



PaNOSC Closing Event

Paving the way towards the PaN FAIR Data Commons

29-30 November 2022

Grenoble - France

What PaNOSC changed for European XFEL

Luis Maia

European X-Ray Free Electron Laser



30th November, 2022



PaNOSC has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 823852

Main Achievements

- Scientific Data Policy review
- Open Proposals metadata publicly available via OAI-PMH and Loopback APIs
- Adoption of NeXus MX
- Libraries/Frameworks for Data Analysis
- Libraries/Frameworks for simulation
- Open proposals metadata, data and computing resources available for community at large
- Dissemination of information and trainings

Scientific Data Policy review

European XFEL Scientific Data Policy (2017)

Updated European XFEL Scientific Data Policy


FAIR concepts not mentioned	Explicitly mention FAIR and the objectives
Data Management Plan not mentioned	Define the Data Management Plan as mandatory
Reduced data not mentioned	Explicitly mention the possibility of storing reduced data as raw data
Processed data are not included in the long-term storage	Consider storing processed data for long-term
Some particular auxiliary data were defined (geometries, calibration constants)	Define general auxiliary data term, expand on this data and introduce rules to manage it
Electronic logbook not mentioned	Explicitly mention the electronic logbook as part of the metadata capture
ORCID not mentioned	Encourage usage ORCID as a means of linking users to data DOIs

European XFEL Scientific Data Policy (2017)

Updated European XFEL Scientific Data Policy

Data format is not explicitly mentioned	Define HDF5 as the preferred data format is HDF5
Only metadata from European XFEL software accepted	Allow the possibility of metadata from non-European XFEL software especially for processed data
Persistent identifiers (PIDs) mentioned in general	Explicitly define use of DOIs as PIDs
Granularity of PID is experiment and dataset	Explicitly mention PID can refer to a bespoke collection of datasets in addition to experiment and dataset automatically generated
Personal Data Protection included but GDPR not explicitly mentioned	Make explicit reference to the GDPR regulations
Few good practices included	Expand on the “Good practices” according to the new recommendations

Metadata publicly available (1)

 **myMdC** (<https://in.xfel.eu/metadata>) OAI-PMH [available](#) and DOI automatic generation and publishing [available](#)

This XML file does not appear to have any style information associated with it. The document tree is shown below.

```
<?OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/ http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">
  <responseDate>2022-11-23T10:38:13Z</responseDate>
  <request>https://in.xfel.eu/metadata("verb"=>"Identify")</request>
  <Identify>
    <repositoryName>European XFEL Proposals and Datasets Repository</repositoryName>
    <baseURL>https://in.xfel.eu/metadata</baseURL>
    <protocolVersion>2.0</protocolVersion>
    <adminEmail>luis.maia@xfel.eu</adminEmail>
    <adminEmail>krzysztof.wrona@xfel.eu</adminEmail>
    <earliestDatestamp>2019-02-14T14:27:59Z</earliestDatestamp>
    <deletedRecordno/>deletedRecordno
    <granularity>YYYY-MM-DDThh:mm:ssZ</granularity>
  </Identify>
</OAI-PMH>
```

This XML file does not appear to have any style information associated with it. The document tree is shown below.

```
<?OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/ http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">
  <responseDate>2022-11-23T10:40:59Z</responseDate>
  <request>https://in.xfel.eu/metadata("identifier"=>"oai:xfel.eu:30", "metadataPrefix"=>"panosc", "verb"=>"GetRecord")</request>
  <GetRecord>
    <record>
      <header>
        <identifier>oai:xfel.eu:30</identifier>
        <datestamp>2022-10-27T09:59:47Z</datestamp>
        <setSpec>openaire_data</setSpec>
      </header>
      <metadata>
        <panosc:panoscType xmlns:panosc="http://scicat.ess.se/panosc/" xsi:schemaLocation="http://scicat.ess.se/panosc https://raw.githubusercontent.com/panosc-eu/fair-data-api/master/panosc.xsd">
          <panosc:id>10.22003/XFEL.EU-DATA-700000-00</panosc:id>
          <panosc:name>Example Data</panosc:name>
          <panosc:description>The European XFEL (EuXFEL) example data proposal contains experimental datasets from various original beam-times, currently covering the techniques of serial femtosecond crystallography (SFX), coherent diffraction imaging (single particle imaging, SPI), X-ray powder diffraction, small-angle X-ray scattering (SAXS) and X-ray photon correlation spectroscopy (XPCS). </panosc:description>
          <panosc:owner>Fabio Dall'Antonia</panosc:owner>
          <panosc:contributor>Luca Gelleisio</panosc:contributor>
          <panosc:embargoEndDate>2018-01-01T00:00:00+01:00</panosc:embargoEndDate>
          <panosc:startDate>2017-11-08T00:00:00+01:00</panosc:startDate>
          <panosc:path>https://in.xfel.eu/metadata/doi/10.22003/XFEL.EU-DATA-700000-00</panosc:path>
          <panosc:technique>coherent diffraction imaging, serial femtosecond crystallography, small angle x-ray scattering, x-ray photon correlation spectroscopy, x-ray powder diffraction</panosc:technique>
          <panosc:sampleName>Water, Lysozyme (201804 small crystals), Lithium Titanate, No Sample, Lysozyme, Silica 50nm, Vycor, Sucrose Solution 3% v/v, Potassium hexacyanoferrate(II) trihydrate, Cu foil, Xenon, 2-Co8_pt4_8fold - 30nm Pt cap, 1-Co10_Pt_6fold, Ni-20 MLs - b, Ni75-11 MLs-b, Bolometer</panosc:sampleName>
          <panosc:chemicalFormula>AGIFD, Calibration - Dark HG, Calibration - Dark LG, Calibration - Dark MG, Configuration Tests, Dark, Diffraction, Diffraction data, General, Sample, SAXS 500kHz // no pump laser, scattering, SFX Jetting, Single Particle Diffraction, Test DAQ, XGN, XGN calibration, XPCS</panosc:chemicalFormula>
          <panosc:isid>9074975744</panosc:isid>
          <panosc:wavelength/>
        </panosc:panoscType>
      </metadata>
    </record>
  </GetRecord>
</OAI-PMH>
```

