DAta from PHoton and Neutron Experiments

Bridget Murphy
CAU - Kiel

C A U
Christian-Albrechts-Universität zu Kiel

DAPHNE 4NFDI

Bring together

- Universities
- Research institutions
- KFS-Committee for Synchrotron Research
- KFN-Committee for Neutron Research
- Large-scale photon and neutron facilities
- Wider community

DFG - NFDI

DAPHNE4NFDI

DAta from PHoton and Neutron Experiments

Nationale Forschungsdaten Infrastruktur

Consortium

Brings together 18 partners:

- University user groups
- Large scale facilities
- In addition: KFN + KFS
- > 60 participants (without funding)

Task Area Leaders:

Bridget Murphy TA1 (CAU, **Speaker**)
Anton Barty TA3, TA6 (DESY, **Speaker**)
Astrid Schneidewind TA4, TA5 (FZJ, **Dep. Speaker**)

Lisa Amelung (DESY, Coordinator)

Wiebke Lohstroh TA1 (TUM)
Sebastian Busch TA2 (Hereon)
Tobias Unruh TA2 (FAU)
Frank Schreiber TA3 (U Tübingen)
Jan-Dierk Grunwaldt TA4 (KIT)
Christian Gutt TA5 (U Siegen)

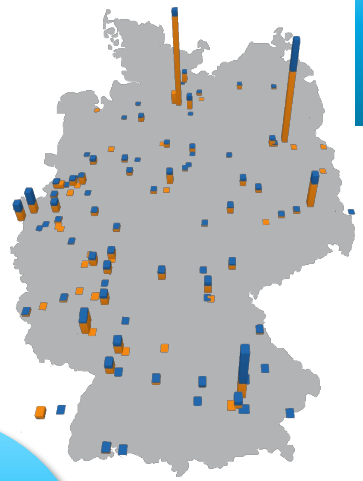
Consortium Proposal

www.DAPHNE4NFDI.de

Research with photons and neutrons in numbers

8 sources in Germany
33 sources in Europe
94 sources worldwide

per year in Germany



3000
participants at facility
user meetings

3000 experiments

28 PB data

3000 publications

5500 users

from 50 companies,
100 universities,
115 research institutions

Reaching a
community of
over 50 000

250 patents
In Europe

DAPHNE4NFDI aims

to make the growing volume of valuable measured data FAIR for the DAPHNE4NFDI community, for the whole NFDI and the scientific community.

These key objectives will be achieved within DAPHNE:

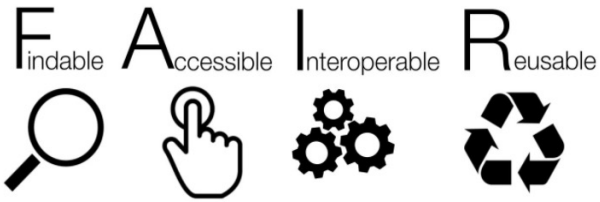
Reuse data by:

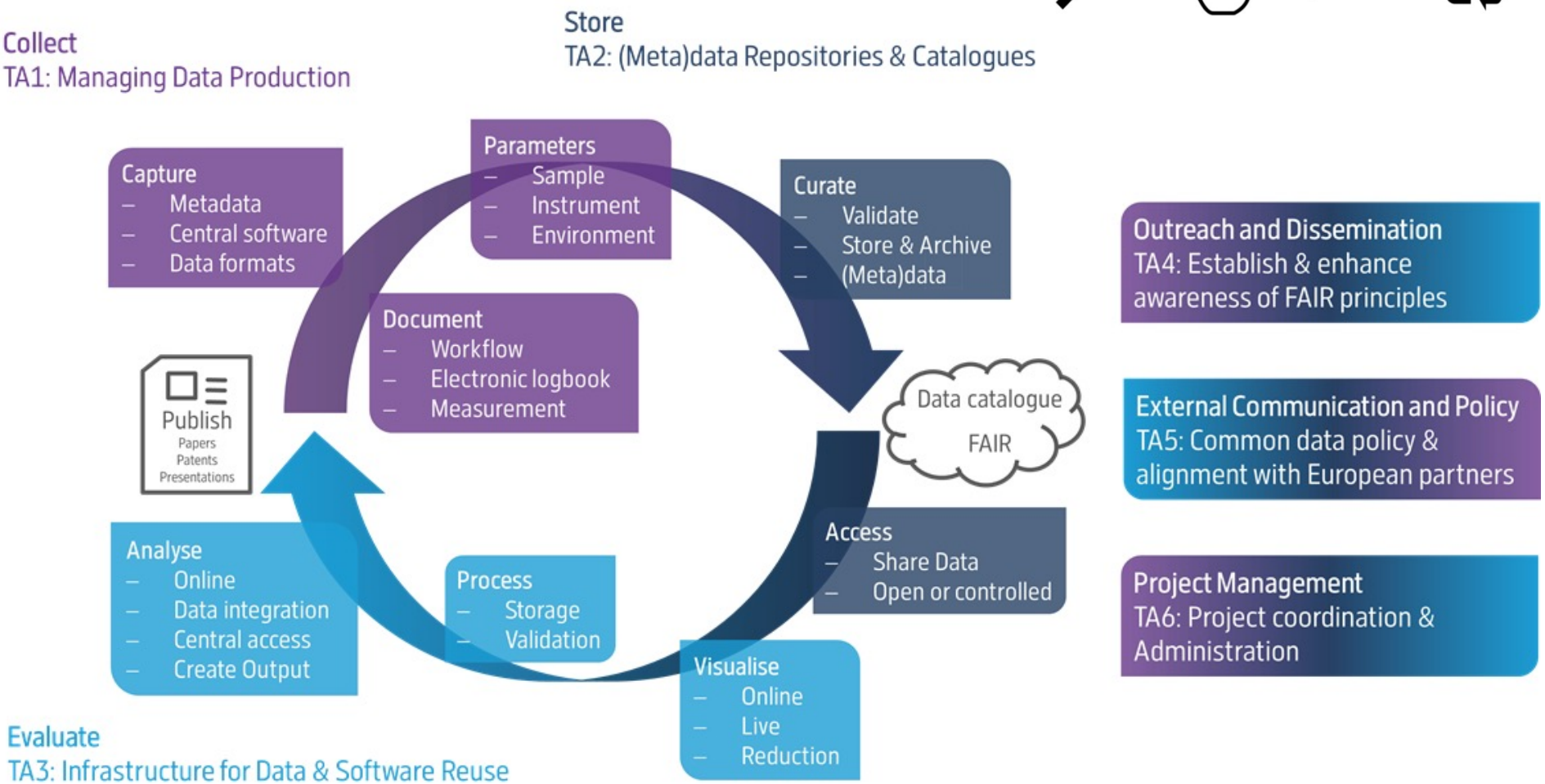
1. **Collection** of **(meta)data**
2. **Curated** repository **databases** of data and metadata
3. **Curated** repository of **managed software**
4. **Education** and **training** in research data management
5. **Connection** to other NFDI consortia, national and international organisations

Nationale Forschungsdaten Infrastruktur

DAPHNE4NFDI

DAPHNE4NFDI - DAta from PHoton and Neutron Experiments





Managing metadata collection

Enabling re-use and repeatability of results, ideally searchable

- Specification of **metadata standards** (whitepaper)
- Electronic log books (**ELN**) at facilities/labs (whitepaper)
- Link to Sample Persistent Identifier **IGSN**
- **Automatic capture** of: instrument, sample environment
- and sample data @experiment
- Data format (Nexus, OpenPMD)

ExPaNDS ontologies

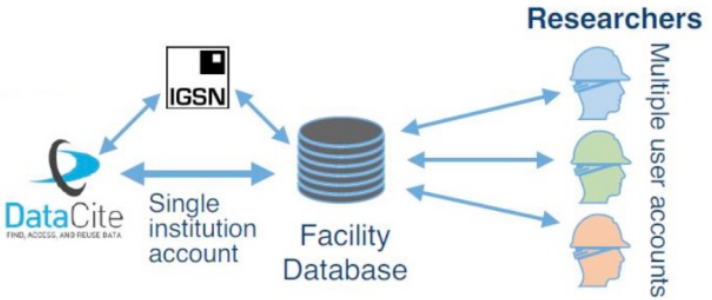
Document Control Information

Settings	Value
Document Identifier:	D3.2
Project Title:	ExPaNDS ontologies
Work Package:	WP3

