

FLASH metadata workflow with SciCat

Dmytro Kutnyakhov

SciCat meeting

Hamburg, 13.11.2023

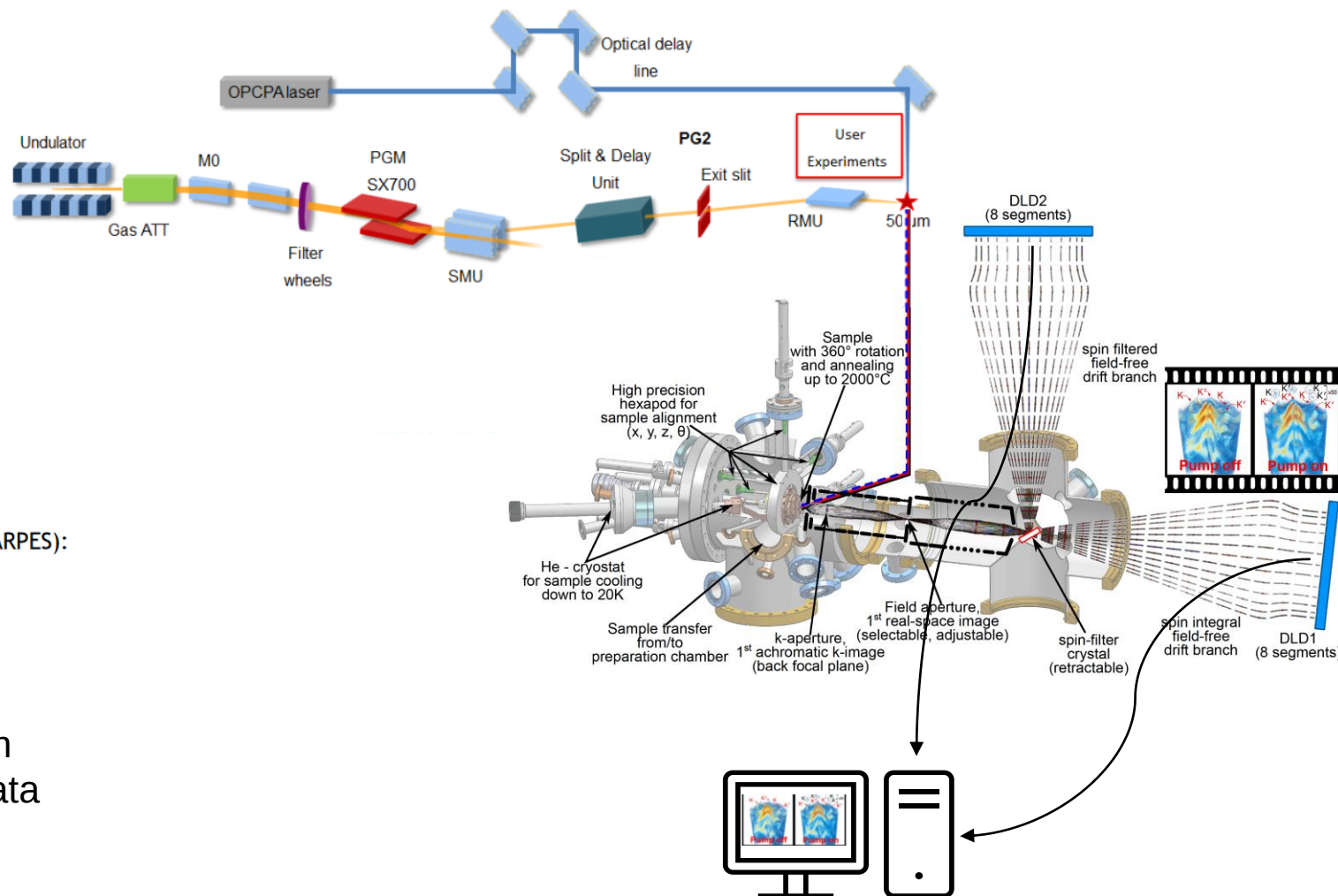
Multidimensional photoemission spectroscopy

Extending the parameter space

$$I \rightarrow I(\mu_{mat}, E, k_x, k_y, S, t, \omega_p, \sigma, x, y, T, \dots)$$

- ⌚ Dependence on materials and preparation procedures (μ_{mat})
- ⌚ Spin-resolving detectors: S
- ⌚ Photon energy and polarization ω_p
- ⌚ Electronic structure of non-equilibrium states (time-resolved ARPES): time t , pump/probe photon energy ω_p , polarization, ...
- à Dependence on sample parameters: strain σ , sample position x, y , temperature T
- ⌚ ...

Goal to develop flexible, community-driven data + metadata format for such MPES data



M. Martins et al., *Rev. Scient. Instr.* **77**, 115108 (2006)

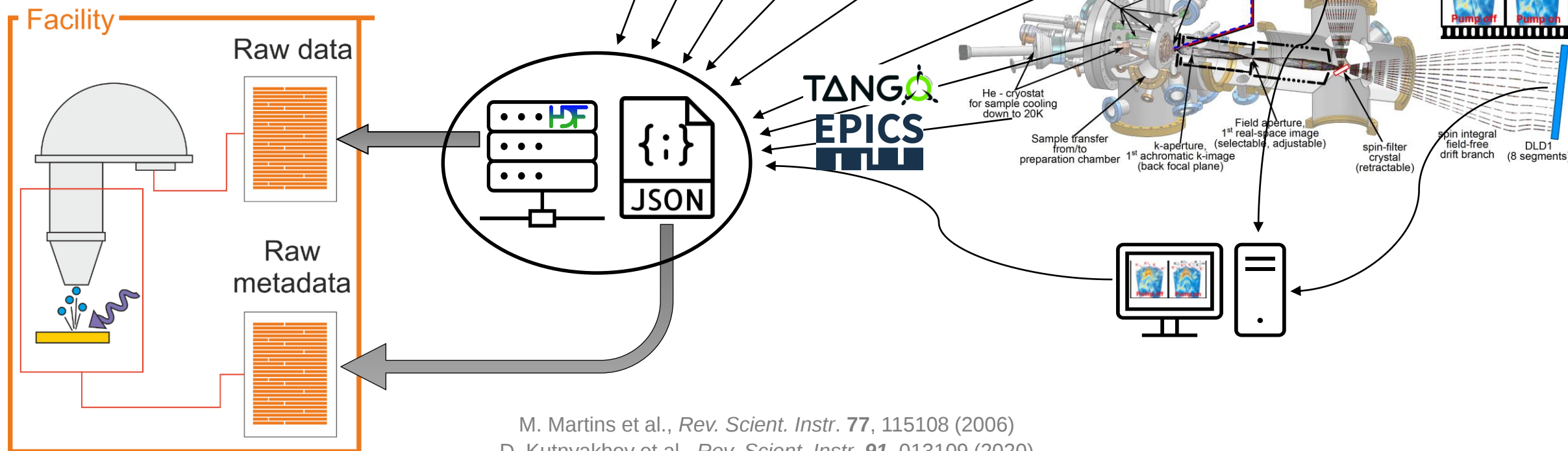
D. Kutnyakhov et al., *Rev. Scient. Instr.* **91**, 013109 (2020)