Proposal no. 700000

DOI: 10.22003/XFEL.EU-DATA-700000-00
Proposal Number: 700000
Name: p700000
Title: Example Data
Abstract: The European XFEL (EuXFEL) example data proposal contains experimental datasets from various original beam-times, currently covering the techniques of serial femtosecond crystallography (SFX), coherent diffraction imaging (single particle imaging, SPI), X-ray powder diffraction, small-angle X-ray scattering (SAXS) and X-ray photon correlation spectroscopy (XPCS).
Beamtime 1: 2017-11-08 00:00:00 +0100 - 2017-12-31 23:59:59 +0100
Instrument: Example Data
Instrument Cycle: 201750
Principal Investigator: Luca Gelleisio
Main Proposer: Fabio Dall'Antonia
Local Contact: ---
Expected end of Embargo: 2018-01-01 00:00:00 +0100
Open data since: 2022-10-27 11:24:32 +0200
open data?: Yes

This proposal data is open

Would you like to get access to this proposal datasets?

Please contact us through the open.data@xfel.eu email address.

Thank you for visiting!

Metadata publicly available (2)

myMdC RESTful and PaNOSC Loopback APIs [available](#)

PaNOSC XFEL APIs 1.0.0 OAS3
panosc/swagger.yaml

See [Loopback documentation](#) for more details.

Note: gt, lt, and, or, between, like, nlike, like, nlike Operators in [Where Filter](#) are yet to be implemented
Contact the developer

Servers
https://in.xfel.eu/metadata/api - Production server (uses live data) Authorize

panosc_instruments Operations for Panosc Instruments

- GET /panosc/v1/instruments Get all the instruments
- GET /panosc/v1/instruments/{id} Get instrument by ID
- GET /panosc/v1/instruments/count Get count of the instruments

panosc_documents Operations for Panosc Documents

- GET /panosc/v1/Documents Get all the documents
- GET /panosc/v1/Documents/{id} Get document by ID
- GET /panosc/v1/Documents/count Get count of the documents

panosc_datasets Operations for Panosc Datasets

- GET /panosc/v1/Datasets Get all the datasets
- GET /panosc/v1/Datasets/{id} Get datasets by ID
- GET /panosc/v1/Datasets/count Get count of the datasets
- GET /panosc/v1/Datasets/{dataset_id}/files Get all the files in given dataset
- GET /panosc/v1/Datasets/{dataset_id}/files/count Get count of the files in dataset

panosc_techniques Operations for Panosc Techniques

- GET /panosc/v1/Techniques Get all the Techniques
- GET /panosc/v1/Techniques/{id} Get Technique by ID
- GET /panosc/v1/Techniques/count Get count of the technique

<https://in.xfel.eu/metadata/api-docs/index.html>

metadata APIs 0.0.0 OAS3

This is the myMdC API documentation
Contact the developer

Servers
https://in.xfel.eu/metadata/api - Production server (uses live data) Authorize

facilities Operations about the facilities

- GET /facilities Get all the facilities
- POST /facilities Create facility
- GET /facilities/{id} Get facility by ID
- PUT /facilities/{id} Update facility by ID
- DELETE /facilities/{id} Delete facility by ID

units Operations about the units

- GET /units Get all the available units
- POST /units Create a unit
- GET /units/{id} Get unit by ID
- PUT /units/{id} Update unit by ID
- DELETE /units/{id} Delete unit by ID

topics Operations about the topics

- GET /topics Get all the available topics
- POST /topics Create a topic
- GET /topics/{id} Get topic by ID
- PUT /topics/{id} Update topic by ID
- DELETE /topics/{id} Delete topic by ID

data_group_types Operations about the data group types

- GET /data_group_types Get all the data group types
- POST /data_group_types Create data group types
- GET /data_group_types/{id} Get data group type by ID
- PUT /data_group_types/{id} Update data group type by ID
- DELETE /data_group_types/{id} Delete data group type by ID