https://zenodo.org/record/4806026#.YUh-kOdCRno

Recommendations on metadata capture and specifications

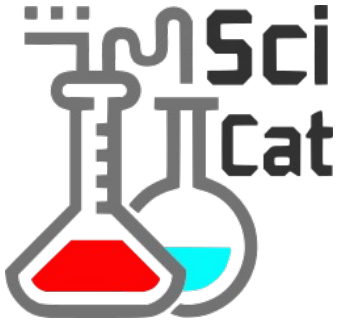
Collect

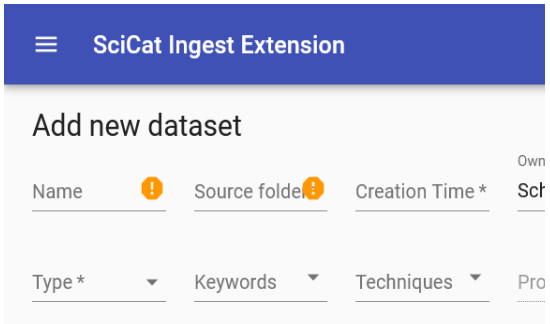


Community data repositories

A place to find published data - and in some cases the ability to reprocess data

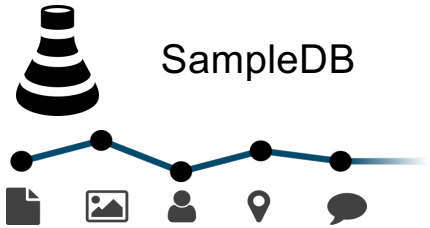
- Specification of **metadata standards**
- Sample description and sample **PID**
- **Repository** for processed and analysed data
- **Reference** data bases running at facilities and UNIs
- Public MetaData catalogue
  - https://public-data.desy.de