Data Acquisition at FLASH

>250 parameters registered:

- experiment data
- photon/beamline diagnostics
- Access to large parameter space

~30 MB/s data acquisition rate @5kHz



M. Martins et al., *Rev. Scient. Instr.* **77**, 115108 (2006)

D. Kutnyakhov et al., *Rev. Scient. Instr.* **91**, 013109 (2020)

NXmpes – application definition

- > HDF5 – entirely agnostic about what data is stored!
- > NeXus – adds meaning and set of rules on how data should be organized in a HDF5 file



NXmpes

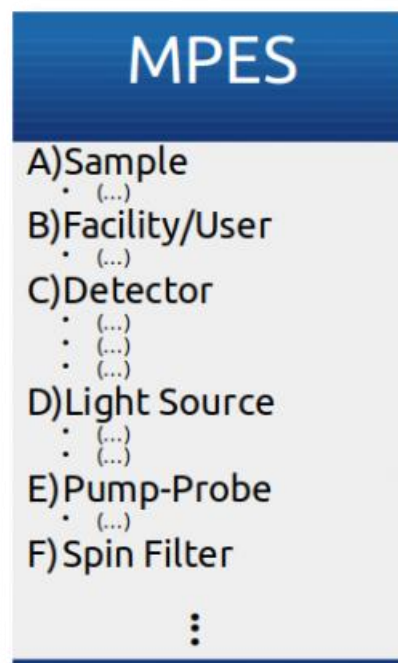
Status:

application definition, extends [NXObject](#)

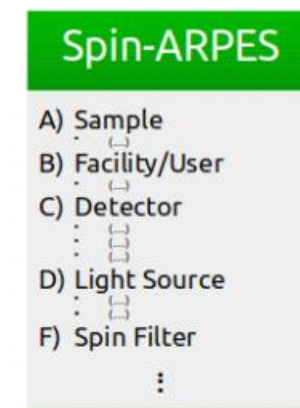
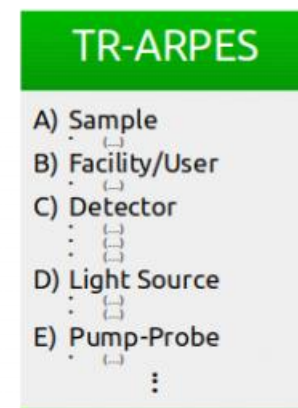
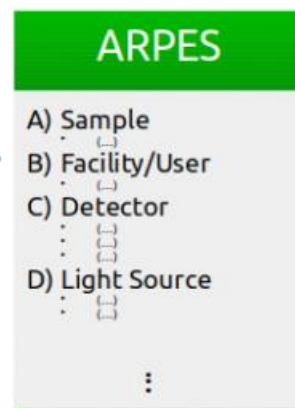
Description:

This is the most general application definition for multidimensional photoelectron spectroscopy.

Universal MPES metadata dictionary



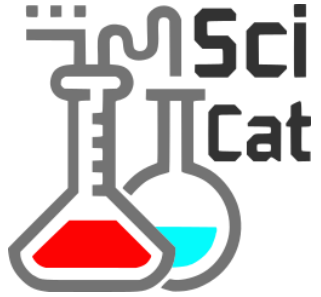
Experiment-specific examples
(in NeXus: *Application Definitions*),
using only the **relevant subset of metadata**
compiled from the universal MPES dictionary.



L. Rettig and T. Pincelli FHI Berlin, <https://mpes.science>

SciCat – test instance at DESY

<https://scicat-flash-test1.desy.de/>



Manage and annotate
your scientific data
and publish as open
data!



<https://scicatproject.github.io/documentation/>

SciCat ingestor

scicatIngestor_demo Last Checkpoint: vor einem Tag (autosaved)

File Edit View Insert Cell Kernel Widgets Help

SciCatIngestor - demo

- ☒ get token
- ☒ get ingested proposal
- ☒ push dataset to scicat
 - ☒ create file list
 - ☐ send file list
- ☒ fetch dataset from scicat by PID
- ☐ delete dataset from scicat by PID --> AUTHORIZATION_REQUIRED
- ☒ save json/request also on disk
- ☒ patch dataset

```
In [1]: from src.flashscicat import SciCatIngestor
```

create SciCat ingestor instance

```
In [2]: proposal_id = "11016608" # beamtimeID
source_folder = '/asap3/fs-flash-o/gpfs/wespe/2022/data/11016608/'
json_dump_folder = '.dumps/'

scicat_ingestor = SciCatIngestor(proposal_id, source_folder)
scicat_ingestor.activate_json_dump(json_dump_folder)
# scicat_ingestor.deactivate_json_dump()
```

scicatIngestor_demo Last Checkpoint: vor einem Tag (autosaved)

File Edit View Insert Cell Kernel Widgets Help

push dataset

```
In [11]: run_ID = "demo_017"
scientific_metadata = {
    "length": {
        "value": "500",
        "unit": "meters"
    },
    "wavelength": {
        "value": "10.0",
        "unit": "nm"
    }
}

scicat_ingestor.push_dataset(run_ID, scientific_metadata)

Dataset demo_017 successfully ingested.
```

```
In [12]: print(scientific_metadata)

{'length': {'value': '500', 'unit': 'meters'}, 'wavelength': {'value': '10.0', 'unit': 'nm'}}
```

scicatIngestor_demo Last Checkpoint: vor einem Tag (autosaved)