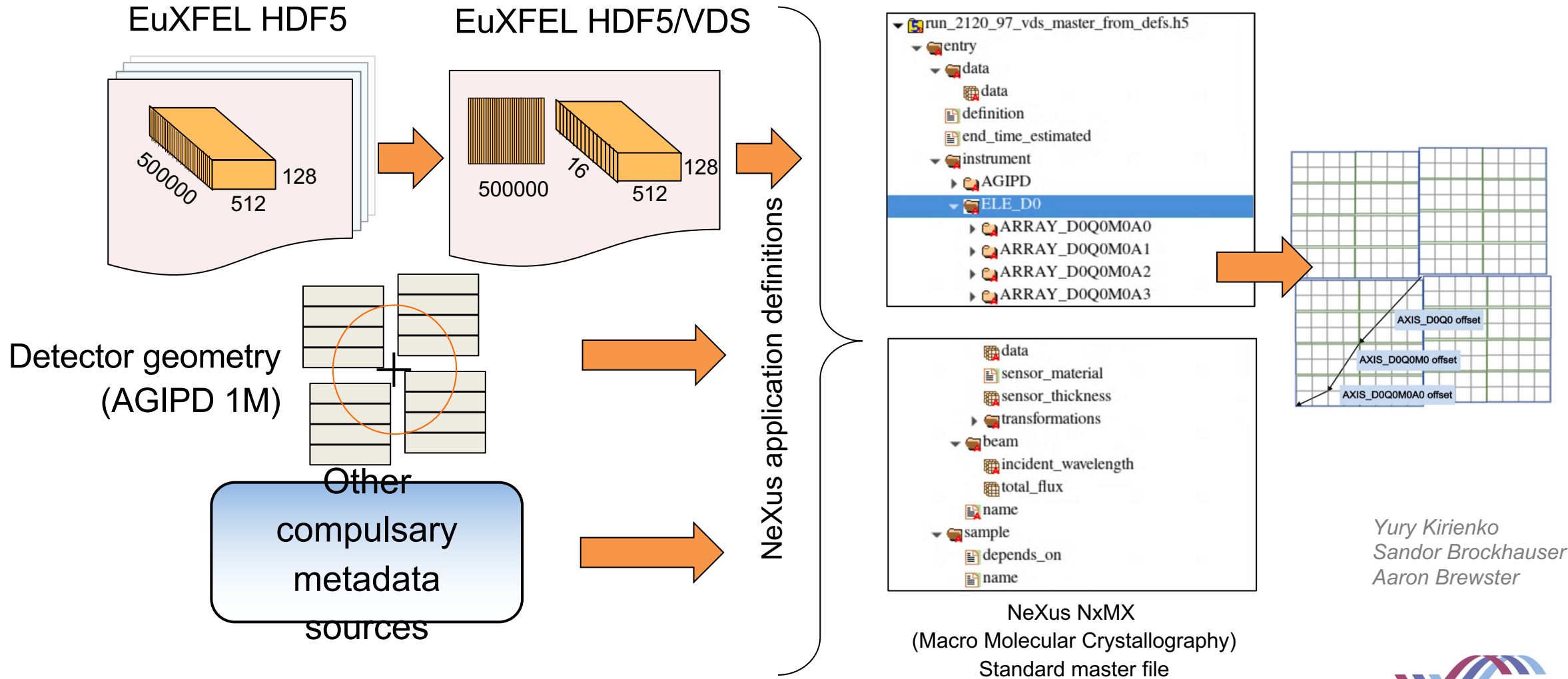
data_types Operations about the data types

- GET /data_types Get all the data types
- POST /data_types Create data type
- GET /data_types/{id} Get data type by ID
- PUT /data_types/{id} Update data type by ID
- DELETE /data_types/{id} Delete data type by ID

experiment_types Operations about the experiment types

- GET /experiment_types Get all the experiment types
- POST /experiment_types Create experiment type

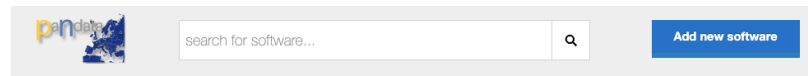
Adoption of NeXus MX, SFX use case



Yury Kirienko
Sandor Brockhauser
Aaron Brewster

Libraries/Frameworks for Data Analysis

Python packages EXtra-data and EXtra-geom



EXtra-geom

Python tools to work with geometry information for multi-module detectors at European XFEL, and to use assemble detector data into images.

Website <https://github.com/European-XFEL/EXtra-geom>

Licenses BSD 3-Clause "New" or "Revised" License (BSD-3-Clause)

Categories Data analysis Library

Software Requirements -

Hardware Requirements -

Platforms Mac OS Linux Windows

Languages Python

Input Formats There are no input formats associated to this software.

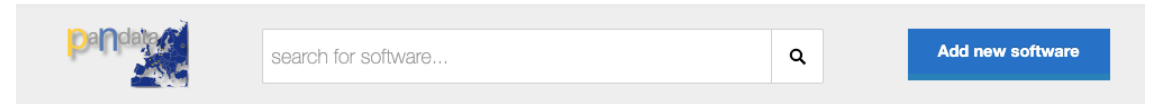
Output Formats There are no output formats associated to this software.

Contact email da-support@xfel.eu

How-to
There is not a how-to for this software.

