used with more than 200 techniques to look deep into matter for **Biology, Material Science, Cultural Heritage** and what not
- In 2019 PETRA III at DESY had 3150 national and international users at 21 beamlines
- Those beam times produce **an extremely precious data, under the governance of the PI.**
- After the '**embargo period**' all data become public '**Open Data**'
- Data which wasn't used in publications
- After a publication, the future of that data is **undetermined.**
- No one, except the original group, understands the details of those collected datasets.










xkcd.com/2582

The Solution!



SciCat

advance your data to advance your science

 EUROPEAN SPALLATION SOURCE <i>Max Novelli</i>	 PSI <i>Carlo Minotti</i>	 <i>Carla Takahashi</i>
 The Rosalind Franklin Institute <i>Laura Shemilt</i>	 Heinz Maier-Leibnitz Zentrum <i>Björn Pedersen</i>	 ADVANCED LIGHT SOURCE <i>Dylan McReynolds</i>
 <i>Regina Hinzmann</i>	 SYNCHROTRON <i>Patrick Madela</i>	 <i>Regina Hinzmann, Nicolas Hayen</i>
 <i>Oliver Knodel</i>	 HELMHOLTZ ZENTRUM DRESDEN ROSSENDORF <i>Oliver Knodel</i>	

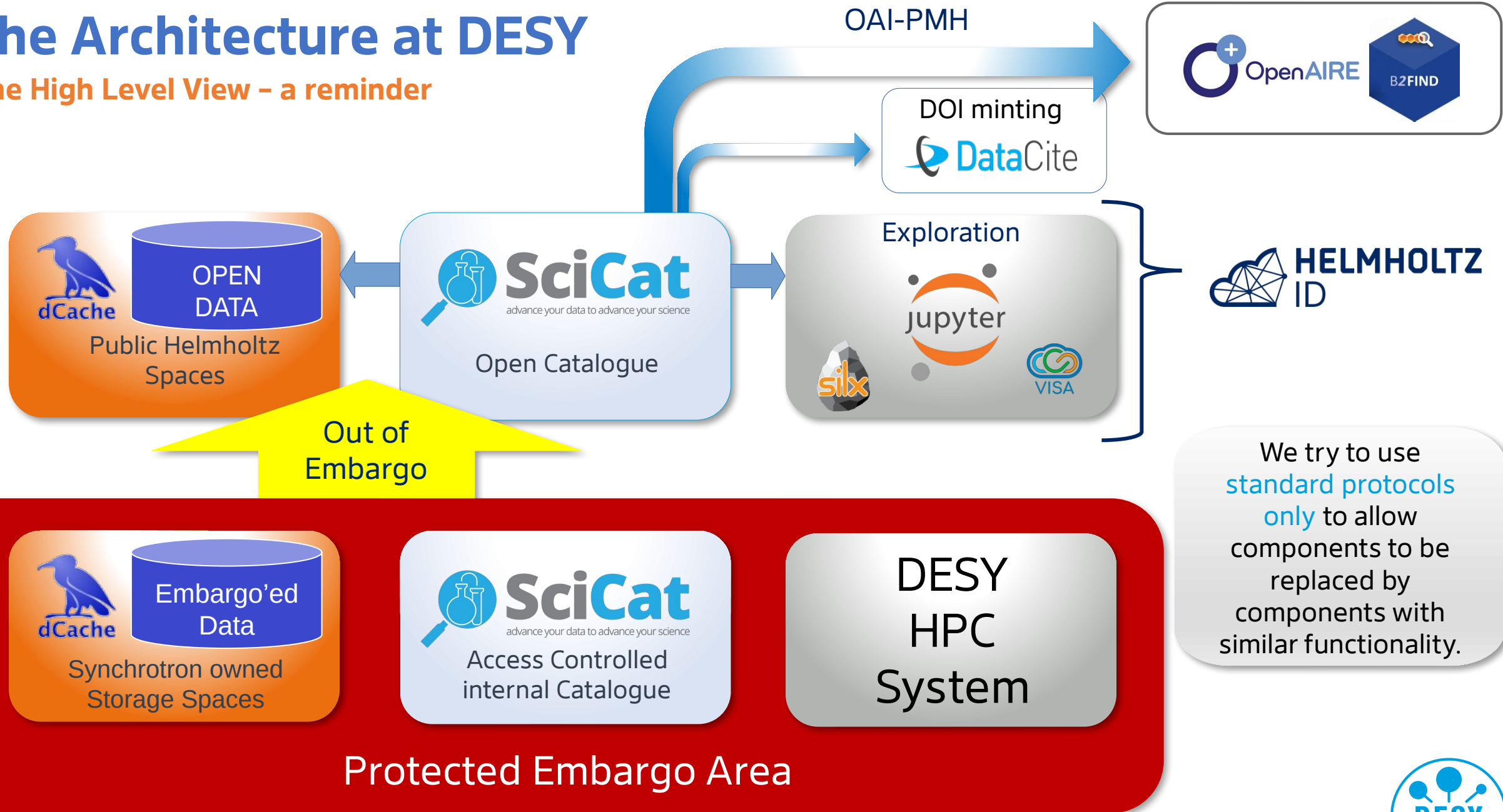
Mission

SciCat is a community-driven scientific metadata catalogue designed to simplify metadata management, enabling data sharing, discovery, and collaboration. By providing a central hub for metadata, SciCat makes data easily findable and accessible for the entire community. With flexible integration into diverse infrastructures and a clear path to data publication, SciCat empowers organizations to foster an open and innovative data culture, driving high-quality science forward.