File Edit View Insert Cell Kernel Widgets Help

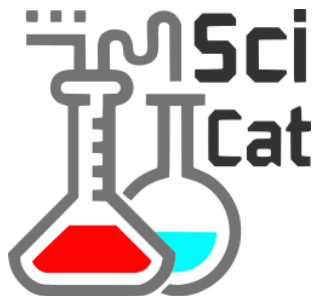
get ingested proposal

```
In [3]: scicat_ingestor.proposal

Out[3]: {'proposalId': '11016608',
'pi_email': 'stefan.duesterer@desy.de',
'pi_firstname': 'Stefan',
'pi_lastname': 'Duesterer',
'email': 'stefan.duesterer@desy.de',
'firstname': 'Stefan',
'lastname': 'Duesterer',
'title': 'Photon Science Group Proposal - FS-FLASH-O (Duesterer)',
'startTime': '2022-11-10T10:00:00.000Z',
'endTime': '2022-11-23T10:00:00.000Z',
'ownerGroup': '11016608-dmgt',
'accessGroups': ['11016608-dmgt',
'11016608-clbt',
'11016608-part',
'fsfl',
'f11dmgt'],
'createdBy': 'proposalIngestor',
'updatedBy': 'proposalIngestor',
'MeasurementPeriodList': [{'id': '11016608',
'instrument': '/fs-flash-o/wespe',
'start': '2022-11-10T17:39:31.000Z',
'end': '2022-11-25T09:11:29.000Z',
'comment': 'None',
'updatedBy': 'proposalIngestor',
'createdAt': '2022-11-10T17:39:31.382Z',
'updatedAt': '2022-11-25T09:11:30.024Z'}]},
'createdAt': '2022-11-10T17:39:31.382Z',
'updatedAt': '2022-11-25T09:11:30.027Z'}
```

Thanks to Christopher Passow!

SciCat – test instance at DESY



Manage and annotate
your scientific data
and publish as open
data!

<https://scicat-flash-test1.desy.de/>



Datasets / [/11013427/43463](#) /

Details

Datafiles

Lifecycle

Jupyter Hub

General Information

Name

43463

PID

/11013427/43463

Type

raw

Creation Time

2022-12-06 06:34

Keywords

Creator Information

Owner

Michael Martins

Principal Investigator

Matz Nissen

Contact Email

michael.martins@desy.de

Owner Group

11013427-dmgt

Access Groups

11013427-dmgt,11013427-clbt,11013427-part,fl1staff,fl1dmgt

File Information

Source Folder

/asap3/flash/gpfs/pg2/2022/data/11013427

Related Documents

Proposal

Time-resolved core-level photoelectron spectroscopy of graphitic carbon nitride, a catalyst for photocatalytic water splitting

Creation Location

/flash/pg2

Scientific Metadata

<https://scicatproject.github.io/documentation/>

Ingest metadata to SciCat and to eLog

scicat_logbook_pydocs_V1_WESPE

Last Checkpoint: vor 9 Stunden (autosaved)

File Edit View Insert Cell Kernel Widgets Help

Python 3 (jupyter)

