



PaNOSC Closing Event

Paving the way towards the PaN FAIR Data Commons

29-30 November 2022

Grenoble - France

Federated Search - Post Project Plans

Massimiliano Novelli, Tobias Richter
European Spallation Source, DMSC

29 November 2022



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



Connecting with third party EOSC data repositories

DOI Landing Page

Published Datasets / 10.17199/BRIGHTNESS/V200115 /

Publication Status

Status registered
Registered Time 2020-09-01, 16:22

General Information

Title Sample Data from V20

Abstract This data was collected as part of BrightnESS, funded by the European Union Framework Programme for Research and Innovation Horizon 2020, under grant agreement 676548. It consists of test data for the detector.

DOI 10.17199/BRIGHTNESS/V200115

URL doi.ess.eu/detail/10.17199/BRIGHTNESS/V200115

OpenAIRE EXPLORE Search Deposit Link Data sources Sign in

View all 124 versions
Research Data . Dataset . 2018

Sample Data from NMX

Pfeiffer, Dorothea;

OPEN ACCESS

DOI: 10.17199/brightness/nmx0016, 10.17199/brightness/nmx0042, 10.17199/brightness/nmx0065, 10.17199/brightness/nmx0018, 10.17199/brightness/nmx0126, 10.17199/brightness/nmx0006, ...

Published: 01 Jan 2018
Publisher: ESS

Summary

Abstract
This data was collected as part of BrightnESS, funded by the European Union \ Framework Programme for Research and Innovation Horizon 2020, under grant agreement 676548. It consists of test data for the detector.

Organization

ESS EUROPEAN SPALLATION SOURCE

ESS
The European Spallation Source ESS is a European Research Infrastructure Consortium (ERIC), a multi-disciplinary research facility based on the world's most powerful neutron...

read more

Social
Twitter
Facebook

Dataset

B2Find

Sample Data from V20

DOI
This data was collected as part of BrightnESS, funded by the European Union Framework Programme for Research and Innovation Horizon 2020, under grant agreement 676548. It consists of test data for the detector.

PaN Photon and neutron ...

Identifier

DOI https://doi.org/10.17199/BRIGHTNESS/V200177
Metadata Access https://scicat.esss.se/openaire/oai?verb=GetRecord&metadataPrefix=oai_datacite&identifier=10.17199/BRIGHTNESS/V200177

Provenance

Creator Richter, Tobias (ESS)
Publisher ESS
Publication Year 2018
Rights OpenAccess; info.eu-repo/semantics/openAccess
OpenAccess true

Representation

Resource Type Dataset
Size 155 MB
Version 1
Discipline Particles, Nuclei and Fields

Data Portal

The screenshot displays the panosc Data Portal interface. At the top left is the panosc logo. A search bar contains the text "COVID" and shows "33 documents found". On the left side, there are several filter sections: Facility (set to "all"), Technique (set to "Select a technique..."), Chemical Formula (empty), Incident Wavelength (set to "min", "max", "nm"), Incident Photon Energy (set to "min", "max", "eV"), Temperature (set to "min", "max", "K"), and Pressure (set to "min", "max", "Pa"). The main content area shows a search result for document ID "10.15151/ESRF-ES-436648953" with a score of 0.900. The title is "Insight into effects of COVID-19 on human heart and lung tissue at sub-cellular level". The abstract text describes the study of COVID-19 effects on heart and lung tissue using X-ray cryo-nanoholotomography. Below the abstract, there are metadata fields: Released (7. May 2024), Facility (European Synchrotron Radiation Facility), and Type (LS). A "Services" section lists VISA, Jupyter SLURM, and PaNdata Software Catalogue. A "Members" section lists Nina Kølln WITTIG, Dmitry KARPOV, and Alexandra JOITA PACUREANU.

<https://data.panosc.eu>

Federated search architecture

Data Portal

The screenshot shows the PaNOSC Data Portal interface. At the top, it says "European Photon and Neutron Open Data Search Portal". Below that is a search bar with the text "diffraction" and a magnifying glass icon. Underneath the search bar, it says "... or try one of these queries: diffraction, lung". The main content area contains text about the project's goals and a list of participating facilities: European Synchrotron Radiation Facility, European Spallation Source, Institut Laue-Langevin, MAX IV, Paul Scherrer Institut, Central European Research Infrastructure Consortium, and European XFEL. At the bottom, there is a paragraph about the mission of the PaN data search portal.

Federated Search

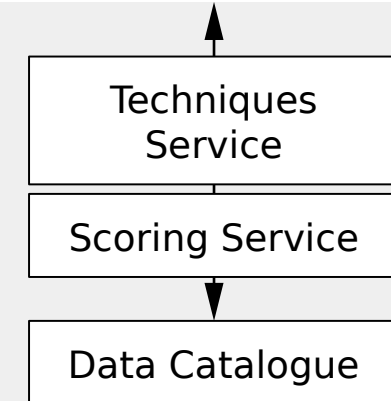
The screenshot shows the StrongLoop API Explorer interface for the "panosc-federated-search-api". It lists several endpoints under two main categories: "Dataset" and "Document". The "Dataset" endpoints include: /Datasets (Find all instances of the model matched by filter from the data source), /Datasets/{id} (Find a model instance by {id} from the data source), /Datasets/{id}/files (Queries files of Dataset), /Datasets/{id}/files/count (Counts files of Dataset), and /Datasets/count (Count instances of the model matched by where from the data source). The "Document" endpoints include: /Documents (Find all instances of the model matched by filter from the data source), /Documents/{id} (Find a model instance by {id} from the data source), and /Documents/count (Count instances of the model matched by where from the data source). There is also an "Instrument" endpoint: /Instrument (Beam line where experiment took place).