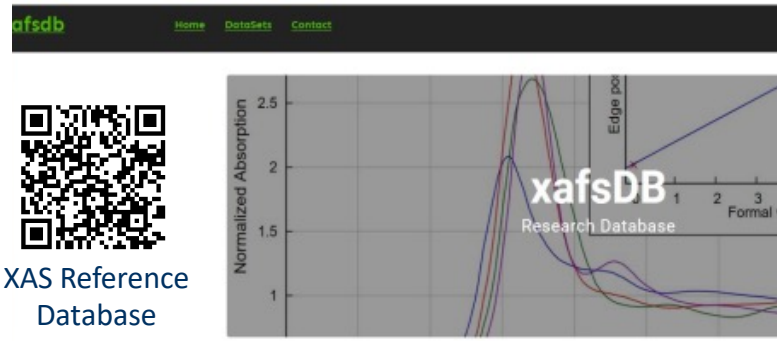


ICat

openBIS



Store

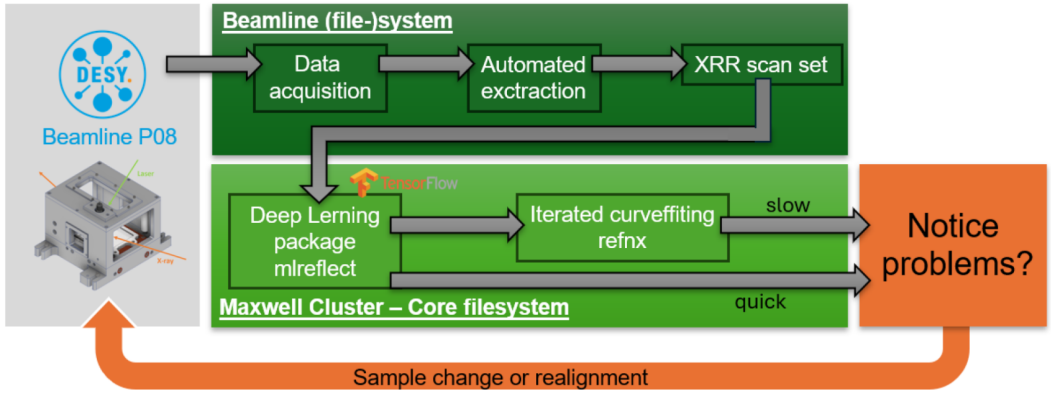


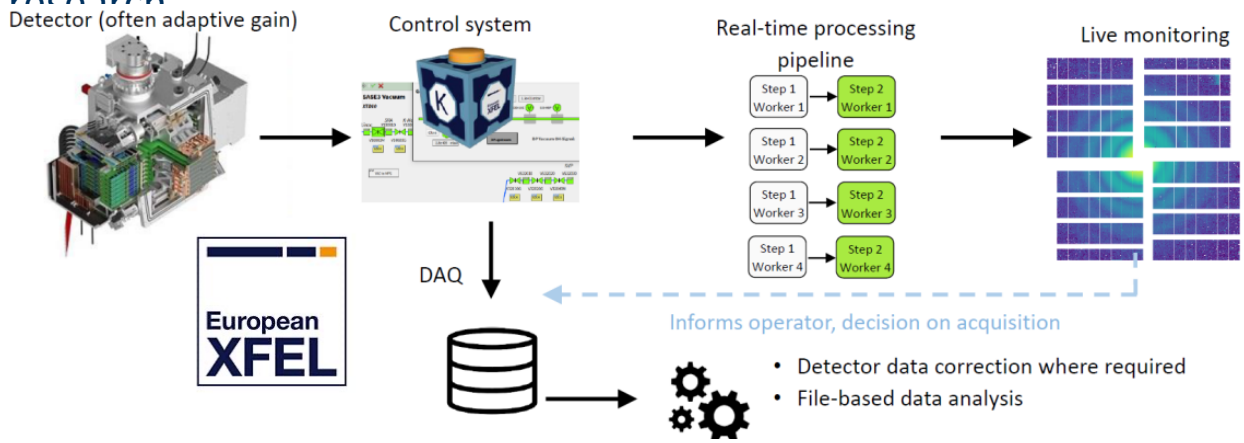
XAS Reference Database

Infrastructure for data and software re-use

Community data analysis software and data mining strategies including machine learning

- **Curate and deploy** analysis software on cloud like infrastructures
- Shared Gitlab
- **Automatise** data analysis chains
- Adoption of common **data models** e.g. by using NeXus
- **Teach** and practice sustainable research software development





Evaluate

L. Pithan et al. J. Synchrotron Rad. 30 (2023) 1064, L. Petersdorf et al. in prep.

9

Use Cases – DAPHNE as a Role Model and Educator

Reusable powder refinement
Neutron TOF diffraction
FZJ – MLZ – ESS - RWTH

Energy and battery materials, catalysis
Tomography
TUM – MLZ – BAM – Hereon – HZB - KIT

Amorphous materials for catalysis
x-ray absorption spectroscopy
KIT – TUB – Uni Wuppertal

Soft matter and liquid interfaces
x-ray reflectivity
Uni Kiel – Uni Tübingen

Proteins & Food science
Diffraction (small and wide angle) Spectroscopy
FAU– Uni Tübingen – EMBL - Uni Kiel

Biomaterials
x-ray imaging
LMU – Uni Göttingen

Dynamics
Correlation spectroscopy - XPCS
Uni Siegen - EuXFEL

Correlated electron systems
Spectroscopy
KIT – FZJ - MLZ

Magnetic structures
Ultrafast / Magnetic x-ray scattering
DESY – Uni Siegen

Chemical systems
x-ray emission spectra, RIXS etc.
KIT – ESRF - DESY

Catalysis and materials
High energy x-ray diffraction
HZDR – Uni Kiel - DESY

X-Ray Reflectivity use case

Collect

Store

Evaluate

Proposal
Pier reviewed

Documentation
electronic logbook

Sample

Raw data

Saved as .nxs

Metadata catalogue

Extraction

Fitting & Modelling

Software
Open Data
IGSN
Publication
X-ray studies of bidirectional structural changes induced by photoswitching in p… me

Persistent sample identifier service

IGSN **@ CAU**

www.igsn.uni-kiel.de

Service provided by CAU central computing
Thorge Petersen
Axel Wedemeyer
Thilo Paul-Stüve

Open data

Machine learning based online analysis
for closed loop experiment in cooperation with Uni Tübingen at ID10, ESRF and P08, PETRA III

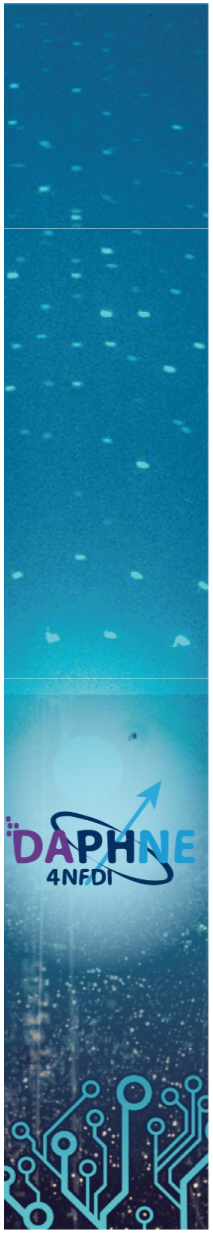
L. Pithan, J. *Synchr. Rad.* **30**, 1064 (2023)