Documentation / Tutorials

[EXtra-geom documentation](#)



EXtra-data

A Python library for accessing and inspecting data in European XFEL's HDF5 files.

European XFEL saves data in multiple HDF5 files with a moderately complex structure. EXtra-data aims to provide a simple interface to access data from a run directory, and conveniently work it in popular Python libraries such as Dask, Xarray and pandas.

Website <https://github.com/European-XFEL/EXtra-data>

Licenses BSD 3-Clause "New" or "Revised" License (BSD-3-Clause)

Categories Data analysis Library

Software Requirements -

Hardware Requirements -

Platforms Mac OS Linux Windows

Languages Python

Input Formats HDF5

Output Formats There are no output formats associated to this software.

Contact email da-support@xfel.eu

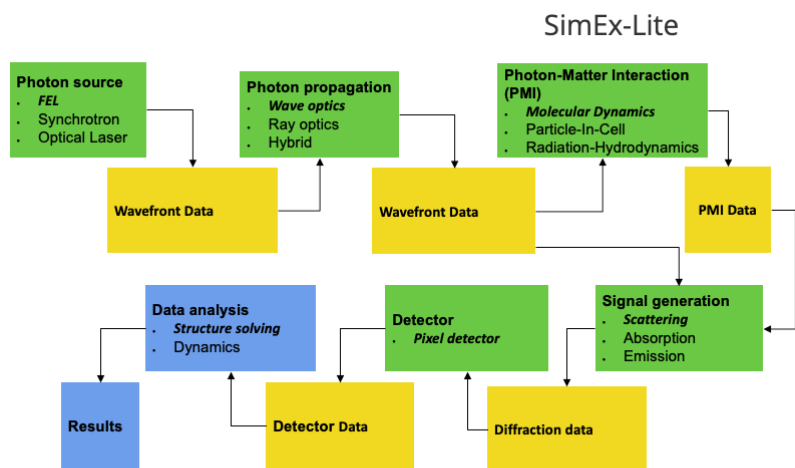
How-to
There is not a how-to for this software.

Documentation / Tutorials

[EXtra-data documentation](#)

Libraries/Frameworks for simulation

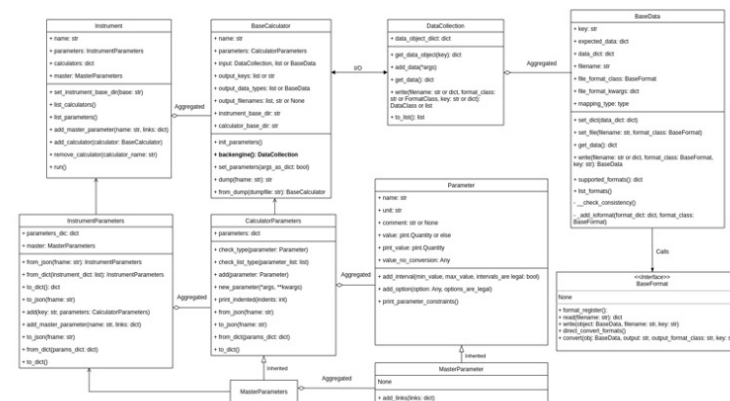
Python packages libpyvinyl and SimEx-Lite



Calculators
Data interfaces
<https://github.com/PaNOSC-ViNYL/SimEx-Lite>

- It is the core package of the SIMEX platform providing the calculator interfaces and data APIs.
- It is built based on **libpyvinyl**
- A **calculator** can be easily constructed within SimEx-Lite with the definition of the corresponding **data class** and the **format class**.
- Users can choose which back engine software to implement to make the installation minimal per needs.

libpyvinyl



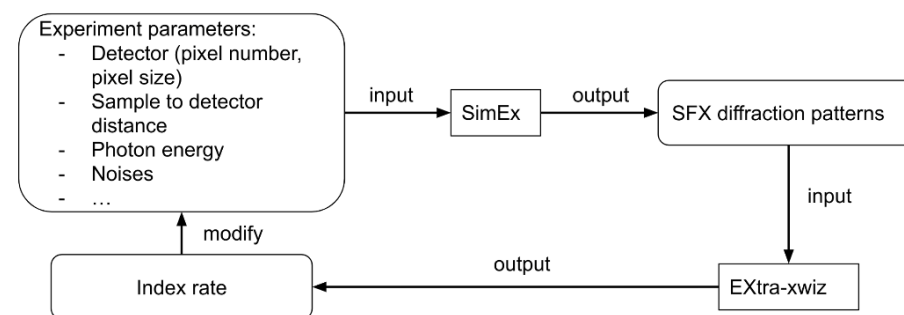
- Harmonize the user interfaces for **neutron** and **X-ray** simulation
- **Base classes:**
 - BaseCalculator
 - BaseData
 - BaseFormat
- **Auxiliary classes:**
 - Parameter
 - CalculatorParameters
 - DataCollection
 - Instrument

- It provides the developers of start-to-end simulation platforms in neutron and X-ray community with a framework to integrate various backengine software.
- It reduces the effort needed to integrate their software to a start-to-end simulation platform for simulation software developers.