The Architecture at DESY

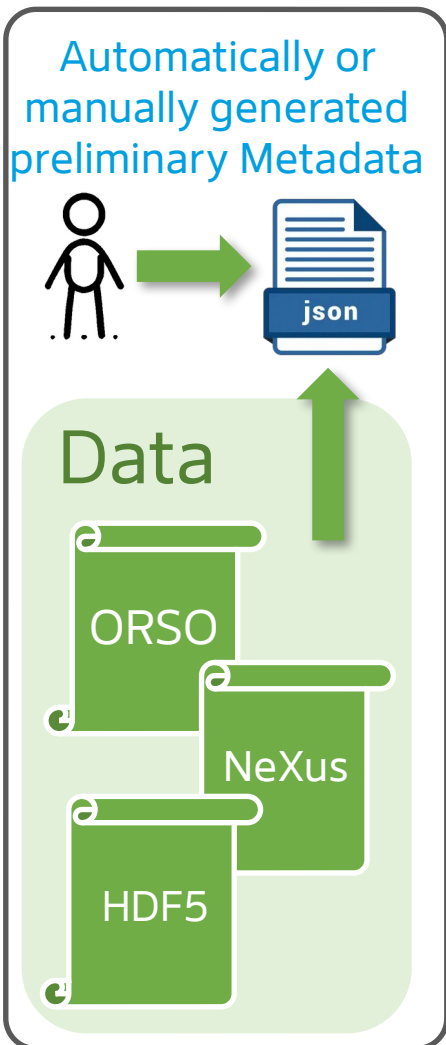
The High Level View – a reminder



The Publishing

sisyphos.desy.de

User Space



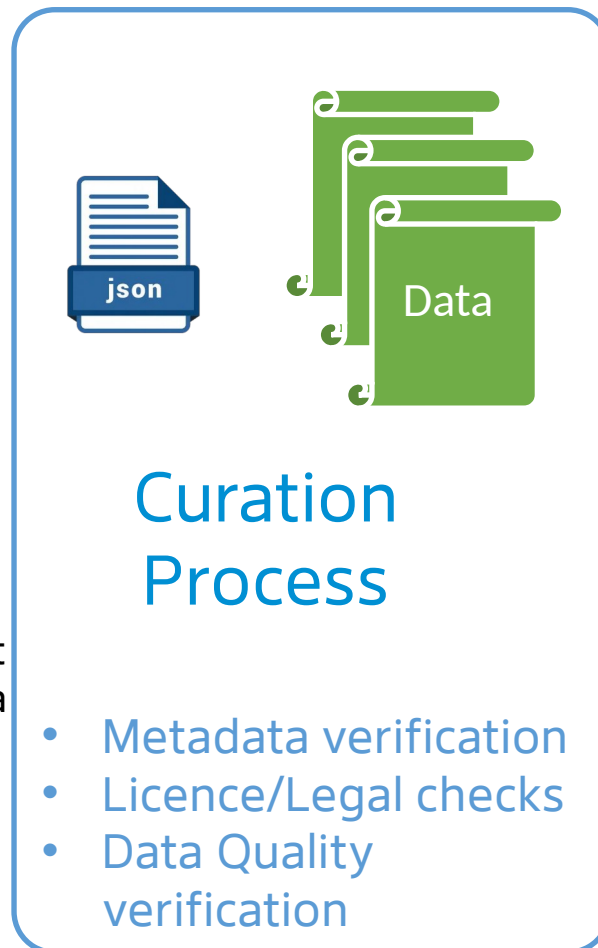
Daphne
Sisyphos
PaN Reflectivity
Database
Upload Tool



Metadata
validation against
defined schemata
enforced.

