

Number of Users shared between facilities																	
	BER II	BESSY II	DESY	DLS	ELETTRA	ESRF	ILL	ISIS	LLB	SINQ	SLS	SOLEIL	FRM-II	ANKA	neutron	photon	all
BER II	2761	284	152	54	39	319	556	302	156	130	29	38	207	15	814	657	2761
BESSY II	284	7068	487	110	298	810	165	94	72	46	263	260	89	57	475	1601	7068
DESY	152	487	3563	88	121	735	194	91	55	44	155	130	103	43	402	1259	3563
DLS	54	110	88	3494	72	739	213	336	35	18	145	149	20	12	448	989	3494
ELETTRA	39	298	121	72	2731	455	85	43	23	4	66	316	9	20	162	906	2731
ESRF	319	810	735	739	455	10728	886	406	235	92	600	1069	144	80	1389	3463	10728
ILL	556	165	194	213	85	886	4338	741	343	229	69	176	349	10	1577	1280	4338
ISIS	302	94	91	336	43	406	741	2755	120	119	43	52	155	5	958	740	2755
LLB	156	72	55	35	23	235	343	120	1348	34	12	131	92	3	455	375	1348
SINQ	130	46	44	18	4	92	229	119	34	726	96	9	97	0	348	221	726
SLS	29	263	155	145	66	600	69	43	12	96	2424	182	18	18	177	974	2424
SOLEIL	38	260	130	149	316	1069	176	52	131	9	182	3656	14	26	309	1524	3656
FRM-II	207	89	103	20	9	144	349	155	92	97	18	14	1087	5	522	281	1087
ANKA	15	57	43	12	20	80	10	5	3	0	18	26	5	452	29	157	452

Source <http://www.pan-data.eu/CountingUsers>

Neutron Sources			
Facility	# Users	Users using at least one <u>other</u> Neutron Source	Users using at least one Photon Source
BER II	2761	814 (29.4%)	657 (23.7%)
ILL	4338	1577 (36.3%)	1280 (29.5%)
ISIS	2755	958 (34.7%)	740 (26.8%)
LLB	1348	455 (33.7%)	375 (27.8%)
SINQ	726	348 (47.9%)	221 (30.4%)
FRM-II	1087	522 (48.0%)	281 (25.8%)

Numbers of distinct neutron scattering users for the period Jan 2010 – Jun 2011 per participating facility

Total number of unique users: **35968** Photons: 28073 Neutrons: **10324**
 Using Neutrons **and** Photons: 2429 or 6.7% of all unique users
 Using more than one facility: 7757 or 21.6% of all users
 Using more than one Neutron source: 1983 or **19.2%** of all neutron users

Photon and Neutron Data Infrastructure

History

Who are we ?

The problems

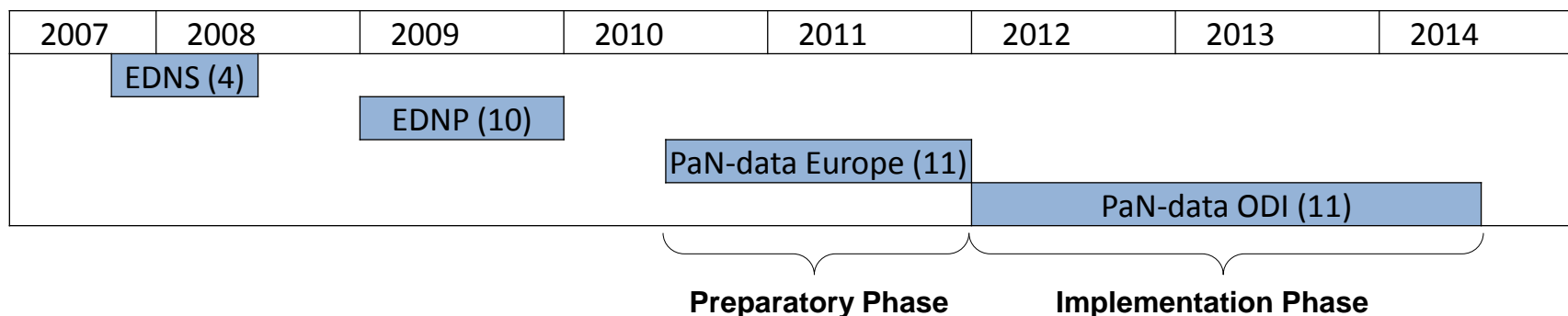
Aims

What has been achieved so far ?



Established 2007 with 4 facilities

Expanded since to 11 facilities



Who are we ?



Joining soon : MAX IV ? JCNS ?