In [45]:

```
#runnumber definition
# just use some number for tests
#DOocs_run_number=18
# the real DOocs DAQ run number
# here i would leave the pydocs.read
# if it can not read the run number something with the DAQ is wring ...
DOocs_run_number_all= pydocs.read('FLASH.DAQ/FLUSER3.RUN.CTR/DB.MGR/RC.RUN.NUMBER') #all the info
DOocs_run_number=DOocs_run_number_all['data'] # extract the runnumber
DOocs_run_start=try_pydocs_read('FLASH.DAQ/FLUSER3.RUN.CTR/DB.MGR/LAST.RUN.START')

# list of parameters inserted by hand
Morning_shift='Olina, Nils'
Late_shift='Steffen, Sebastian'
Night_shift='Lukas, Matz'
authors= Night_shift
#authors = Morning_shift + ', ' + Late_shift+Michael M.'
#authors = ', '.join(Morning_shift+Night_shift)
aim_of_run = 'FEL+Optical, O1s, D20'/'O1s, D20'
#comments = '1200l/mm grating, photon peak'
#comments = 'Collecting statistics'
#comments2 = ''
#comments = 'fixed delay, +40 ps'
#comments = 'short delay range scan around 10'
#sample_name = 'CNA_H2_flash_11'
Sample_name = 'TGCN_Flash_08'
Sample_chemical_name = 'g_C3N4'
Sample_position = [0.50,1.5,30.7,55.0] # test something with a list
Chamber_position = [398.0,243.00]
Sample_Temperature = 'RT'
Prep_pressure = '4.0e-8'
Sample_pressure = '6.6e-8'
Spectrum_Mode = ['LAD','LAD']#['DM','DM']
#Spectrum_Themis1001_param = ['Ekin = 180 eV','Epass = 100 eV','U_conv = 0 V','U_dld = 400 V','U_det = 1950 V','Iris = '']
Spectrum_Themis1001_param = ['Ekin = 85 eV','Epass = 30 eV','U_conv = -1 V','U_dld = 400 V','U_det = 1950 V','Iris = 30']
#Spectrum_Themis1001_param = ['Ekin = 342 eV','Epass = 30 eV','U_conv = -1 V','U_dld = 400 V','U_det = 1950 V','Iris = 30']
Spectrum_Themis1001_counts = 3500
Spectrum_Themis1002_param = ['Ekin = 85 eV','Epass = 30 eV','U_conv = -1 V','U_dld = 400 V','U_det = 1880 V','Iris = 30']
#Spectrum_Themis1002_param = ['Ekin = 342 eV','Epass = 30 eV','U_conv = -1 V','U_dld = 400 V','U_det = 1880 V','Iris = 34']
Spectrum_Themis1002_counts = 800

#PPLaser_DelayScan_param = ['fixed','Set=1284.7 ps','Encoder=1288.072 ps','','']
#PPLaser_DelayScan_param = ['step function','1323.7ps to 1325.5ps','1step','100bs','','']
#PPLaser_DelayScan_param = ['scan','1320.Sps to 1327.Sps','120steps','1s']
with open('/home/hlxss/DelayStage/delay_stage.json') as json_file:
    data = json.load(json_file)
PPLaser_DelayScan_param.json = [data['delay_start'], data['delay_stop'], data['number_of_steps'], data['waiting_time']]
PPLaser_Wavelength = 343
PPLaser_Powermeter = 0.10

#read DOocs parameters
# FEL pulse energy
rep_rate=try_pydocs_read('FLASH.DTAG/TIMER/FLASHCPUIME1.0/REP_RATE_KHZ.1')
Pulse_energy_tunnel=try_pydocs_read('FLASH.FEL/XGM.PHOTONFLUX/FL1.TUNNEL/PHOTONFLUX.U3')
Pulse_energy_BDA=try_pydocs_read('FLASH.FEL/XGM.PHOTONFLUX/FL1.BDA/PHOTONFLUX.U3')
Number_of_bunches=try_pydocs_read('FLASH.DTAG/TOROID/3GUN/BUNCHES.FLASH1')
Bunch_charge=try_pydocs_read('FLASH.DTAG/TOROID/3GUN/CHARGE.FLASH1')
Tunnel_aperture1=try_pydocs_read('TTF2.FEL/HASY.MOTOR/STPEPLTB/APERIVAL')
Tunnel_aperture2=try_pydocs_read('TTF2.FEL/HASY.MOTOR/STPEPLTB/APERIVAL')
FEL_wavelength=try_pydocs_read('FLASH.DAQ/ENERGY.DOGLG/LAMBDA_MEAN.VAL')
FEL_photon_energy=try_pydocs_read('TTF2.FEL/HASY.MONOCHR/MONOCHR/ENERGY')
FEL_cff=try_pydocs_read('TTF2.FEL/HASY.MONOCHR/MONOCHR/CF.INFO')
FEL_grating=try_pydocs_read('TTF2.FEL/HASY.MONOCHR/MONOCHR/LINESGRAT1')
FEL_dispersion=try_pydocs_read('TTF2.FEL/HASY.MONOCHR/MONOCHR/ENERGY.DISPERSION.PG2')
PG_WAU_V=try_pydocs_read('TTF2.FEL/HASY.WAU/NAU/VERHIDTH')
PG_WAU_H=try_pydocs_read('TTF2.FEL/HASY.WAU/NAU/HORHIDTH')
PG_Exitslit1=try_pydocs_read('TTF2.FEL/HASY.EXITSLIT/EXITSLIT/CURRENTWIDTH')
PG_Filterwheel1=try_pydocs_read('TTF2.FEL/BOAF1/BDA.F1/FW.CURRENT.DESC')
PG_Filterwheel2=try_pydocs_read('TTF2.FEL/BOAF2/BDA.F2/FW.CURRENT.DESC')
PG_Filterwheel3=try_pydocs_read('TTF2.FEL/PGFILTERS/PGFILTERS/FW.CURRENT.DESC')
PG_Filterwheel4=try_pydocs_read('FLASH.FEL/FILTER/PG2.UNIT4/FW.CURRENT.DESC')
PG_Filterwheel5=try_pydocs_read('TTF2.FEL/MOVESCREEN/MOVESCREEN/MCP.MOTOR.POS')
PG_GasAttenuator=try_pydocs_read('FLASH.FEL/GAS.DOSTING/FL1.ATTENUATOR/ACTIVESP')
# PP Laser parameters
PPLaser_Number_of_bunches=try_pydocs_read('FLASH.LASER/FLACUPGLASER1.PULSEENERGY/DTAGOUT/PULSECOUNT')
PPLaser_Transmission=try_pydocs_read('FLASH.LASER/FLASHMODPG1.FEEDFW/Attenuation_FL1GL/INPUT.V')
PPLaser_Filterwheel=try_pydocs_read('FLASH.LASER/FLACUPGLASER1.FLTHRM/ATT.WHEEL.POS')
PPLaser_Polarization=try_pydocs_read('FLASH.LASER/PGLASER/MOTOR2.PGLASER/PPDS')
PPLaser_EnergyMean=try_pydocs_read('FLASH.LASER/FLACUPGLASER1.PULSEENERGY/DTAGOUT/PULSEENERGY.MEAN')
PPLaser_PulseEnergy=try_pydocs_read('FLASH.LASER/FLACUPGLASER1.PULSEENERGY/PG2.Incoup1/PULSEENERGY.MEAN')
PPLaser_P1iter=try_pydocs_read('FLASH.SYNGLASER.LOCK.EXP/FL1.PG.DSC/CURRENT_INPUT_ZITTER.RD')
PPLaser_PulseSig = try_pydocs_read('FLASH.LASER/FLASHMODPG1.CAN/DTAG1.32.AC/ROD1_SPECTRUM.X.STG')
PPLaser_Shutter = try_pydocs_read('FLASH.LASER/TIMER/PGLASER1.RTH.TRG9.CONTROL')
if PPLaser_Shutter == 32783:
    PPLaser_State = 'OFF'
elif PPLaser_Shutter == 15:
    PPLaser_State = 'ON'
```

DOocs eLogBook

2023 52 02 10.01 09.01 01 2022 2021 2020 2019 2018 2017 2016 2015 news comments

View Current Logbook Search Logbook Help Your Feedback Example Link Printer: wespelog

DOocs eLogBook

FLASH Status: Laser heater commission last 8 hours: 0.37 nC; 74.1

News: Logbook server maintenance

06.12.2022 05:34 Lukas, Matz Run 43463 - FEL+Optical, O1s, D20

Sample: TGCN_Flash_08

Note: automatic elog print

Comment: Collecting statistics

Comment2:

Environment: P_prep = 4.0e-8 mbar P_main = 6.6e-8 mbar T_sample = RT

Sample Position: x = 0.5 y = 1.5 z = 30.7 omega = 55.0° Level1 = 398.0 Level2 = 243.0

Themis 1001 green LAD Ekin = 85 eV Epass = 30 eV U_conv = -1 V U_dld = 400 V U_det = 1950 V Iris = 30 counts: 3500

Themis 1002 blue LAD Ekin = 85 eV Epass = 30 eV U_conv = -1 V U_dld = 400 V U_det = 1880 V Iris = 30 counts: 800

FEL settings from pannel: bunches = 440 bunchcharge = 0.26 nC GMD_Tunnel = 5.48 μJ GMD_BDA = 5.97 μJ

Beamline settings: aperture1: 1.00 mm aperture2: 1.00 mm WAU (horiz.) 5.00 mm, (vert.) 0.50 mm

Gas-absorber = --- FW1 = OPEN FW2 = OPEN FW3 = OPEN FW4 = OPEN MCP_5 (I0) = 70.00

Mono: 3rd order exit slit = 50.0 grating = 1200 l/mm cff = 2.00 energy = 630.0834 eV Dispersion = 2.1155 eV/mm