UI with field
annotations

HIFIS Staging Space



Final
Publishing



Open Data Portal



DOI minting



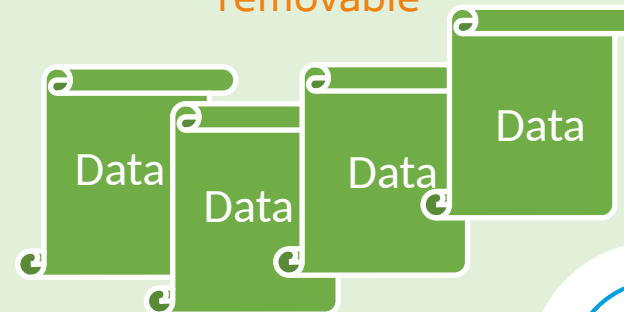
Landing Page

OAI-PMH

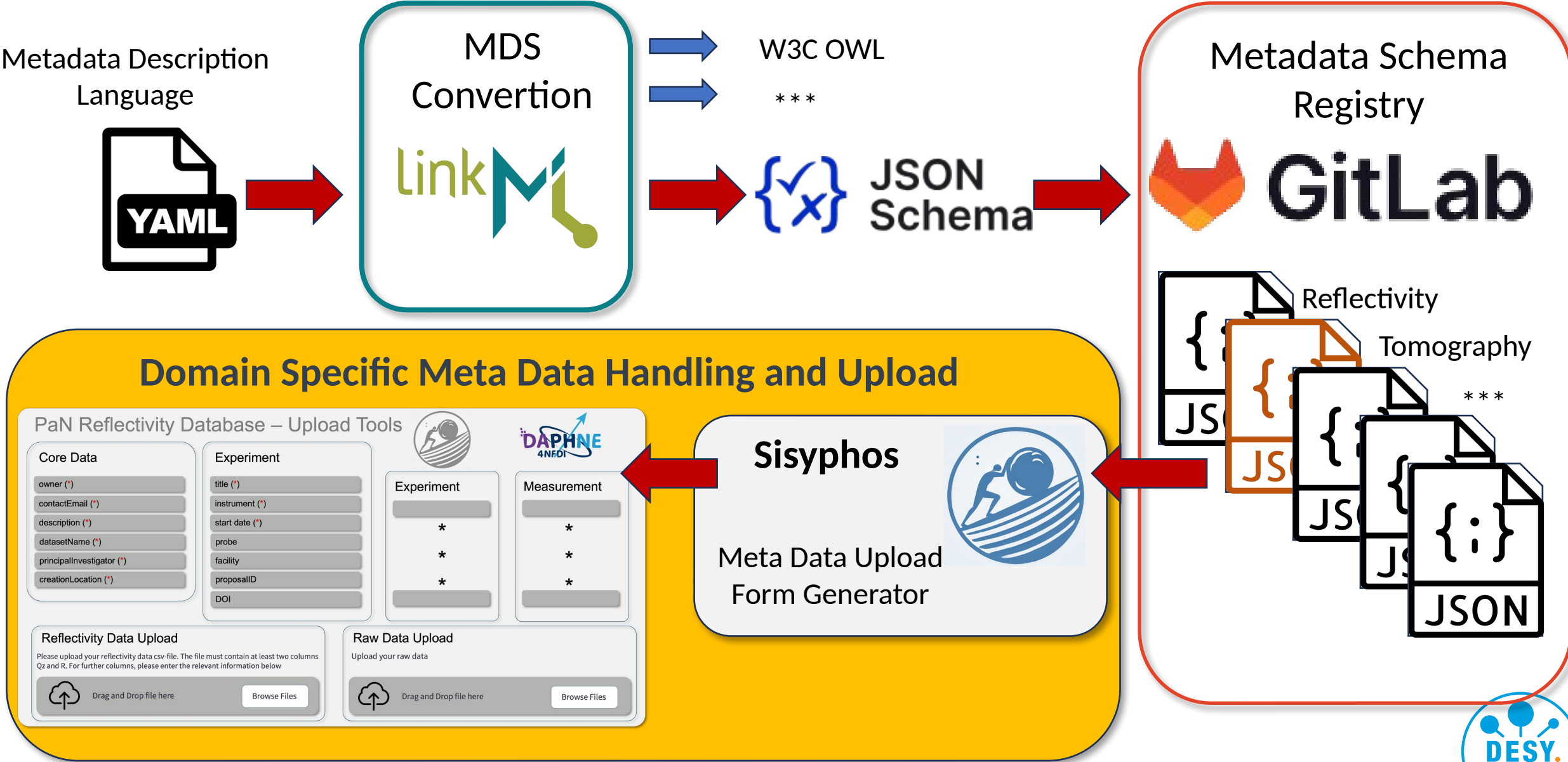


Persistent HIFIS Space

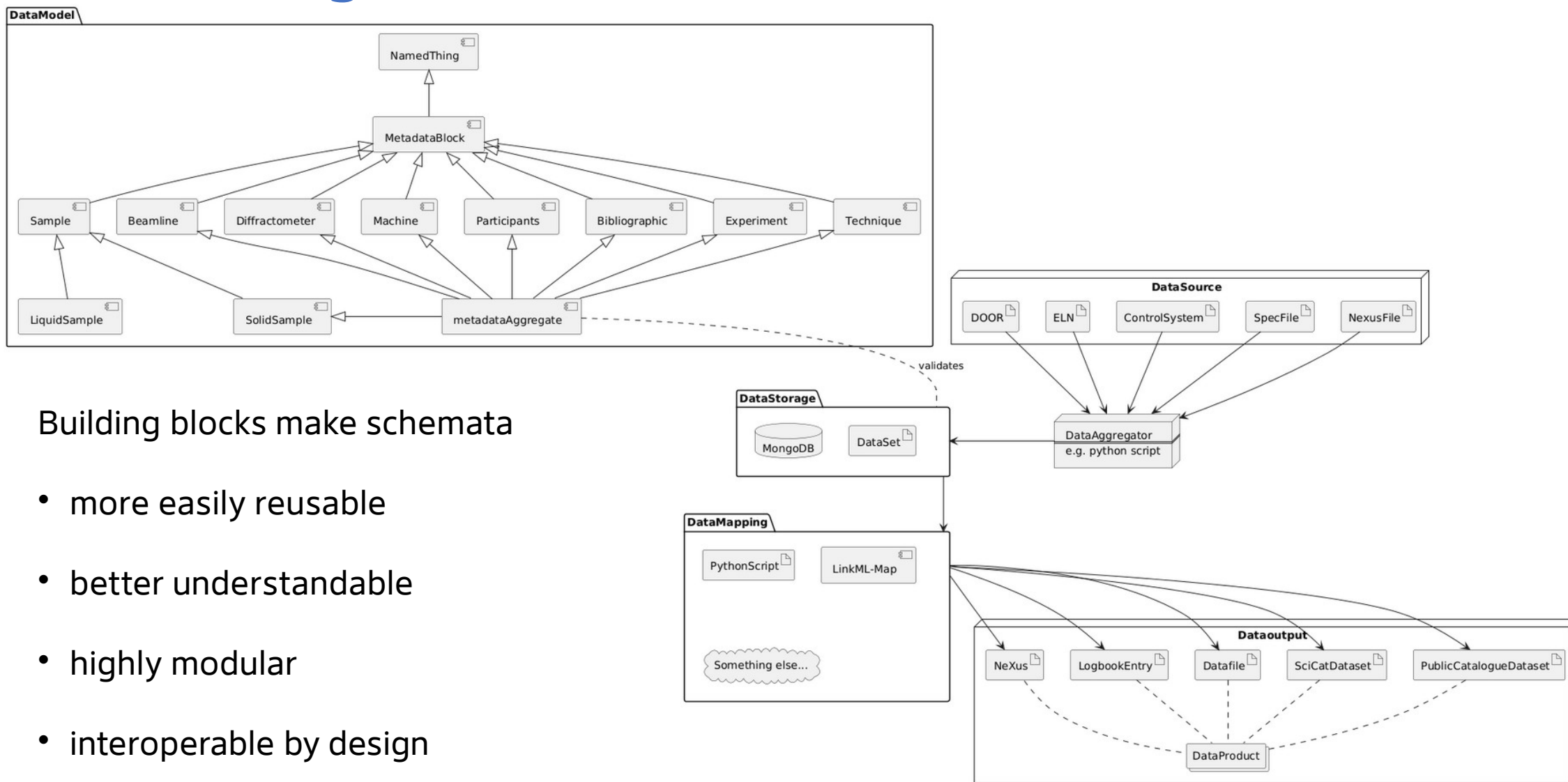
Data is immutable and no longer
removable



The Schemata



The building blocks



Building blocks make schemata

- more easily reusable
- better understandable
- highly modular
- interoperable by design

Source: P. Jordt, CAU, priv. comm.

The Form – sisypchos.desy.de

PaN Reflectivity Database – Upload Tools

Core Data

owner (*)

contactEmail (*)

description (*)

datasetName (*)

principalInvestigator (*)

creationLocation (*)

Experiment

title (*)

instrument (*)

start date (*)

probe

facility

proposalID

DOI



Experiment

*

*

*

Measurement

*

*

*

Reflectivity Data Upload

Please upload your reflectivity data csv-file. The file must contain at least two columns Qz and R. For further columns, please enter the relevant information below



Drag and Drop file here

Browse Files

Raw Data Upload

Upload your raw data



Drag and Drop file here

Browse Files



The Form – sisypchos.desy.de

PaN Reflectivity Database – Upload Tools

Core Data

owner (*)

contactEmail (*)

description (*)

datasetName (*)

principalInvestigator (*)

creationLocation (*)

Experiment

title (*)

instrument (*)

start date (*)

probe

facility

proposalID

DOI

Experiment

*

*

*

Measurement

*

*

*

Reflectivity Data Upload

Please upload your reflectivity data csv-file. The file must contain at least two columns Qz and R. For further columns, please enter the relevant information below