Participants : IT persons and scientists (too few ?)

The problems

1. scientific data is often considered private property
2. open access to scientific data is almost impossible
3. scientific data is not managed professionally
4. barriers to interact with data must be lowered
5. scientists use several facilities for their research
6. the data deluge makes everything worse (mostly X-Ray)

- Data is not on-line
- Data is poorly or not described
- No search tools
- No way to identify/cite them
- Authentication/authorisation for scientists is difficult
- Open access is not (yet) well accepted
- Institutions lack infrastructure



ESFRI Position Paper on Digital Repositories:

“Research Infrastructures should guarantee that raw research data are made available through portals and databases.”

06/09/2007 – e-IRG ESFRI



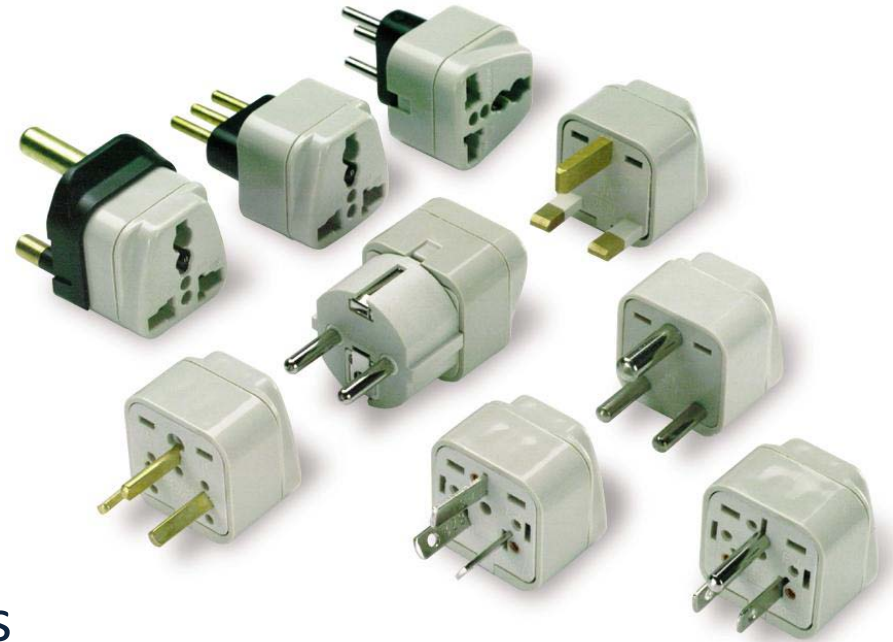
Data's shameful neglect

“Research cannot flourish if data are not preserved and made accessible. All concerned must act accordingly”

Nature **461**, 145 (10 September 2009) | doi:10.1038/461145a

Barriers to interact with data must be lowered Scientists use several facilities for their research

- Common data file format
- Federated data bases
- Search tools, visualization tools
- EU wide authentication and authorization of scientists
- ...



- Harmonize authentication and authorization
- Standardize data formats and annotation of data
- Allow transparent and secure remote access to data
- Establish sustainable and compatible distributed data catalogues (cross search engine)
- Allow long term preservation of data
- Provide compatible data analysis software
- Promote data policies in laboratories
- The Infra should cover the whole continuum (from proposal to publication)