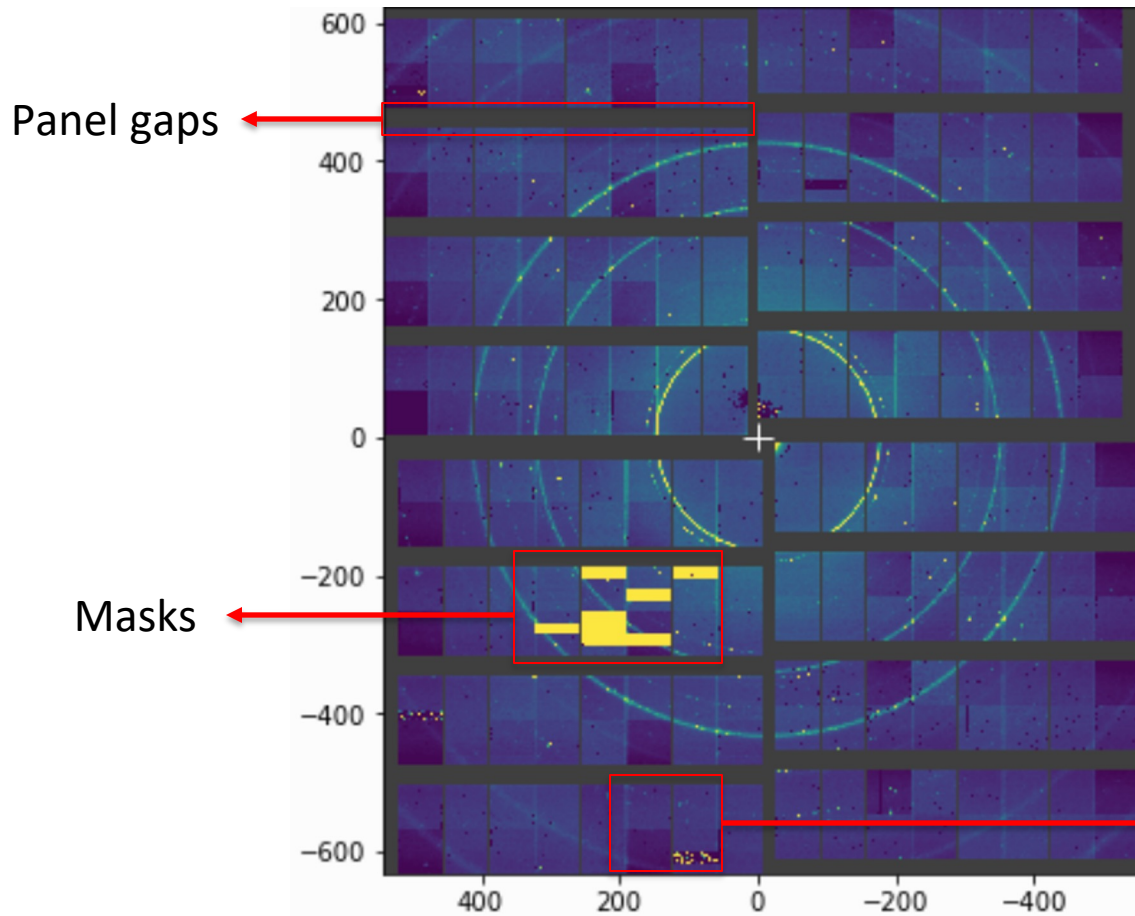
<https://github.com/PaNOSC-ViNYL/libpyvinyl>

A simple usage example: <https://github.com/PaNOSC-ViNYL/libpyvinyl/tree/master/tests/integration/plusminus>

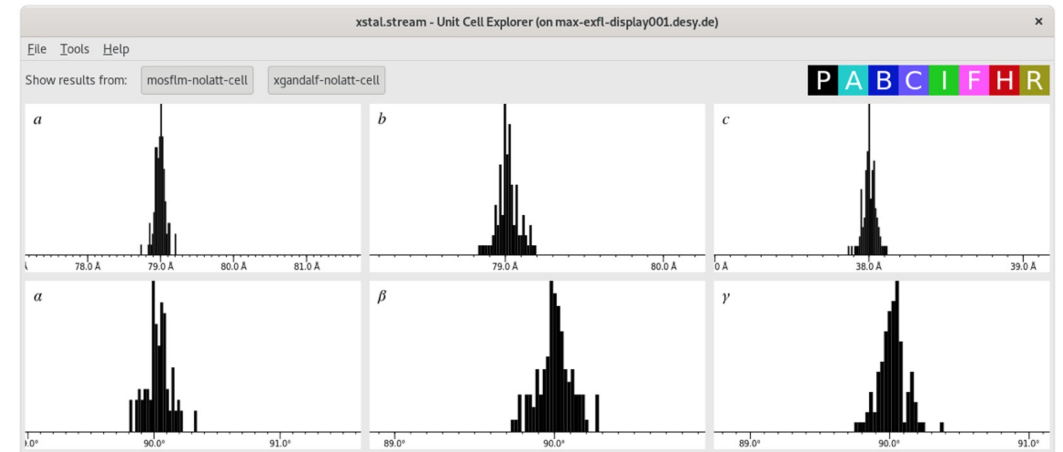
Optimize sample to detector distance in different energy densities



Serial Crystallography Simulation and Analysis



Demonstrate the effects of noise, panel gaps and masks on the results of crystallography analysis with SimEx simulation diffraction pattern results



Scenarios:

- New researchers want to get hands-on experience of data processing for serial crystallography.
- Starting from a working example of a simulation coupled to an analysis pipeline
- Explore how different experimental conditions and different analysis options affect the results.