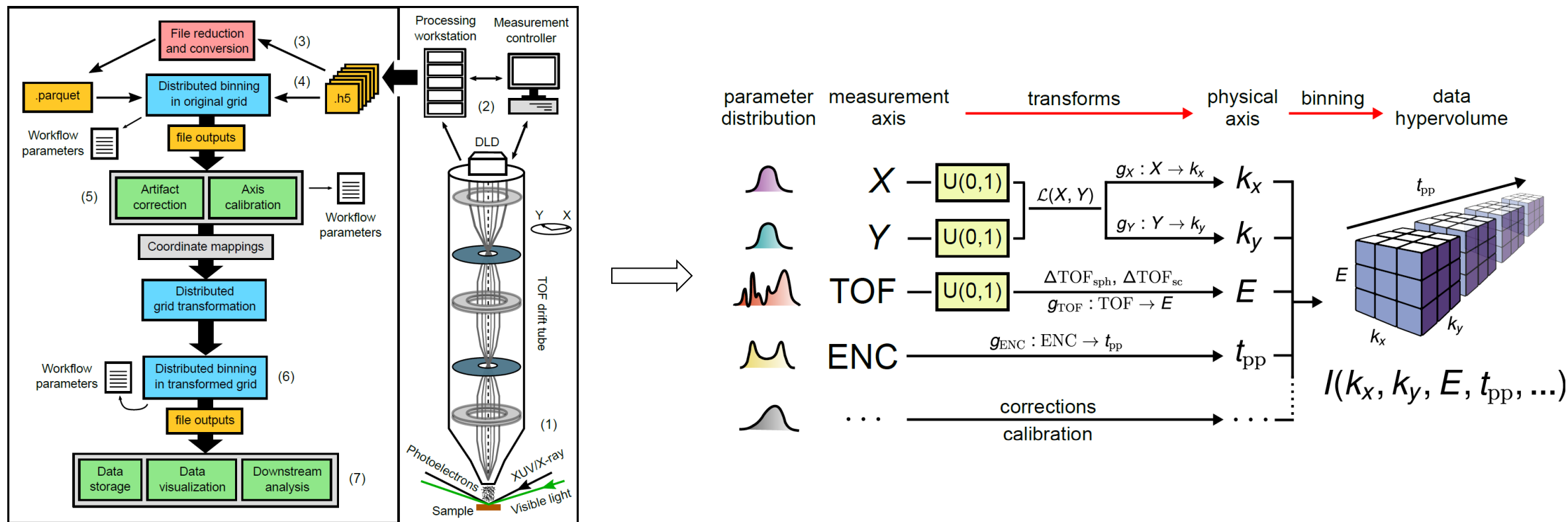
Scientific Metadata

Search

title	Run 43463 - FEL+Optical, O1s, D20
start_time	2022-12-06 05:34:34
run_number	43463
aim_of_run	FEL+Optical, O1s, D20
comments	Collecting statistics
Instrumentation	
instrument_name	WESPE
instrument_description	in preparation
instrument_position	
Source	
type	Free-Electron Laser
name	FLASH
probe	x-ray
center_wavelength	5.94 (nm)
photon_energy	630.0834 (eV)
pulse_energy	5.48 (μJ)
dispersion	2.1155 (eV/nm)
cff	2.00
repetition_rate	1003 (kHz)
Beamline	
WAU_H	5.00 (mm)
WAU_V	0.50 (mm)
exit_slit	50 (μm)
filter_wheel1	OPEN
filter_wheel2	OPEN
filter_wheel3	OPEN

Single event data processing

Single-event data calibrated and converted into multidimensional volumetric data.