Drag and Drop file here

Browse Files

Raw Data Upload

Upload your raw data

Drag and Drop file here

Browse Files

Meta Data

→

SciCat

advance your data to advance your science

xkcd.com/2054

HELMHOLTZ

FH SciComp Workshop 2025

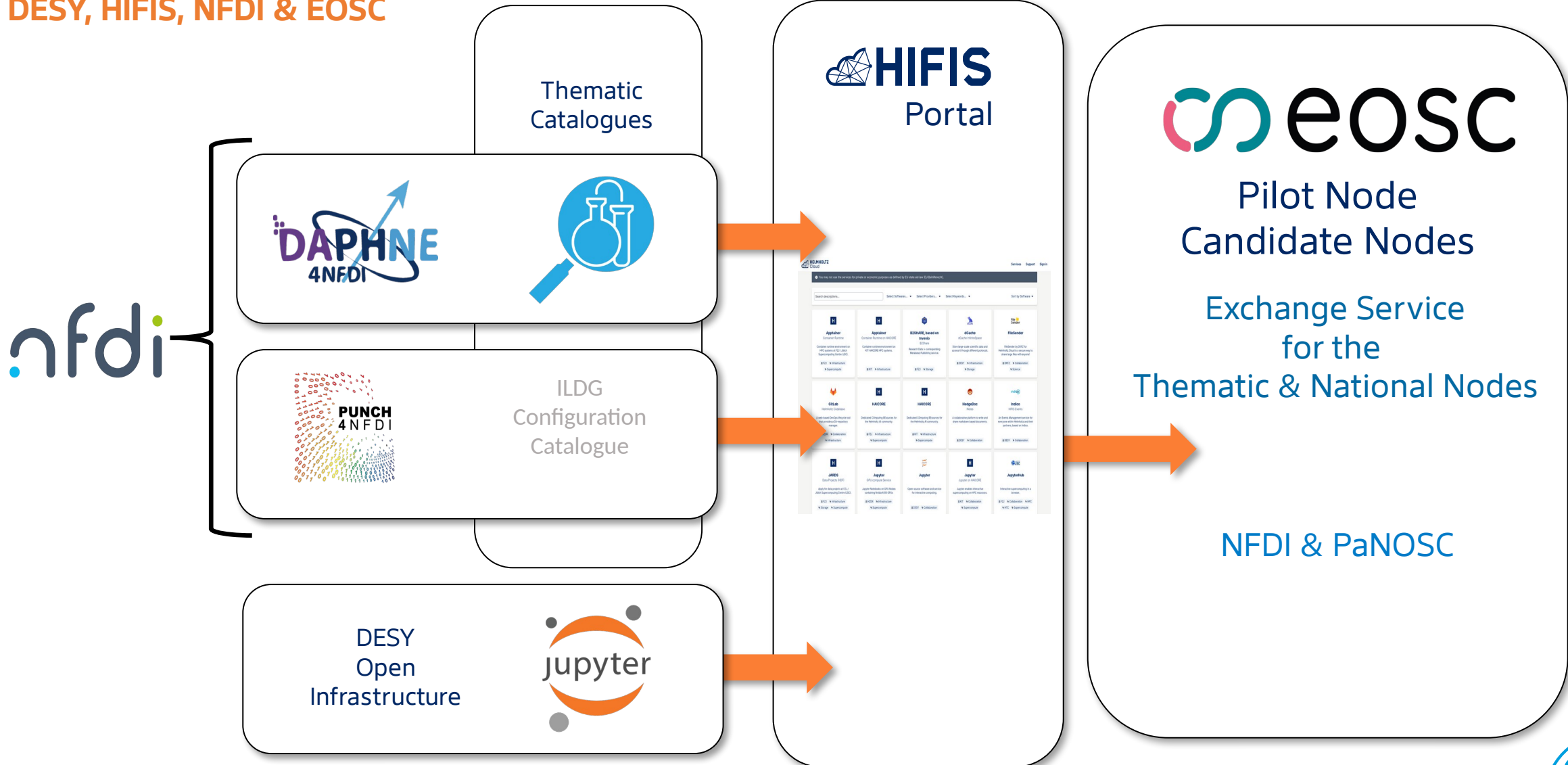
DESY Open Data Portal

Fuhrmann, Wetzel

14

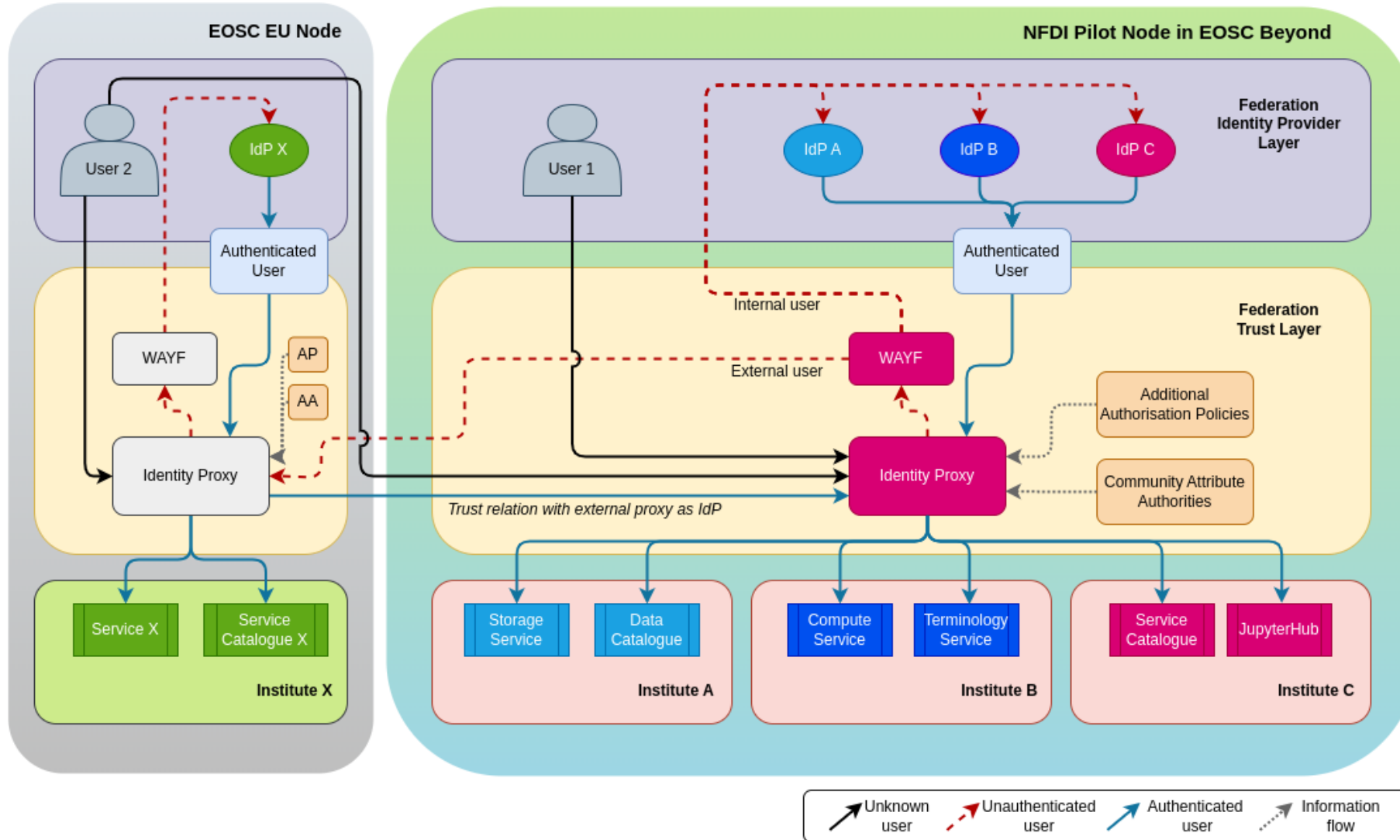
The further Landscapes

DESY, HIFIS, NFDI & EOSC



The Federations

Interfacing with other European communities and services - WIP



The European Open Science Cloud is currently being set up as a network of community-provided platforms ("Nodes")