Open data DB
Databank for XRR/NR data and their fits

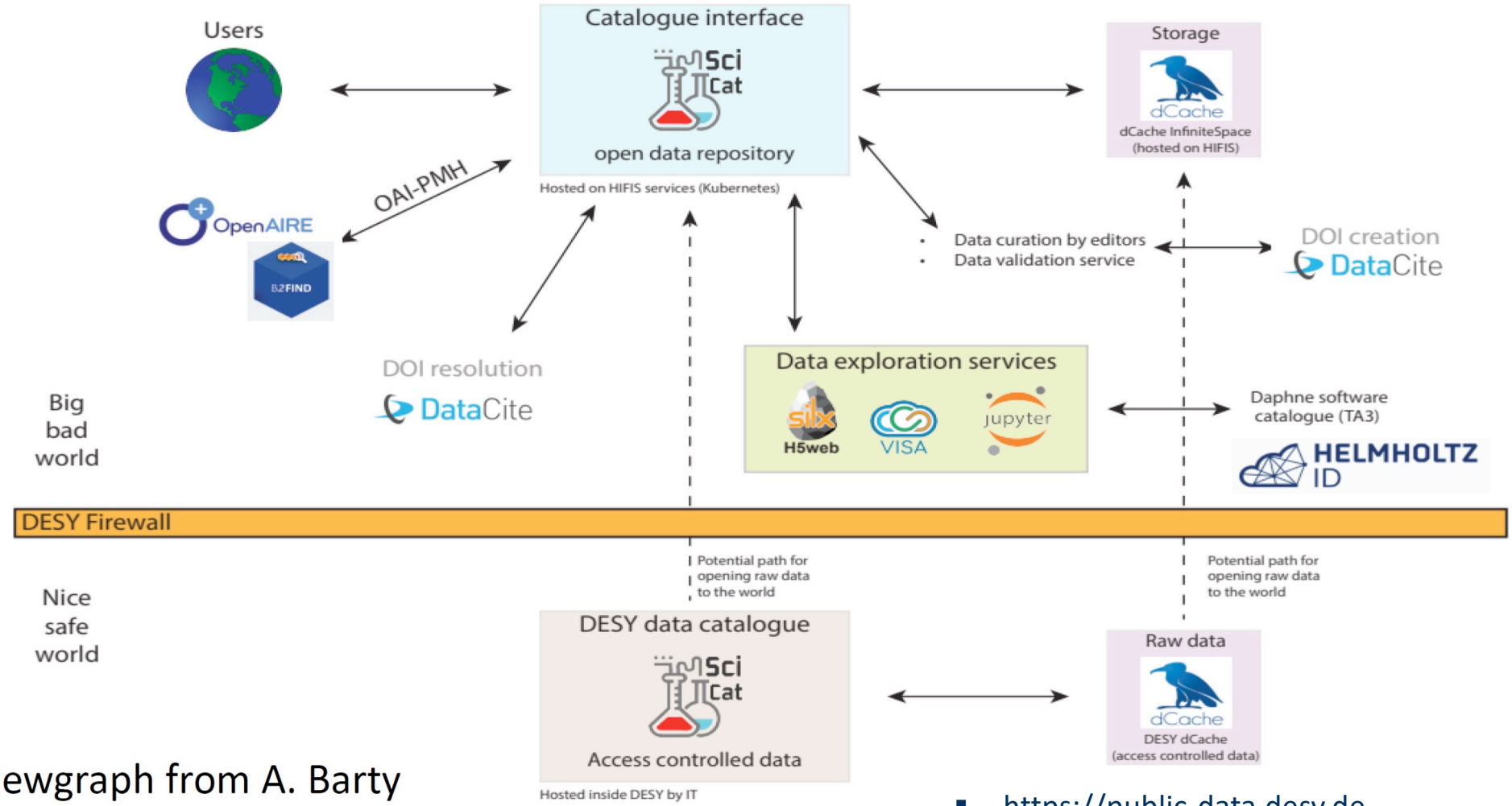
Example FAIR data publication

S. Hövelmann, *IUCr J* **11**, 4 (2024)

Findable Accessible Interoperable Reusable


DAPHNE
4NFDI