Data open to Scientific Community (1)

- Visa portal visa.xfel.eu for community at large (e.g. EGI Check-in which includes Umbrella)



Data Analysis, in the cloud

VISA (Virtual Infrastructure for Scientific Analysis) makes it simple to create compute instances on the data analysis infrastructure to analyse your experimental data using just your web browser

[Sign in with your user account](#)

Analyse your data

Create a new compute instance and use your web browser to access a Remote Desktop or JupyterLab to start analysing your experimental data

Collaborate with your team

Share your compute instance with other members of your team to collaborate together in real time

No need to install software

The compute instances come with pre-installed data analysis software so you can start analysing your experimental data immediately



CLOUD APPLICATIONS AT DESY

English ▾

Sign in to your account

Username

Password

Remember me

[Sign in](#)

Or sign in with

GitHub

EGI Check-in

Helmholtz AAI



Compute instances

[CREATE A NEW INSTANCE](#)

What is VISA?

VISA (Virtual Infrastructure for Scientific Analysis) is a data analysis portal that allows you to create compute instances to analyse your experimental data. Once you have created a new instance, you can then access it remotely using only a web browser from anywhere in the world.



Looks like you don't have any instances

Why not create one?

[CREATE A NEW INSTANCE](#)



New compute instance

Search for experiments

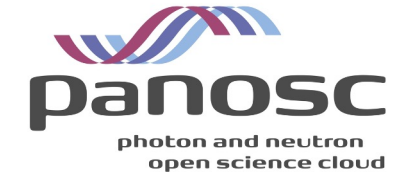
Search for your experiments using the filters below

Instrument All instruments ▾ between 2017 ▾ and 2021 ▾ with open data included ▾ sort by date (newest first) ▾

Proposal	Title	Instrument	Start Date	End Date	
p700000	SFX on sucrose solution, AGIPD detector at SPB instrument	XMPL	01 Jun 2021	01 Jun 2021	SELECT
p700000	SFX on Hen egg-white lysozyme, AGIPD detector	XMPL	15 Apr 2021	15 Apr 2021	SELECT
p700000	SAXS on vycor sample, AGIPD detector at MID instrument	XMPL	10 Apr 2021	10 Apr 2021	SELECT
p700000	pnCCD detector calibration with copper foil, SCS instrument	XMPL	14 Aug 2020	14 Aug 2020	SELECT
p700000	pnCCD detector calibration with copper foil, SCS instrument	XMPL	14 Aug 2020	14 Aug 2020	SELECT