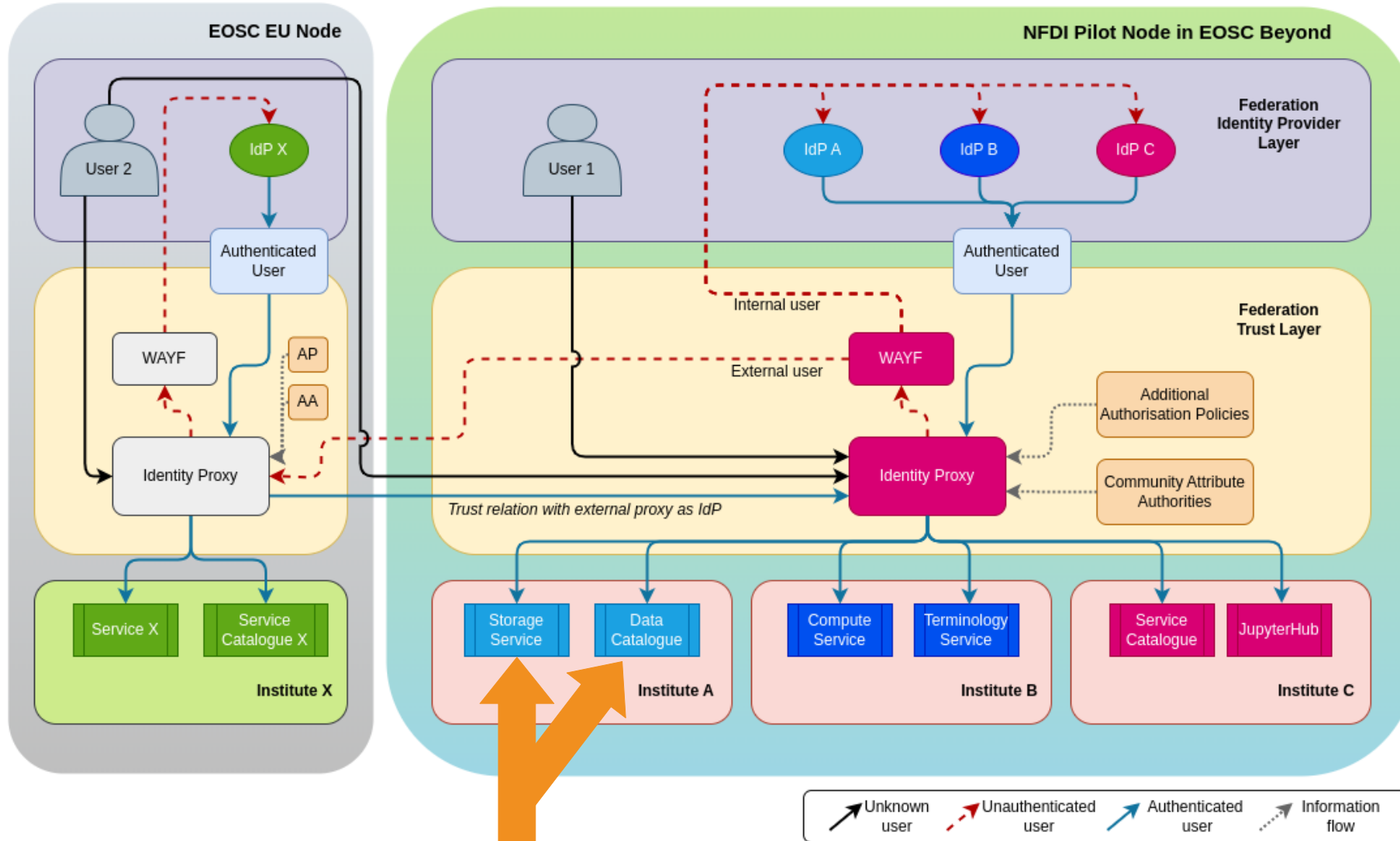
Each node brings their own AAI and services behind it

The nodes are designed to be interoperable in order to bring synergies and interconnections for a more efficient scientific landscape



The Federations

Interfacing with other European communities and services - WIP



Mutual trust for authentication and authorization between communities allows for cross-community usage of

- Identities
- Services
- Data sources
- Infrastructure
- ...

In a modular, unified and transparent way to make it useful and easy for the scientists

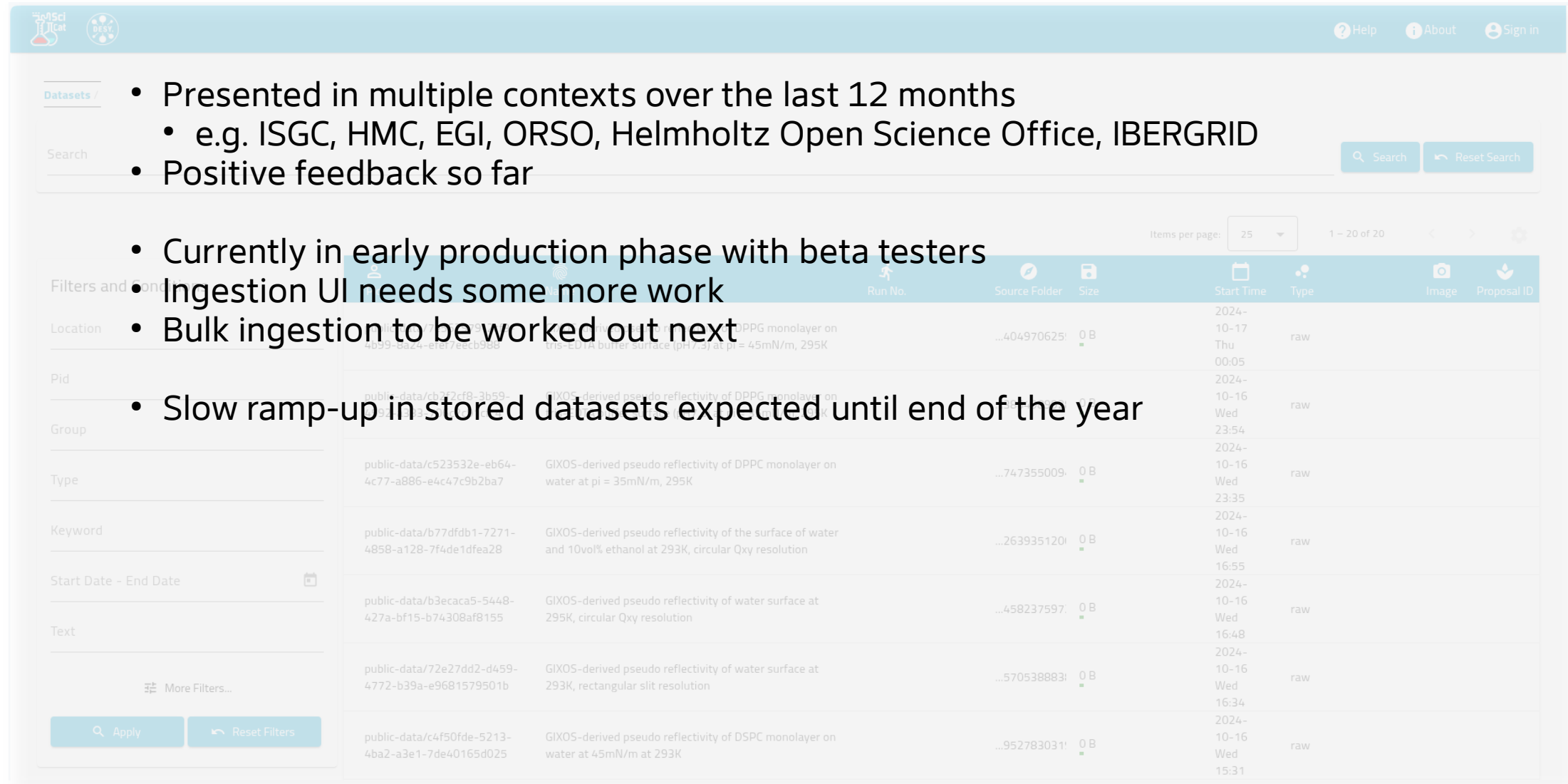


What does the catalogue look like?



The Outlook

Where do we go from here?