Raw-Data / Internal-Data / Public-Data Catalogues



viewgraph from A. Barty

- https://public-data.desy.de

NFDI Consortium

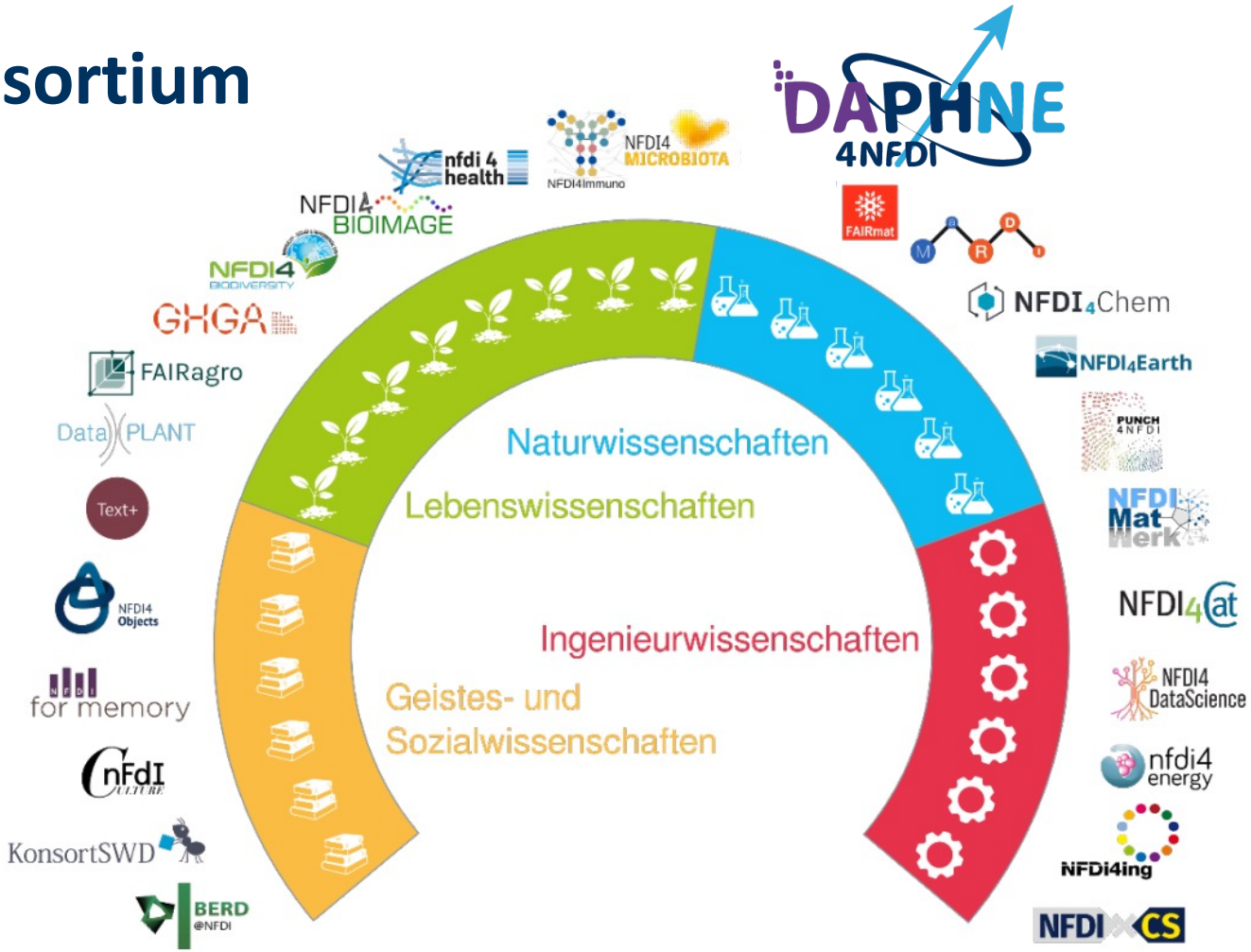
26+1 consortium

Nationale Forschungsdateninfrastruktur (NFDI) e.V.