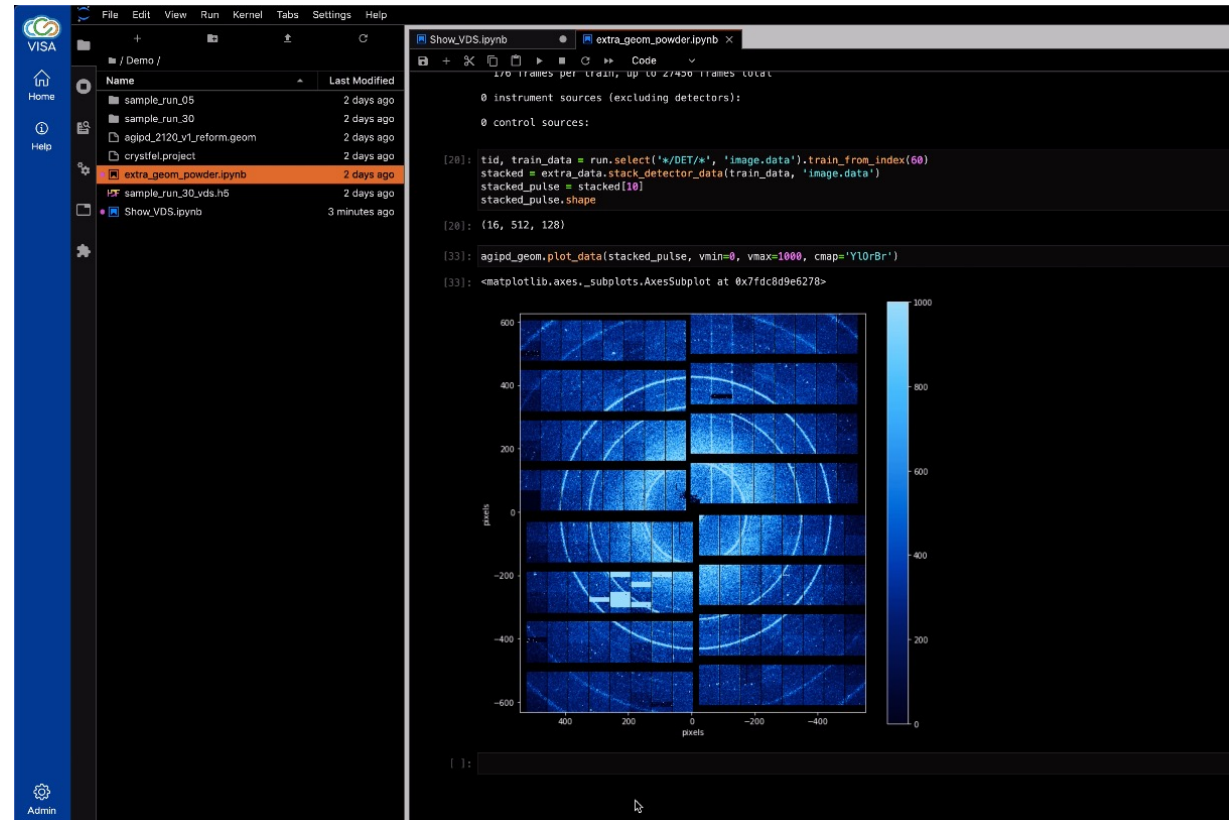
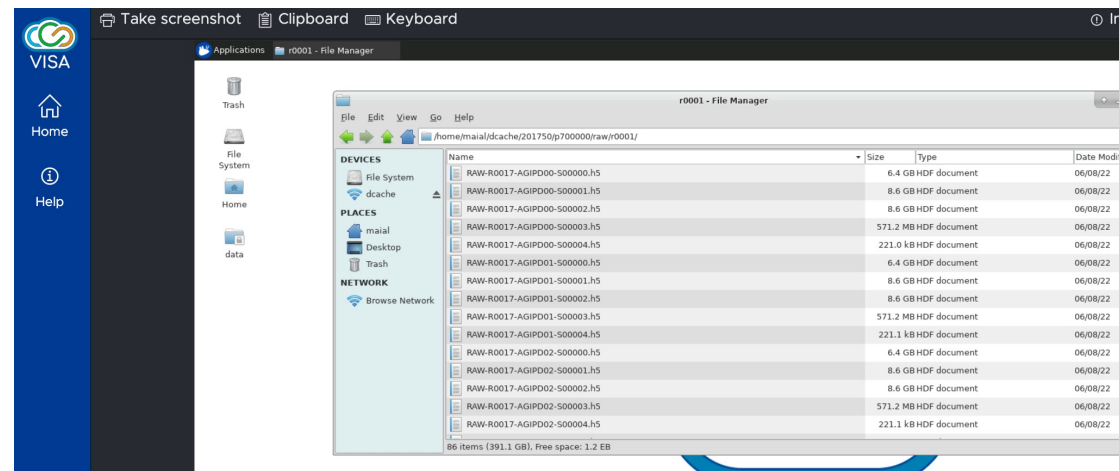

Results per page 5 ▾ 1 - 5 of 39 experiments < < 1 / 8 > >

Close



Data open to Scientific Community (2)

Visa portal allows access to proposals data



Dissemination of information and trainings

Production of scientific papers and trainings

Structural Dynamics

ARTICLE

scitation.org/journal/sdy

Expected resolution limits of x-ray free-electron laser single-particle imaging for realistic source and detector properties

Cite as: Struct. Dyn. 9, 064101 (2022); doi: 10.1063/4.0000169
Submitted: 7 September 2022 · Accepted: 31 October 2022 ·
Published Online: 16 November 2022



Juncheng E,¹ Y. Kim,¹ J. Bielecki,¹ M. Sikorski,¹ R. de Wijn,¹ C. Fortmann-Grote,^{1,2} J. Sztuk-Dambietz,¹ J. C. P. Kolyadu,¹ R. Letrun,¹ H. J. Kirkwood,¹ T. Sato,¹ R. Bean,¹ A. P. Mancuso,^{1,3,a)} and C. Kim^{1,b)}

AFFILIATIONS

¹European XFEL, Holzkoppel 4, 22869 Schenefeld, Germany

²Max Planck Institute for Evolutionary Biology, August-Thienemann-Straße 2, 24306 Plön, Germany

³Department of Chemistry and Physics, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, Victoria 3086, Australia

^{a)}Electronic mail: adrian.mancuso@xfel.eu

^{b)}Author to whom correspondence should be addressed: chan.kim@xfel.eu

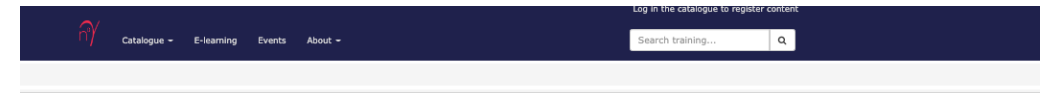
scientific reports

OPEN

Effects of radiation damage and inelastic scattering on single-particle imaging of hydrated proteins with an X-ray Free-Electron Laser

Juncheng E^{1,2,3}, Michal Stransky^{1,2,3}, Zoltan Jurek^{3,4}, Carsten Fortmann-Grote^{3,5}, Libor Juha^{6,7}, Robin Santra^{3,4,8}, Beata Zijia^{2,3,9} & Adrian P. Mancuso^{1,9,10}