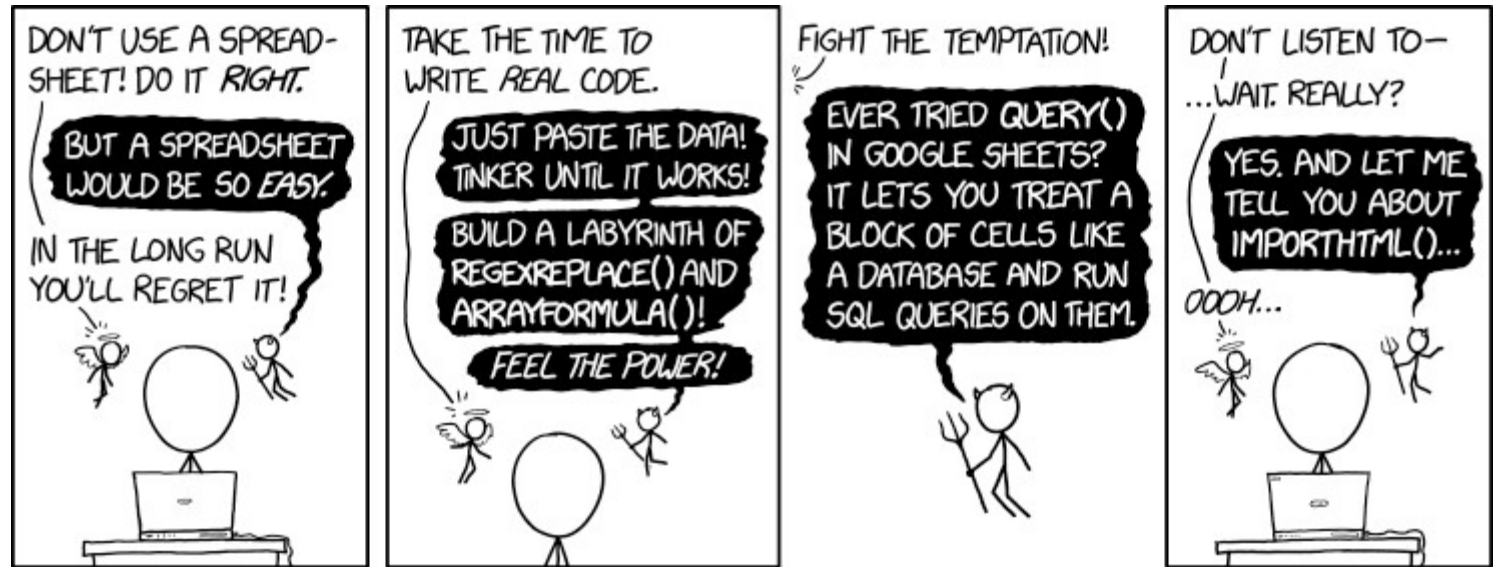


The screenshot displays the DESY Open Data Portal interface. At the top, there's a header with the DESY logo and navigation links for Help, About, and Sign in. Below the header, a 'Datasets' section is visible, featuring a search bar and a list of datasets. The datasets are presented in a table format with columns for Run No., Source Folder, Size, Start Time, Type, and Proposal ID. The table shows several entries, including GIXOS-derived pseudo reflectivity of DPPG monolayer on water at pi = 45mN/m, 295K, and GIXOS-derived pseudo reflectivity of the surface of water and 10vol% ethanol at 293K, circular Qxy resolution. On the left side, there's a 'Filters and Conditions' panel with various filter options like Location, Pid, Group, Type, Keyword, Start Date - End Date, and Text. The interface also includes a 'Search' button and a 'Reset Search' button.

- Presented in multiple contexts over the last 12 months
 - e.g. ISGC, HMC, EGI, ORSO, Helmholtz Open Science Office, IBERGRID
- Positive feedback so far
- Currently in early production phase with beta testers
- Ingestion UI needs some more work
- Bulk ingestion to be worked out next
- Slow ramp-up in stored datasets expected until end of the year



Thank you



xkcd.com/2180

Patrick Fuhrmann, Tim Wetzel

DESY IT

Patrick.Fuhrmann@desy.de, Tim.Wetzel@desy.de

Additional slides







The Files






hifis-storage.desy.de




dCache.org

Rootdesypublic-dataupload








Type	Name	Creation time	File location	Size
	daphne4nfdi	29/11/2023, 14:17:40	Disk	--
	it-ric	29/11/2023, 15:24:43	Disk	--
	punch4nfdi	29/11/2023, 14:18:05	Disk	--



Jupyter Hub

General Information

Dataset Name	GIXOS-derived pseudo reflectivity of the surface of water and 10vol% ethanol at 293K, circular Qxy resolution
Description	pseudo reflectivity derived from grazing incidence off-specular scattering data from the surface of water and 10vol% ethanol at 293K
PID	public-data/b77dfdb1-7271-4858-a128-7f4de1dfea28 
Type	raw
Creation Date	2024-10-16 16:55
Keywords	

Creator Information

Owner	Chen Shen
PI	Chen Shen
Contact	chen.shen@desy.de
Group	ingestor
Access Group	

File Information

Source Folder	/OpenPortal/public-data/ChenShen-Grazingincidenceoffspeculardiffusescatteringmeasurementfromliquidsurfacesatfixedincidentangleanewmethodtoextractspecularreflectivity--9165335642639351200
---------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------



[? Help](#)

[i About](#)

[Sign in](#)

Scientific Meta Data

Q SearchX

▼ Experiment

Title	Grazing incidence off-specular diffuse scattering measurement from liquid surfaces at fixed incident angle: a new method to extract specular reflectivity
Instrument	P08 / Langmuir trough GID setup
Start Date	2023-02-19
Probe	x-ray
Facility	PETRA III
Proposal ID	11017009
DOI	null

▼ Sample

Name	surface of the mixture of water and 10 vol% ethanol
Catagory	vapour / liquid
Composition	90% water and 10% ethanol vol/vol
Description	surface of the mixture of water and 10 vol% ethanol
Environment	customized Kibron G4 Langmuir trough (P08 version)
Sample Parameters	tension: 53mN/m, 293K



The Published Data

public-data.desy.de



[Help](#) [About](#)   ingestor

[Datasets](#) / [Batch](#) /

[Empty Cart](#) [Publish](#) [Share](#) [Archive](#) [Retrieve](#)

 Remove	PID	Source Folder	Creation Date
	public-data/7e562979-2cf9-4b99-8a24-efef7eecb988	/OpenPortal/public-data/ChenShen-GIXOSmeasurementproofofconceptofpseudorefectivity-2883921514049706255	2024-10-17 00:05
	public-data/cb2f2cf8-3b59-4592-a383-706c2c91cfc4	/OpenPortal/public-data/ChenShen-GIXOSmeasurementproofofconceptofpseudorefectivity-5675188953864989928	2024-10-16 23:54





Datasets / Batch / Publish /

Your are about to publish 2 datasets

You are about to publish 2 datasets.

Title*

Pseudo reflectivity reference

Creators*

Chen Shen 

Publisher*

DESY

Resource Type*

raw

Description*

pseudo reflectivity derived from grazing incidence off-specular scattering of DPPG monolayer on the surface of tris-EDTA buffer (pH7.3) at surface pressure of 45mN/m, 295K



Abstract*

pseudo reflectivity derived from grazing incidence off-specular scattering of DPPG monolayer on the surface of tris-EDTA buffer (pH7.3) at surface pressure of 45mN/m, 295K


Download link

Publish






Published Datasets / [10.83065/43dcf67b-9d09-44c6-a377-e07ea5d81b8d](#)

 Publication Status

StatusRegistered

Registration Time2025-03-17. T. 9:30

 General Information

Title

Pseudo reflectivity reference

Abstract

pseudo reflectivity derived from grazing incidence off-specular scattering of DPPG monolayer on the surface of tris-EDTA buffer (pH7.3) at surface pressure of 45mN/m, 295K

DOI


10.83065/43dcf67b-9d09-44c6-a377-e07ea5d81b8d

URL

[public-doi-dev.desy.de10.83065%2F43dcf67b-9d09-44c6-a377-e07ea5d81b8d](#)