Nationale Forschungsdaten Infrastruktur

NFDI Consortium

26+1 consortium

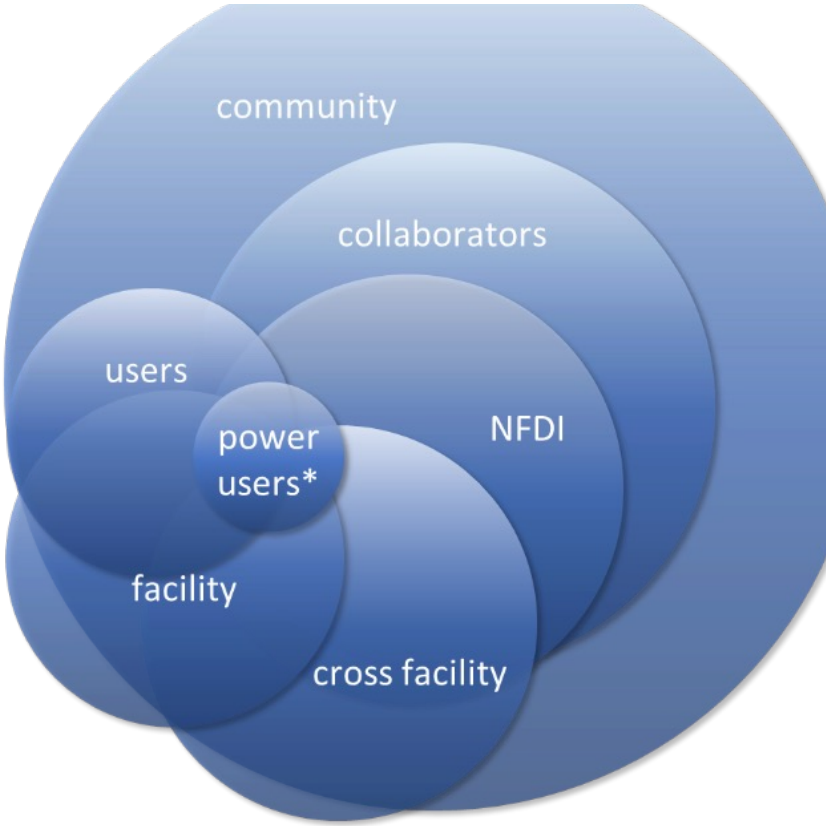


ErUM

Nationale Forschungsdaten Infrastruktur

Dissemination and outreach

The NFDI consortium as a role model and educator



Dissemination

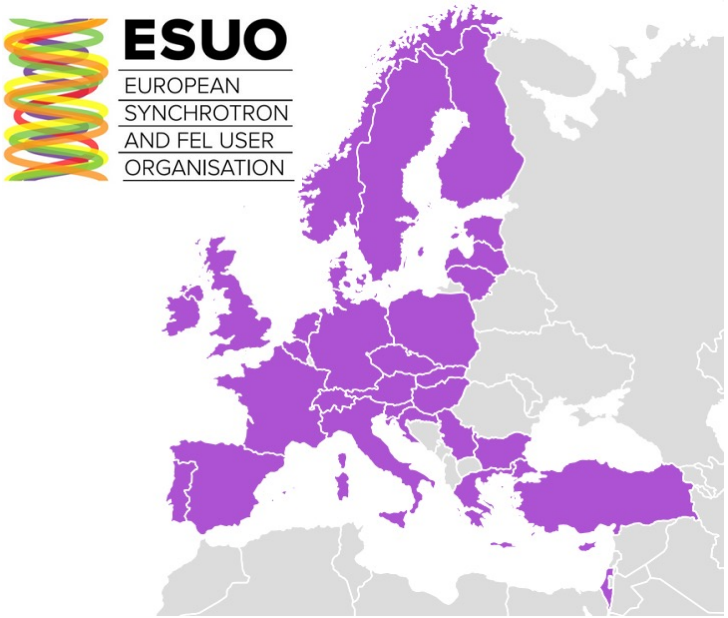
- Project results within and outside DAPHNE4NFDI
- FAIR data concepts
- Broad synchrotron and neutron techniques with data related tasks
- Interaction with facilities and beamlines

Outreach

- Communication platform in the DAPHNE4NDI community
- Via KFS and KFN, ENSA and ESUO, ErUM-data, NFDI e.V
- Conferences
- Homepage, communication to the public

Integration into international research community

- DAPHNE4NFDI worldwide network > 30.000 synchrotron and neutron users
- European user organizations
- European Open Science: PaNOSC and ExPaNDS, OSCARS



ESUO/ENSA/DAPHNE4NFDI meeting

ESUO and ENSA with DAPHNE4NFDI held a European Data policy at PSI November 13th – 15th 2023

- Photon and neutron community white paper on the importance of FAIR and Open data principles
- Describe how users and facilities can work to provide open science.