Check for updates



European XFEL

The European XFEL in the Hamburg area is a new international research facility of superlatives: 27,000 X-ray flashes per second and a brilliance that is a billion times higher than that of the best conventional X-ray sources open up completely new opportunities for science. Research groups from around the world are able to map the atomic details of viruses, decipher the molecular composition of cells, take three-dimensional "photos" of the nanoworld, "film" chemical reactions, and study processes such as those occurring deep inside planets. European XFEL has a workforce of more than 350 employees and started user operation September 2017. With construction and commissioning costs of 1.25 billion euro (at 2005 price levels) and a total length of 3.4 kilometres, the European XFEL is one of the largest and most ambitious European new research facilities to date. At present, 12 countries have signed the European XFEL convention: Denmark, France, Germany, Hungary, Italy, Poland, Russia, Slovakia, Spain, Sweden, Switzerland, and the United Kingdom.

Materials (3) Events (0) Activity log

Showing 3 materials.

- Jupyter notebooks on Machine Learning for scientific data analysis
Jupyter Notebooks serving as supplementary material for a tutorial on Machine Learning, originally presented at the 2022 European XFEL user meeting.
Resource type: jupyter notebook
- A deep dive into the mathematics of Machine Learning
Crash course on Machine Learning
Keywords: machine learning
Resource type: slides
- Offline data analysis tutorial
Tutorial on European XFEL offline data analysis as per example of MID instrument data
Scientific topics: small angle x-ray scattering
Keywords: Extra-data, Extra-geom, pyFAT, data analysis
Resource type: jupyter notebook

17th Int. Conf. on Acc. and Large Exp. Physics Control Systems ISBN: 978-3-95450-209-7 ICALEPCS2019, New York, NY, USA JACoW Publishing doi:10.18429/JACoW-ICALEPCS2019-TUCPR02 ISSN: 2226-0358

DATA EXPLORATION AND ANALYSIS WITH JUPYTER NOTEBOOKS

H. Fangohr¹, M. Beg, M. Bergemann, V. Bondar, S. Brockhauser^{2,3}, C. Carinan, R. Costa, F. Dall'Antonia, C. Danilevski, J. C. E. W. Ehsan, S. G. Esenov, R. Fabbri, S. Fangohr, G. Flucke, C. Fortmann⁴, D. Fulla Marsa, G. Giovanetti, D. Goeries, S. Hauf, D. G. Hickin, T. Jarosiewicz⁵, E. Kamil, M. Karnevskiy, Y. Kirienko, A. Klimovskaia, T. A. Kluyver, M. Kuster, L. Le Guyader, A. Madsen, L. G. Maia, D. Mamchik, L. Mercadier, T. Michelat, J. Möller, I. Mohacs, A. Parenti, M. Reiser, R. Rosca, D. B. Rueck, T. Rüter, H. Santos, R. Schaffer, A. Scherz, M. Scholz, A. Silenzi, M. Spitzewski⁵, J. Sztuk, J. Szuba, S. Trojanowski⁵, K. Wrona, A. A. Yaroslavtsev, J. Zhu
European XFEL GmbH, Schenefeld, Germany
J. Reppin, F. Schlünzen, M. Schuh, DESY, Hamburg, Germany
E. Fernandez-del-Castillo, G. Sipo, EGI Foundation, Amsterdam, Netherlands
T. H. Rod, J. R. Selknaes, J. W. Taylor, ESS, Copenhagen, Denmark
A. Campbell, A. Götz, J. Kieffer, ESRF, Grenoble, France
J. Hall, E. Pellegrini, J. F. Perrin, ILL, Grenoble, France

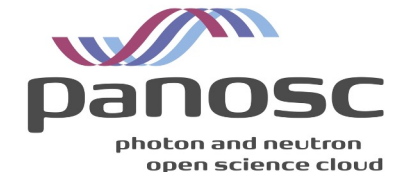
¹ also at University of Southampton, Southampton, United Kingdom

² also at University of Szeged, Szeged, Hungary

³ also at Biological Research Center of the Hungarian Academy of Sciences, Szeged, Hungary

⁴ also at Max-Planck-Inst. for Evolutionary Biology, Plön, Germany

⁵ also at NCBJ, Otwock, Poland



Acknowledgements



- Carsten Fortmann-Grote
- Deike Pahl
- Fabio Dall'Antonia
- Hans Fangohr
- Ivette Bermudez Macias
- Janusz Malka
- Janusz Szuba
- Julian Hörsch
- Juncheng E

- Krzysztof Wrona
- Luca Gelisio
- Luis Maia
- Oleksii Turkot
- Robert Rosca
- Sandor Brockhauser
- Sina Nolle
- Steve Aplin
- Sudhanshu Singh
- Tarcisio Fedrizzi
- Thomas Kluyver
- Thomas Tschentscher

- Yury Kirienko
- Varun Singh



- Humaira Abdul Salam
- Johannes Reppin
- Michael Schuh
- Patrick Fuhrmann
- Tim Wetzel



PaNOSC Closing Event

Paving the way towards the PaN FAIR Data Commons

29-30 November 2022

Grenoble - France

Thank you

luis.maia@xfel.eu



PaNOSC has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 823852