Metadata from SciCat as data attributes



e.g. Run **46082** from Beamtime **11015619**

```
In [2]: # To load the file, we define the path of parquet files which includes the simulations
# The config file needs no changes
processed_file_path = "/asap3/flash/gpfs/pg2/2023/data/11015619/processed/converted/46082.1Q"
sp1Q_processed = SedProcessor(runs=[46082], config="N3_Dye_2023_1Q.yaml", collect_metadata=True, detector="1Q", load_parquet
```

```
In [4]: sp1Q_processed._attributes
```

```
Out[4]: {'principalInvestigator': 'Geoff Thornton', 'creationLocation': '/flash/pg2', 'scientificMetadata': {'title': 'Run 46082 - LAD mode, Ti2p, FEL', 'start_time': '2023-06-02 00:26:36', 'run_number': 46082, 'aim_of_run': 'LAD mode, Ti2p, FEL', 'comments': 'Static with 50 mum exit slit', 'Instrumentation': {'instrument_name': 'WESPE', 'instrument_description': 'in preparation', 'instrument_position': {'value': [400, 241.7], 'unit': ['mm', 'mm'], 'valueSI': [400, 241.7], 'unitSI': ['mm', 'm m']}}, 'Source': {'type': 'Free-Electron Laser', 'name': 'FLASH', 'probe': 'x-ray', 'center_wavelength': {'value': '5.74', 'unit': 'nm', 'valueSI': 5.7400000000000004e-09, 'unitSI': 'm'}, 'photon_energy': {'value': '642.9578', 'unit': 'eV', 'valueSI': 1.030131919443957e-16, 'unitSI': '(kg m^2) / s^2'}, 'pulse_energy': {'value': '20.57', 'unit': 'mJ', 'valueSI': '2 0.57', 'unitSI': 'mJ'}, 'dispersion': {'value': '2.1805', 'unit': 'eV/mm', 'valueSI': 3.4935459999825e-16, 'unitSI': '(kg m) / s^2'}, 'cfft': '2.00', 'repetition_rate': {'value': '1003', 'unit': 'kHz', 'valueSI': 1003000, 'unitSI': 's^-1'}, 'Beamline': {'WAU_H': {'value': '1.50', 'unit': 'mm', 'valueSI': 0.0015, 'unitSI': 'm'}, 'WAU_V': {'value': '0.50', 'unit': 'mm', 'valueSI': 0.0005, 'unitSI': 'm'}, 'exit_slit': {'value': 50, 'unit': 'um', 'valueSI': 50, 'unitSI': 'um'}, 'filter_wheel1': 'OPEN', 'filter_wheel2': 'OPEN', 'filter_wheel3': 'OPEN', 'filter_wheel4': 'Unknown', 'filter_wheel5': '70.00', 'gas_attenuator': '---'}, 'Laser': {'wavelength': {'value': 515, 'unit': 'nm', 'valueSI': 5.15e-07, 'unitSI': 'm'}, 'transmission': '0.20', 'filter_wheel': 5, 'pulse_energy': {'value': '0.00', 'unit': 'mJ', 'valueSI': '0.00', 'unitSI': 'mJ'}, 'pulse_duration': {'value': '68.3', 'unit': 'fs', 'valueSI': 6.83e-14, 'unitSI': 's'}, 'polarization': '-7.00', 'jitter': {'value': '50.53', 'unit': 'fs', 'valueSI': 5.0530000000000007e-14, 'unitSI': 's'}, 'delay_scan': [1483, 1486, 100, 1]}, 'Sample': {'sample': 'TiO2', 'sample_chemical_name': 'TiO2', 'sample_temperature': 'RT', 'sample_pressure': {'value': '6.8e-10', 'unit': 'mbar', 'valueSI': 6.8e-08, 'unitSI': 'kg / (m s^2)'}, 'prep_pressure': {'value': '5.2e-10', 'unit': 'mbar', 'valueSI': 5.1999999999999996e-08, 'unitSI': 'kg / (m s^2)'}, 'Manipulator': {'type': 'Rod', 'sample_position': [0, 1, 27, 55]}, 'Spectrometer': {'Spectrometer_Mode': ['LAD', 'LAD'], 'Themis1001_Parameters': ['Ekin = 169 eV', 'Epass = 20 eV', 'U_conv = 0 V', 'U_dld = 400 V', 'U_det = 1960 V', 'Iris = 18'], 'Themis1002_Parameters': ['Ekin = 169 eV', 'Epass = 20 eV', 'U_conv = 0 V', 'U_dld = 400 V', 'U_det = 1880 V', 'Iris = 32']}, 'runNumber': {'value': '46082'}, 'pid': '/11015619/46082', 'owner': 'Geoff Thornton', 'contactEmail': 'g.thornton@ucl.ac.uk', 'sourceFolder': '/asap3/flash/gpfs/pg2/2023/data/11015619', 'creationTime': '2023-06-02T00:28:01.000Z', 'type': 'raw', 'datasetName': 'Run 46082 - LAD mode, Ti2p, FEL', 'classification': 'IN=medium,AV=low,CO=low', 'version': '3.1.0', 'ownerGroup': '11015619-dmgt', 'accessGroups': ['11015619-dmgt', '11015619-clbt', '11015619-part', 'fl1staff', 'fl1dmgt'], 'createdBy': 'ingestor', 'history': [], 'datasetLifecycle': {'archivable': True, 'retrievable': False, 'publishable': False, 'archiveRetentionTime': '2033-06-01T00:00:00.000Z', 'dateOfPublishing': '2026-06-01T00:00:00.000Z', 'isOnCentralDisk': True, 'archiveStatusMessage': 'datasetCreated', 'retrieveStatusMessage': '', 'retrieveIntegrityCheck': False}, 'createdAt': '2023-06-01T22:28:01.538Z', 'updatedAt': '2023-06-01T22:28:01.554Z', 'proposalId': '11015619', 'techniques': []}
```