DAPHNE: Bridget Murphy Christian Gutt,
Astrid Schneidewind, Ullrich Pietsch,
ENSA: Henrik Ronnow, Lambert van Eijck, A.S,
ESUO: Cormac Mc Guinness, Bridget Murphy
ExPaNDS: Andrew Goetz

In discussion with ENSA and ESUO national delegates

DAPHNE@DESY

DESY is responsible for the consortium co-ordination

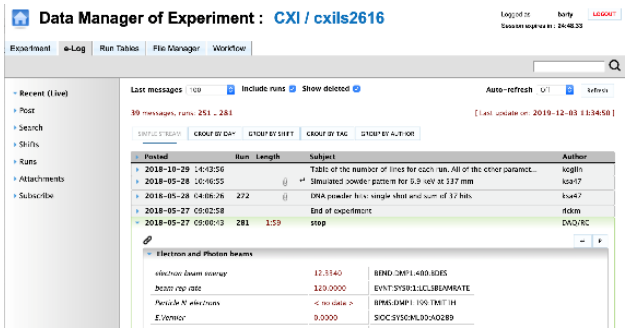
- SciCat for beamline data
  - P08 test case
  - FLASH test case
- PIDs for datasets (dataDOIs)
- Single sign on AAI
- Open data catalogues (public-data.desy.de)
- SciCat development
- Electronic log book for the facility users
- Sample PID infrastructure
- Software ecosystem – development, RSE and distribution (CrystFEL)
- Finances oversight

ExPaNDS ontologies

Document Control Information

Settings	Value
Document Identifier:	D3.2
Project Title:	ExPaNDS ontologies
Work Package:	WP3

https://zenodo.org/record/4806026#.YUh-kOdCRno

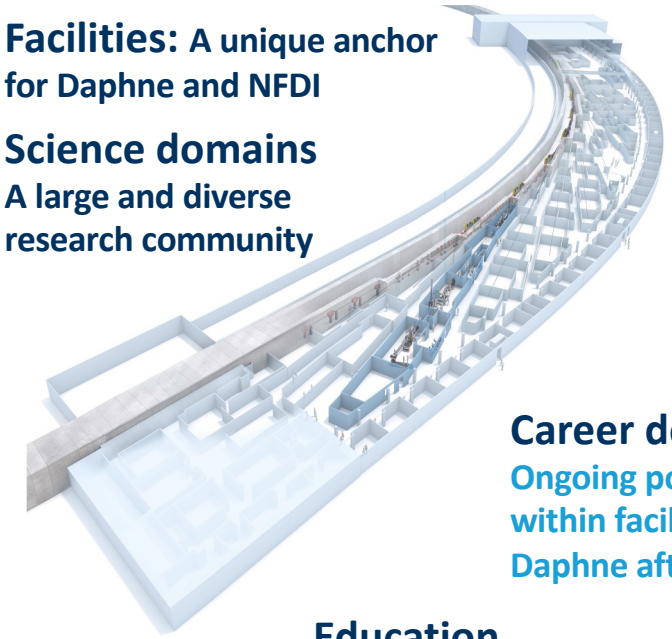


Both DAPHNE employees and in kind contributions are necessary to make DAPHNE work

Sustainability within NFDI beyond initial DFG funding

Facilities: A unique anchor for Daphne and NFDI

Science domains
A large and diverse research community



Integration
Daphne data infrastructure will be closely integrated with facilities, ensuring continued impact long after NFDI is over.

Continued operation of services
Operation of services and maintenance will be provided by the facilities after funding is over.

Career development
Ongoing positions, including long term, within facility computing groups sustain Daphne after initial funding.

Education
A culture of data curation and scientific software development will be embedded in University courses for students and future generations

Leading by example
Power user groups act as role models leading the community in entrenching best practices

© Andreas Heddergott

Thank you!