Facility Infrastructure

Facility Search

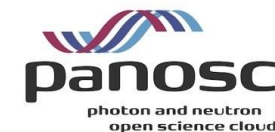
The screenshot shows the LoopBack API Explorer interface for the "panosc-search-api". It lists several endpoints under two main categories: "Dataset" and "Document". The "Dataset" endpoints include: /Datasets (Find all instances of the model matched by filter from the data source), /Datasets/{id} (Find a model instance by {id} from the data source), /Datasets/{id}/files (Queries files of Dataset), /Datasets/{id}/files/count (Counts files of Dataset), and /Datasets/count (Count instances of the model matched by where from the data source). The "Document" endpoints include: /Documents (Find all instances of the model matched by filter from the data source), /Documents/{id} (Find a model instance by {id} from the data source), and /Documents/count (Count instances of the model matched by where from the data source). There is also an "Instrument" endpoint: /Instrument (Beam line where experiment took place).

PaNOSC Infrastructure



ExPaNDS/PaNOSC meeting at SOLEIL:

- Presented achievements
- Presented future plan
- Hackathon
 - Demonstrate Data Portal and Federated Search
 - Discuss datasets curation to improve selected queries
- Future perspective



Future Plans from Facilities

- Full adoption of a catalogue system
- Ingest datasets in facility catalogue
- Improve curation datasets effort:
 - techniques, PaNOSC parameters, uniform metadata
- Maintain and improve PaNOSC scoring:
 - performance, incomplete words matching, pass-through words
- Active participation in support community
- Provide access to datasets not under embargo

Sustainability

- Technology – Code
 - Catalogues - local or collaborative project (SciCat/ICAT)
 - Common APIs - Connectivity
- Services
 - Local catalogues and their running are a facility responsibility
 - Who runs common or federated services?
 - Who is responsible for user support?
- Standards
 - Ontologies
 - Formats
 - Schemas

ExPaNDS and PaNOSC in the PaN EU project landscape



Photon (LEAPS)



Neutron (LENS)



Courtesy : LEAPS and LENS Web Pages

LENS & LEAPS Data Sub-Working Groups (under discussion)

LENS WG4

- Data processing and analysis
- **Open data**
- Artificial Intelligence, Machine Learning
- Atomistic Simulations



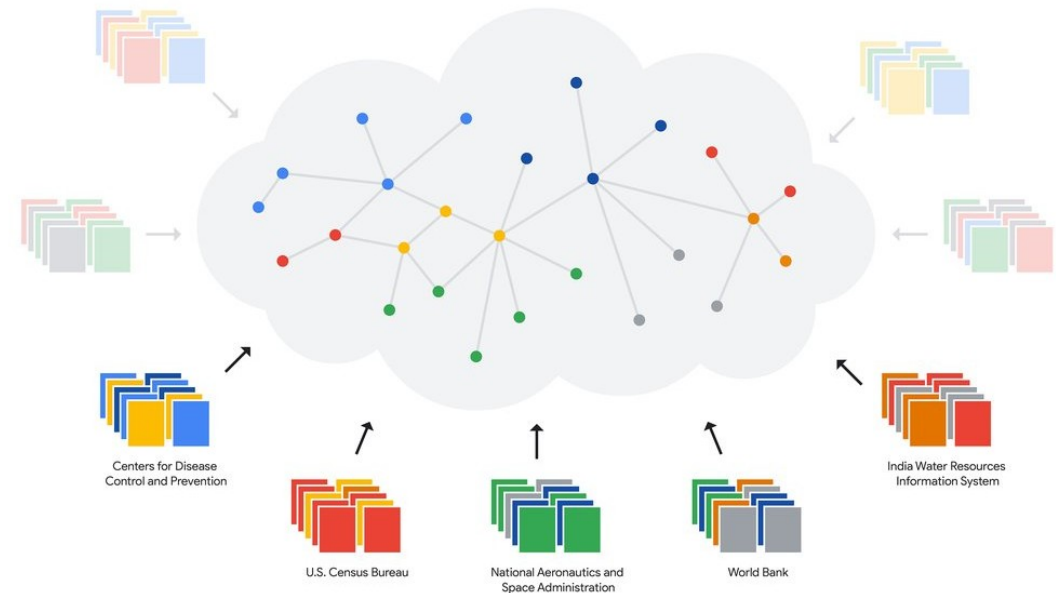
LEAPS WG3

- Data Infrastructure and Provision
- **Data Management and Processing**
- Data Analysis and Simulation



Connectivity - Application of FAIR Data

- Download and transfer services
- Visualization
- Processing and Analysis
- Provenance
- Automated or assisted ingest
- Upload of data
- Sample centric data tracking
- Curating data publications
+ linking with publishers
- Knowledge Graphs
- Connectivity of data
- Data accessibility for algorithms



How YOU can help:

- Work with your facility data curators to create a high quality data repository
- Engage with your PaN communities to standardize data formats, schemas and ontologies
- Reach out to connect PaN data repositories with application domain communities
- Use the PaN Search Portal and provide feedback
- Evangelize use FAIR data to young scientists





PaNOSC Closing Event

Paving the way towards the PaN FAIR Data Commons

29-30 November 2022

Grenoble - France

Thank you

Massimiliano Novelli (max.novelli@ess.eu)

Tobias Richter (tobias.richter@ess.eu)

Special thanks to Alun Ashton



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