What has been achieved so far where do we stand

Data policy

Authentication

Metadata catalogue

Software catalogue

Data format

Persistent Identifiers for data sets

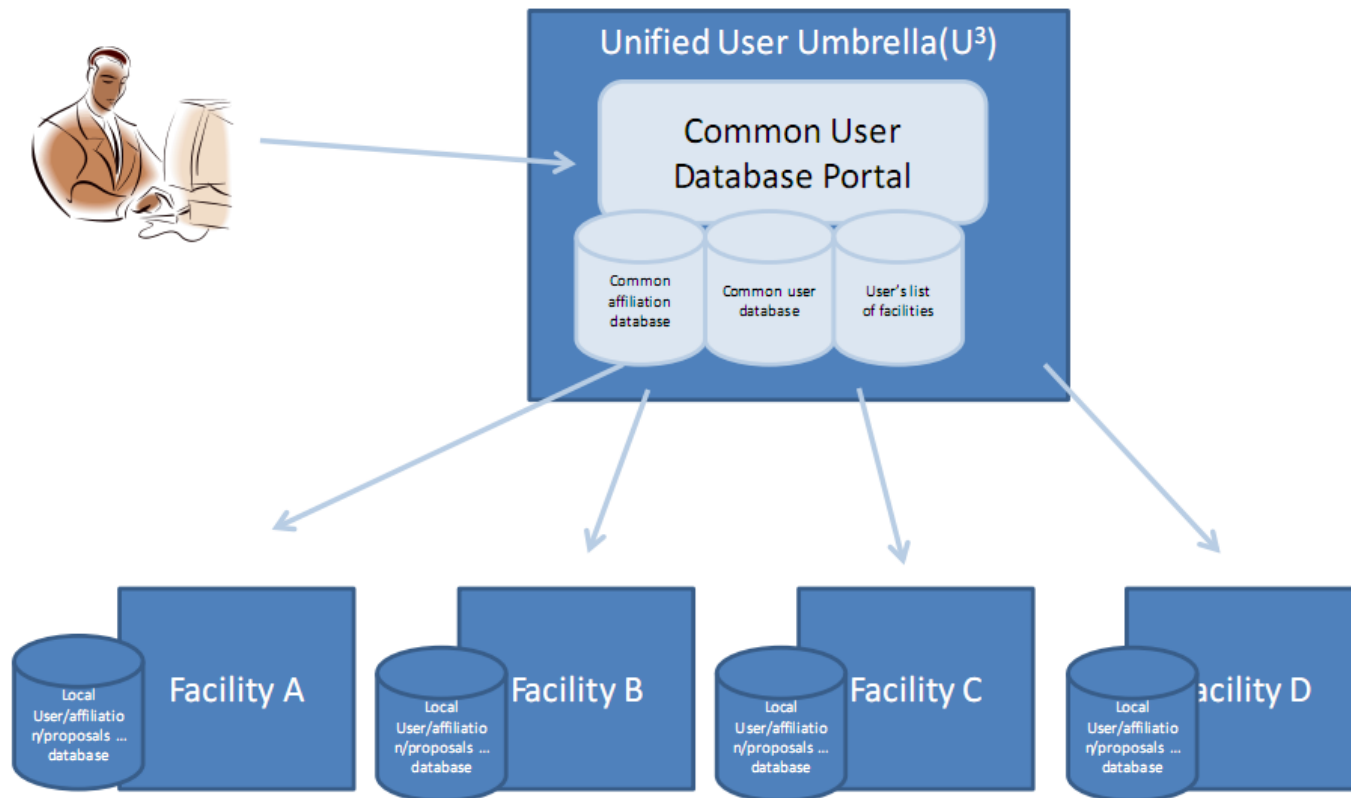
- We need a clear statement on :
 - Who is the owner
 - If we make them publicly available, how do we protect the Experimental team.
- What do we propose :
 1. The facility shall act as a **custodian** for the data.
 2. **All raw data will be curated** in a well-defined format with a unique ID.
 3. **Metadata** is captured automatically and resides either within the raw data files, and/or in an associated on-line catalogue.
 4. **Access to raw data** and the associated metadata obtained from an experiment **is restricted to the experimental team for a maximum period of 3 years**. Thereafter, it will become publicly accessible.
 5. The embargo period can be extended on requests.
 6. Analysis of openly accessible data must acknowledge the source of the data and cite its unique identifier and any publication linked to the same raw data