Publication Year

2025

 Creator Information


Creator

Chen Shen

Authors

Publisher

DESY

 File Information

Resource Type

raw

Data Description

pseudo reflectivity derived from grazing incidence off-specular scattering of DPPG monolayer on the surface of tris-EDTA buffer (pH7.3) at



 DataCite
Fabrica

TEST - OpenData - DESY - Deutsches Elektronen-Synchrotron

Info Settings Prefixes DOIs

 Create DOI

 Export DOI Metadata

State

- | | |
|-------------------------------------|----|
| <input type="checkbox"/> Draft | 22 |
| <input type="checkbox"/> Findable | 3 |
| <input type="checkbox"/> Registered | 2 |

Resource Type

- | | |
|-----------------------------------|----|
| <input type="checkbox"/> Dataset | 26 |
| <input type="checkbox"/> Software | 1 |

Year created

- | | |
|-------------------------------|----|
| <input type="checkbox"/> 2025 | 22 |
| <input type="checkbox"/> 2024 | 5 |

Repository

Type to search. For example 10.4121/17185607.v1


Search

[Reset All](#)

27 DOIs

Sort by Date Updated ▾


Pseudo reflectivity reference Dataset

Chen Shen,
Raw published 2025 via DESY
pseudo reflectivity derived from grazing incidence off-specular scattering of DPPG monolayer on the surface of tris-EDTA buffer (pH7.3)
at surface pressure of 45mN/m, 295K
Created March 11, 2025 at 08:13:29 UTC. [Findable](#)
 <https://handle.test.datacite.org/10.83065/43dcf67b-9d09-44c6-a377-e07ea5d81b8d>

First test doi minting for production instance Dataset

Armando Bermudez Martinez,
Raw published 2025 via DESY
First test doi minting for production instance
Created March 10, 2025 at 17:05:35 UTC. [Findable](#)
 <https://handle.test.datacite.org/10.83065/9d1d353c-5aaa-40c3-8820-da31b284984d>





TEST - OpenData - DESY - Deutsches Elektronen-Synchrotron / DOIs

10.83065/43dcf67b-9d09-44c6-a377-e07ea5d81b8d

- Update DOI (Form)
- Update DOI (File Upload)

Findable

- Metadata Export
- DataCite XML
- DataCite JSON
- Schema.org JSON-LD
- BibTeX

DOI created

March 11, 2025 at 08:13:29 UTC

DOI registered

March 11, 2025 at 08:13:30 UTC

DOI last updated

March 11, 2025 at 08:13:30 UTC

Schema ⓘ

DataCite Metadata Schema 4

URL

https://doi.example.com10.83065%2F43dcf67b-9d09-44c6-a377-e07ea5d81b8d

Metadata

Summary View ▾

Pseudo reflectivity reference Dataset

Chen Shen,
Raw published 2025 via DESY
pseudo reflectivity derived from grazing incidence off-specular scattering of DPPG monolayer on the surface of tris-EDTA buffer (pH7.3)
at surface pressure of 45mN/m, 295K

https://handle.test.datacite.org/10.83065/43dcf67b-9d09-44c6-a377-e07ea5d81b8d

Citation

APA ▾





Shen, C. (2025). *Pseudo reflectivity reference* [Data set]. DESY. https://doi.org/10.83065/43DCF67B-9D09-44C6-A377-E07EA5D81B8D



DESY Public Data Repository

Public Data Repository Dashboard

Items per page: 20 1 – 3 of 3

<input type="checkbox"/>	 Title	 Registered Time	 Creator	 Doi
<input type="checkbox"/>	test 9 doi minting	2025-02-24 18:01	Armando Bermudez Martinez	10.83065/a73734e6-24b8-4c9c-bf13-0a66d4bc852e
<input type="checkbox"/>	Test 10 doi minting	2025-03-07 09:56	Armando Bermudez Martinez	10.83065/22825797-7ae0-4a2e-9376-20714bf41e00
<input type="checkbox"/>	Pseudo reflectivity reference	2025-03-11 09:13	Chen Shen	10.83065/43dcf67b-9d09-44c6-a377-e07ea5d81b8d



DESY Public Data Repository

Pseudo reflectivity reference

Chen Shen; DESY (2025)

Abstract

pseudo reflectivity derived from grazing incidence off-specular scattering of DPPG monolayer on the surface of tris-EDTA buffer (pH7.3) at surface pressure of 45mN/m, 295K

Publication details

DOI <https://doi.test.datacite.org/doi/10.83065%2F43dcf67b-9d09-44c6-a377-e07ea5d81b8d>

Resource Type raw

Related Publications

Datasets

Data Description pseudo reflectivity derived from grazing incidence off-specular scattering of DPPG monolayer on the surface of tris-EDTA buffer (pH7.3) at surface pressure of 45mN/m, 295K

[public-data/7e562979-2cf9-4b99-8a24-efef7eecb988](#)

[public-data/cb2f2cf8-3b59-4592-a383-706c2c91cfc4](#)





DESY Public data download

You requested the files of the dataset with the ID
"public-data/7e562979-2cf9-4b99-8a24-efef7eecb988"

The files are available at hifis-storage.desy.de.

In order to access the files, you can use either of the following methods:

- Click the following link and download the files directly from your browser:

[direct download](#)

- Alternatively, you can mount the directory mentioned above on your computer by following the instructions on using rclone here:

https://hifis.net/doc/cloud-services/Storage_DESY/rclone_oidcagent/































dCache

OpenPortal

public-data

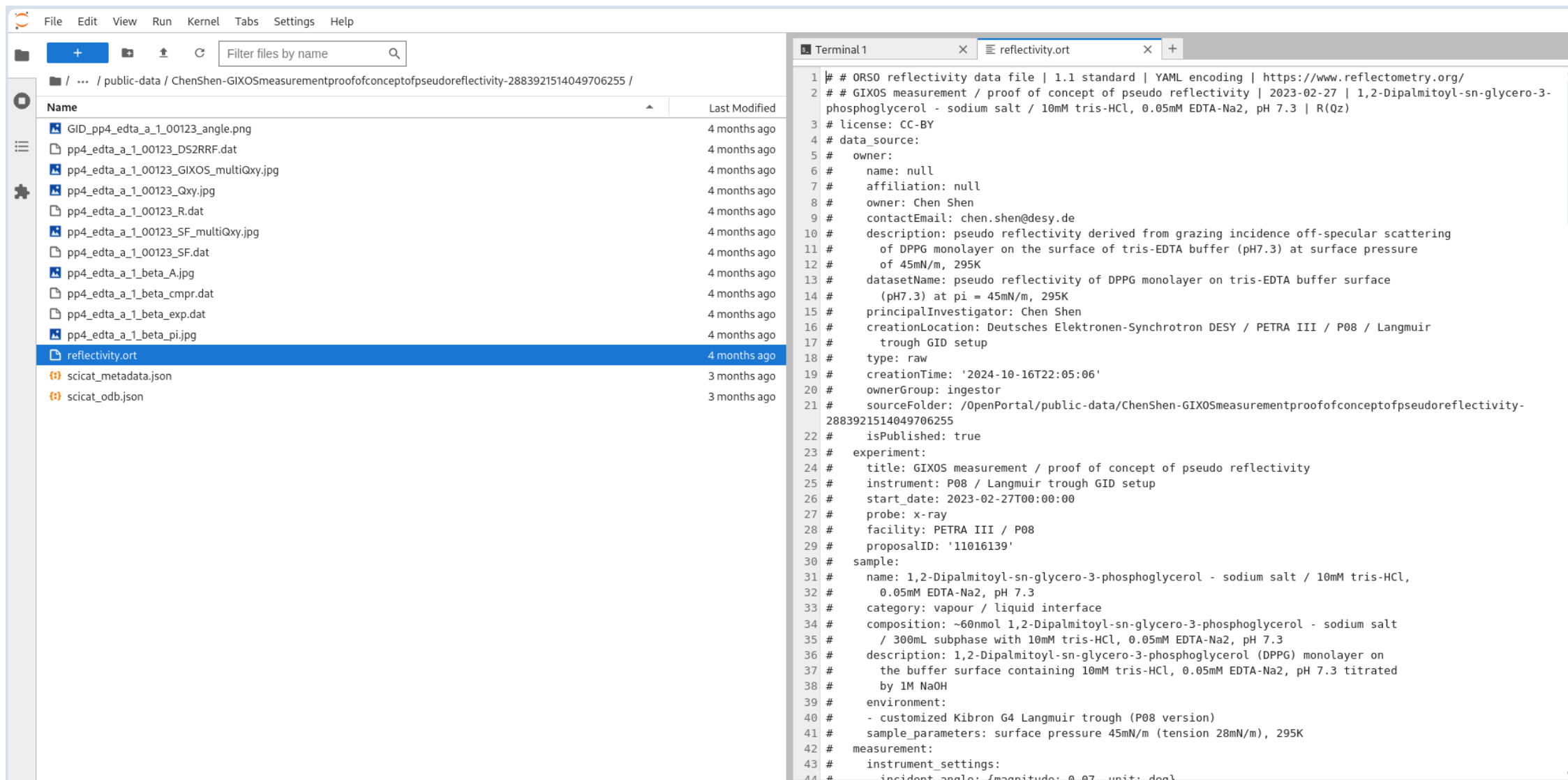
ChenShen-GIXOSmeasurementproofofconceptofpseudorefectivity-2883921514049706255

Name	Size	Last Modified
 GID_pp4_edta_a_1_00123_angle.png 	338512	Thu Oct 17 11:34:44 CEST 2024
 pp4_edta_a_1_00123_DS2RRF.dat 	2169	Thu Oct 17 11:34:44 CEST 2024
 pp4_edta_a_1_00123_GIXOS_multiQxy.jpg 	145231	Thu Oct 17 11:34:44 CEST 2024
 pp4_edta_a_1_00123_Qxy.jpg 	97479	Thu Oct 17 11:34:44 CEST 2024
 pp4_edta_a_1_00123_R.dat 	4163	Thu Oct 17 11:34:44 CEST 2024
 pp4_edta_a_1_00123_SF.dat 	5034	Thu Oct 17 11:34:44 CEST 2024
 pp4_edta_a_1_00123_SF_multiQxy.jpg 	147746	Thu Oct 17 11:34:44 CEST 2024
 pp4_edta_a_1_beta_A.jpg 	430633	Thu Oct 17 11:34:44 CEST 2024
 pp4_edta_a_1_beta_cmpr.dat 	26623	Thu Oct 17 11:34:44 CEST 2024
 pp4_edta_a_1_beta_exp.dat 	25604	Thu Oct 17 11:34:44 CEST 2024
 pp4_edta_a_1_beta_pi.jpg 	479965	Thu Oct 17 11:34:44 CEST 2024
 reflectivity.ort 	12119	Thu Oct 17 11:34:44 CEST 2024
 scicat_metadata.json 	4880	Wed Nov 27 12:27:31 CET 2024
 scicat_odb.json 	3063	Fri Nov 22 09:13:50 CET 2024



The Exploration

jupyter.desy.de



The screenshot displays a JupyterLab environment. On the left, a file browser shows the directory structure: `/ ... / public-data / ChenShen-GIXOSmeasurementproofofconceptofpseudoreflectivity-2883921514049706255 /`. The file list includes various image and data files, with `reflectivity.ort` selected. On the right, a terminal window titled 'Terminal1' shows the content of `reflectivity.ort`, a YAML file with the following metadata:

```
1 |# # ORSO reflectivity data file | 1.1 standard | YAML encoding | https://www.reflectometry.org/
2 |# # GIXOS measurement / proof of concept of pseudo reflectivity | 2023-02-27 | 1,2-Dipalmitoyl-sn-glycero-3-
3 |phosphoglycerol - sodium salt / 10mM tris-HCl, 0.05mM EDTA-Na2, pH 7.3 | R(Qz)
4 |# license: CC-BY
5 |# data_source:
6 |# owner:
7 |#   name: null
8 |#   affiliation: null
9 |#   owner: Chen Shen
10 |#   contactEmail: chen.shen@desy.de
11 |#   description: pseudo reflectivity derived from grazing incidence off-specular scattering
12 |#     of DPPG monolayer on the surface of tris-EDTA buffer (pH7.3) at surface pressure
13 |#     of 45mN/m, 295K
14 |#   datasetName: pseudo reflectivity of DPPG monolayer on tris-EDTA buffer surface
15 |#     (pH7.3) at pi = 45mN/m, 295K
16 |#   principalInvestigator: Chen Shen
17 |#   creationLocation: Deutsches Elektronen-Synchrotron DESY / PETRA III / P08 / Langmuir
18 |#     trough GID setup
19 |#   type: raw
20 |#   creationTime: '2024-10-16T22:05:06'
21 |#   ownerGroup: ingestor
22 |#   sourceFolder: /OpenPortal/public-data/ChenShen-GIXOSmeasurementproofofconceptofpseudoreflectivity-
23 |#     2883921514049706255
24 |#   isPublished: true
25 |# experiment:
26 |#   title: GIXOS measurement / proof of concept of pseudo reflectivity
27 |#   instrument: P08 / Langmuir trough GID setup
28 |#   start_date: 2023-02-27T00:00:00
29 |#   probe: x-ray
30 |#   facility: PETRA III / P08
31 |#   proposalID: '11016139'
32 |# sample:
33 |#   name: 1,2-Dipalmitoyl-sn-glycero-3-phosphoglycerol - sodium salt / 10mM tris-HCl,
34 |#     0.05mM EDTA-Na2, pH 7.3
35 |#   category: vapour / liquid interface
36 |#   composition: ~60nmol 1,2-Dipalmitoyl-sn-glycero-3-phosphoglycerol - sodium salt
37 |#     / 300mL subphase with 10mM tris-HCl, 0.05mM EDTA-Na2, pH 7.3
38 |#   description: 1,2-Dipalmitoyl-sn-glycero-3-phosphoglycerol (DPPG) monolayer on
39 |#     the buffer surface containing 10mM tris-HCl, 0.05mM EDTA-Na2, pH 7.3 titrated
40 |#     by 1M NaOH
41 |#   environment:
42 |#     - customized Kibron G4 Langmuir trough (P08 version)
43 |#   sample_parameters: surface pressure 45mN/m (tension 28mN/m), 295K
44 |# measurement:
45 |#   instrument_settings:
46 |#     incident_angle: {magnitude: 0.07, unit: deg}
```