Parser to NeXus files

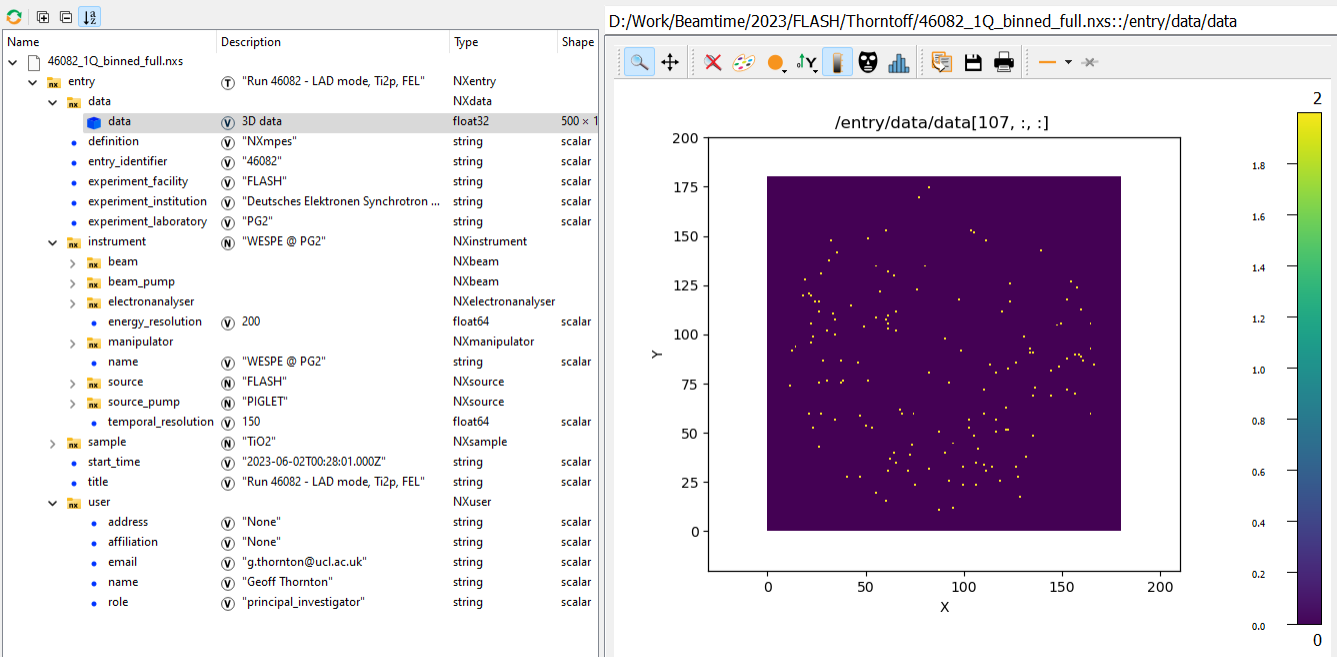


e.g. Run **46082** from Beamtime **11015619**

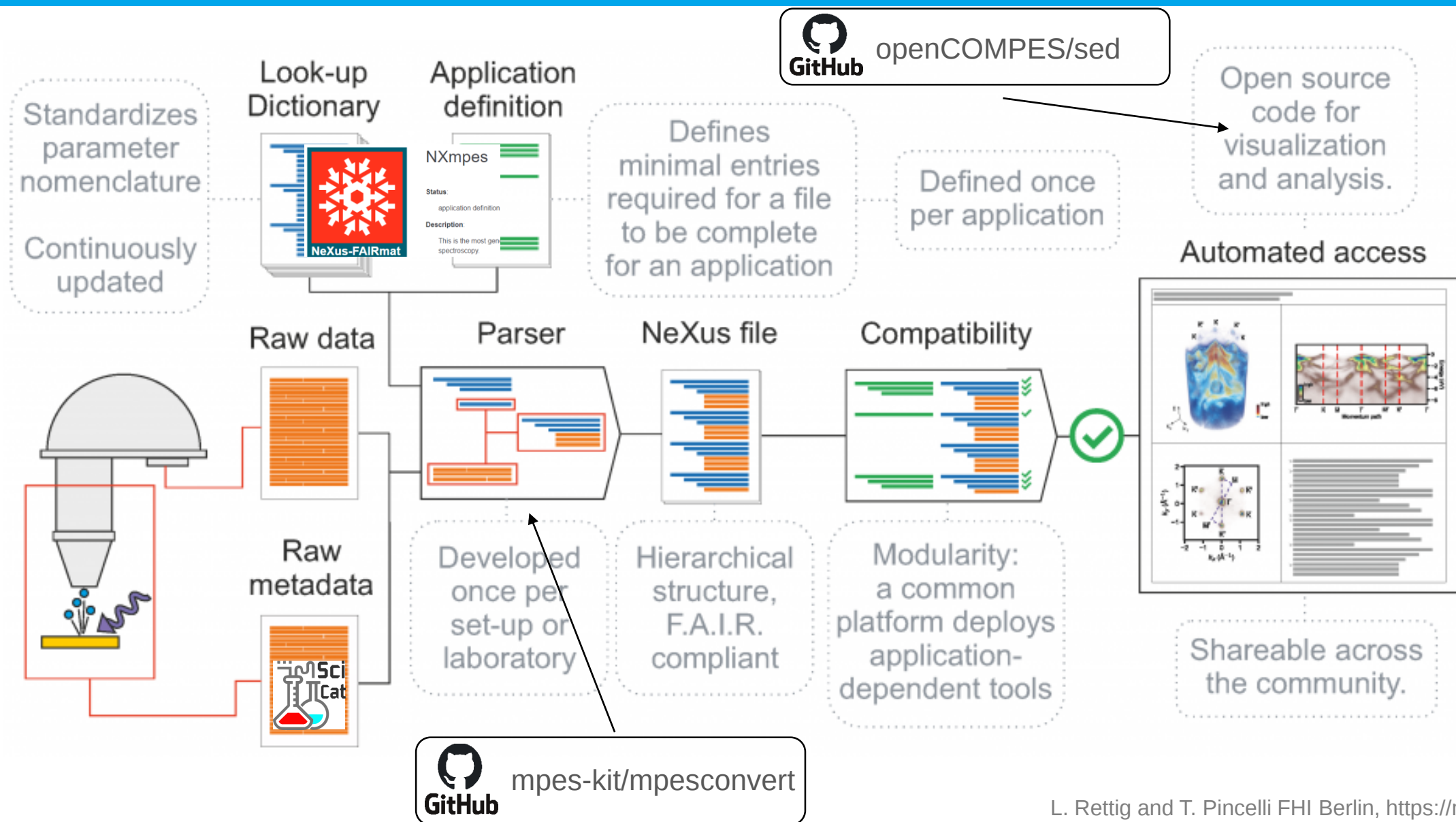
```
1 {
2   "/default": "entry",
3   "/ENTRY[entry]/@default": "data",
4   "/ENTRY[entry]/definition": "NXmpes",
5   "/ENTRY[entry]/definition/@version": "None",
6   "/ENTRY[entry]/title": "@attrs:metadata/scientificMetadata/title",
7   "/ENTRY[entry]/start_time": "@attrs:metadata/creationTime",
8   "/ENTRY[entry]/experiment_institution": "Deutsches Elektronen Synchrotron DESY - Helmholtz-Gemeinschaft",
9   "/ENTRY[entry]/experiment_facility": "FLASH",
10  "/ENTRY[entry]/experiment_laboratory": "PG2",
11  "/ENTRY[entry]/entry_identifier": "@attrs:metadata/scientificMetadata/run_number",
12  "/ENTRY[entry]/USER[user]/name": "@attrs:metadata/principalInvestigator",
13  "/ENTRY[entry]/USER[user]/role": "principal_investigator",
14  "/ENTRY[entry]/USER[user]/affiliation": "None",
15  "/ENTRY[entry]/USER[user]/address": "None",
16  "/ENTRY[entry]/USER[user]/email": "@attrs:metadata/contactEmail",
17  "/ENTRY[entry]/INSTRUMENT[instrument]/name": "WESPE @ PG2",
18  "/ENTRY[entry]/INSTRUMENT[instrument]/name@short_name": "WESPE",
19  "/ENTRY[entry]/INSTRUMENT[instrument]/energy_resolution": 200.0,
20  "/ENTRY[entry]/INSTRUMENT[instrument]/energy_resolution/@units": "meV",
21  "/ENTRY[entry]/INSTRUMENT[instrument]/temporal_resolution": 150.0,
22  "/ENTRY[entry]/INSTRUMENT[instrument]/temporal_resolution/@units": "fs",
23  "/ENTRY[entry]/INSTRUMENT[instrument]/ELECTRONANALYSER[electronanalyser]/description": "Double time-of-flight spectrometer equipped
with delay line detectors at the WESPE endstation of beamline PG2 of FLASH",
24  "/ENTRY[entry]/INSTRUMENT[instrument]/ELECTRONANALYSER[electronanalyser]/energy_resolution": 200.0,
25  "/ENTRY[entry]/INSTRUMENT[instrument]/ELECTRONANALYSER[electronanalyser]/energy_resolution/@units": "meV",
26  "/ENTRY[entry]/INSTRUMENT[instrument]/ELECTRONANALYSER[electronanalyser]/fast_axes": [
27    "x",
28    "y",
29    "e"
30  ],
31  "/ENTRY[entry]/INSTRUMENT[instrument]/ELECTRONANALYSER[electronanalyser]/slow_axes": "None",
32  "/ENTRY[entry]/INSTRUMENT[instrument]/ELECTRONANALYSER[electronanalyser]/COLLECTIONCOLUMN[collectioncolumn]/projection": "reciprocal",
33  "/ENTRY[entry]/INSTRUMENT[instrument]/ELECTRONANALYSER[electronanalyser]/COLLECTIONCOLUMN[collectioncolumn]/scheme": "Angular
dispersive",
34  "/ENTRY[entry]/INSTRUMENT[instrument]/ELECTRONANALYSER[electronanalyser]/COLLECTIONCOLUMN[collectioncolumn]/mode":
"@attrs:metadata/scientificMetadata/Spectrometer/Spectrometer_Mode",
35  "/ENTRY[entry]/INSTRUMENT[instrument]/ELECTRONANALYSER[electronanalyser]/COLLECTIONCOLUMN[collectioncolumn]/working_distance": 20.0,
36  "/ENTRY[entry]/INSTRUMENT[instrument]/ELECTRONANALYSER[electronanalyser]/COLLECTIONCOLUMN[collectioncolumn]/working_distance/@units":
"mm",
37  "/ENTRY[entry]/INSTRUMENT[instrument]/ELECTRONANALYSER[electronanalyser]/COLLECTIONCOLUMN[collectioncolumn]/field_aperture/shape":
"curved slit",
38  "/ENTRY[entry]/INSTRUMENT[instrument]/ELECTRONANALYSER[electronanalyser]/COLLECTIONCOLUMN[collectioncolumn]/field_aperture/size": 30,
39  "/ENTRY[entry]/INSTRUMENT[instrument]/ELECTRONANALYSER[electronanalyser]/COLLECTIONCOLUMN[collectioncolumn]/field_aperture
/size/@units": "mm",
40  "/ENTRY[entry]/INSTRUMENT[instrument]/ELECTRONANALYSER[electronanalyser]/ENERGYDISPERSION[energydispersion]/energy_scan_mode":
"fixed",
41  "/ENTRY[entry]/INSTRUMENT[instrument]/ELECTRONANALYSER[electronanalyser]/ENERGYDISPERSION[energydispersion]/pass_energy": "nan",
42  "/ENTRY[entry]/INSTRUMENT[instrument]/ELECTRONANALYSER[electronanalyser]/ENERGYDISPERSION[energydispersion]/pass_energy/@units": "eV",
43  "/ENTRY[entry]/INSTRUMENT[instrument]/ELECTRONANALYSER[electronanalyser]/ENERGYDISPERSION[energydispersion]/scheme": "tof",
44  "/ENTRY[entry]/INSTRUMENT[instrument]/ELECTRONANALYSER[electronanalyser]/ENERGYDISPERSION[energydispersion]/tof_distance": 0.9,
45  "/ENTRY[entry]/INSTRUMENT[instrument]/ELECTRONANALYSER[electronanalyser]/ENERGYDISPERSION[energydispersion]/tof_distance/@units": "m",
46  "/ENTRY[entry]/INSTRUMENT[instrument]/ELECTRONANALYSER[electronanalyser]/DETECTOR[detector]/amplifier_type": "MCP",
47  "/ENTRY[entry]/INSTRUMENT[instrument]/ELECTRONANALYSER[electronanalyser]/DETECTOR[detector]/detector_type": "DLD",
48  "/ENTRY[entry]/INSTRUMENT[instrument]/ELECTRONANALYSER[electronanalyser]/DETECTOR[detector]/sensor_pixels": [
49    "x",
50    "y",
51    "e"
52  ]
53 }
```

```
sp1Q_processed.save("46082_1Q_binned_full.nxs")

Using mpes reader to convert the given files:
• /home/kutnyakd/_beamtimes/N3_Dye_2023/NXmpes_config_WESPE.json
[info]: Path /ENTRY[entry]/PROCESS[process]/DISTORTION[distortion]/applied not found. Skipping the entry.
[info]: Path /ENTRY[entry]/PROCESS[process]/DISTORTION[distortion]/symmetry not found. Skipping the entry.
[info]: Path /ENTRY[entry]/PROCESS[process]/DISTORTION[distortion]/original_centre not found. Skipping the entry.
[info]: Path /ENTRY[entry]/PROCESS[process]/DISTORTION[distortion]/original_points not found. Skipping the entry.
[info]: Path /ENTRY[entry]/PROCESS[process]/DISTORTION[distortion]/cdeform_field not found. Skipping the entry.
[info]: Path /ENTRY[entry]/PROCESS[process]/DISTORTION[distortion]/rdeform_field not found. Skipping the entry.
[info]: Path /ENTRY[entry]/PROCESS[process]/energy_calibration/applied not found. Skipping the entry.
[info]: Path /ENTRY[entry]/PROCESS[process]/energy_calibration/coefficients not found. Skipping the entry.
[info]: Path /ENTRY[entry]/PROCESS[process]/energy_calibration/fit_function not found. Skipping the entry.
[info]: Path /ENTRY[entry]/PROCESS[process]/energy_calibration/original_axis not found. Skipping the entry.
[info]: Path /ENTRY[entry]/PROCESS[process]/energy_calibration/calibrated_axis not found. Skipping the entry.
The output file generated: 46082_1Q_binned_full.nxs
```



Final goal – publish data in NeXus format



Figures from: Xian, R. P. et al. arXiv:1909.07714 (2019).

L. Rettig and T. Pincelli FHI Berlin, <https://mpes.science>

Problems/issues/wishes

- Access metadata with own (not ingestor) credentials – like via gamma-portal.desy.de or directly from Maxwell
- Access SciCat from outside of DESY
- Publishing data with SciCat metadata – “derived” data type, DOI number
- How to deal with metadata of several Runs combined together?
- Beamtime summary table
- How to update existing metadata entry – “patch” dataset or?
- asapm to SciCat migration?

Acknowledgements



MAX-PLANCK-GESELLSCHAFT
Fritz-Haber-Institut

Ralph Ernstorfer
Laurenz Rettig
Tommaso Pincelli
Rui Patrick Xian
Maciej Dendzik



AARHUS UNIVERSITY

Philip Hofmann
Steinn Agustsson
Davide Curcio

ETH zürich

Yves Acremann
Kevin Bühlmann



Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG

FLASH.

Free-Electron Laser FLASH

Christopher Passow
Stefan Duesterer
Muhammad Zain Sohail
Michael Heber
Nils Wind
Lukas Wenthaus

Thank you for your attention!!!