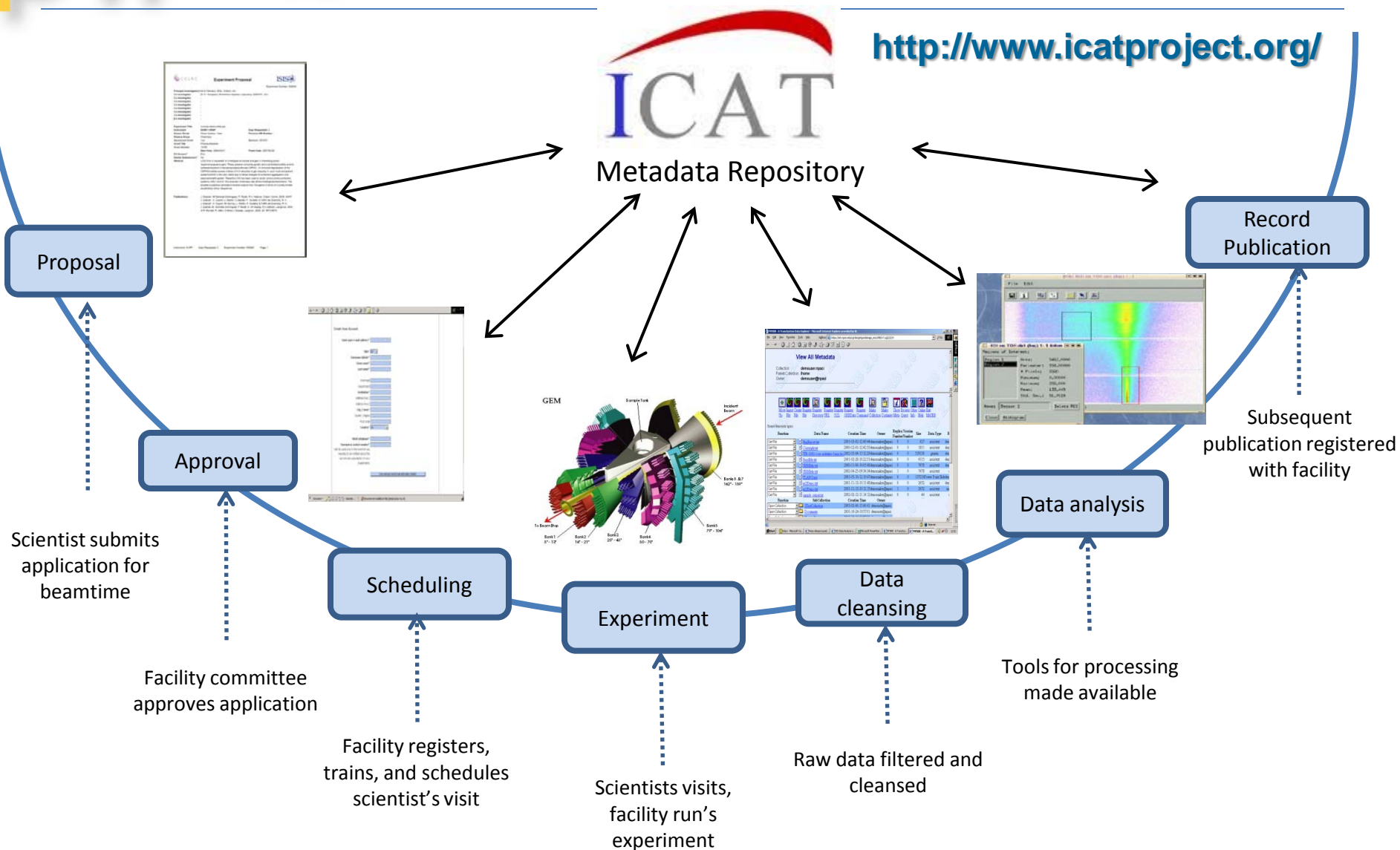
<http://www.pan-data.eu/imagesGHD/0/08/PaN-data-D2-1.pdf>

- A user should have a single id through the different facilities.
- Federation of Authentication. A user once authenticated by Facility A is automatically identified by Facility B
- Decision on Authorization should remain local to the facility.
- Simple to use : avoid WAYF with long list of IdP, delays for getting certificate,
- Should remain persistent when user change its affiliation

The user credentials and attributes are managed centrally.



A prototype implementation (based on Shibboleth) is operational



- Technically
 - Database schema
 - A set of WebServices i.e You can decide for your own user interface
- Functionalities
 - Allows the aggregation of
 - Experimental parameters
 - Information coming from the proposal system
 - references to datafiles
 - DOIs
 - Implement the security model driven by the data policy
 - Enable cross searching between ICAT instances.

Implemented at ISIS, SNS, Diamond and ILL

Many software are developed by individuals.

That's great !

We just want to help to raise up their usability.

- Create / Manage a catalogue of available software (mainly analysis)
- Forster good practice in software development (licences, code versioning, tests, info about support, documentation, ...)
- Through Web Services could be integrate in facility portals or ICAT instances

PaNsoft Catalogue - Mozilla Firefox

File Edit View History Bookmarks Title Tools Help

Back Forward Tile Tabs pandata.ill.fr/app_dev.php/ Reload Stop Google Home grabMyBooks

Les plus visités Débuter avec Firefox À la une English-French Diction... Thinkmap Visual Thes... Import to Mendeley Cambridge SIGAL OCS Inventory eGroupWare [Calendar...

PaNsoft Help & Support Software Catalogue Institutes Search software... Login Register

Photon and Neutron Software Catalogue

PaNsoft is a database of software used mainly for data analysis of neutron and photon experiments. PaNsoft is one element of a larger project, [PaNdata](#), which aims to provide a complete, shared data infrastructure for neutron and photon laboratories.

This database can be freely consulted. It gives an overview of software available for neutron and photon experiments and their use with respect to instruments at experimental facilities.

By [registering](#) and [logging-in](#) new software can be entered and it will appear in the database after moderation. Similarly, feedback can be given on the software presented herein and more generally via the forum hosted here.

[Browse software](#)

Software: [Recent Software](#) [Popular Software](#)

Recent Software

NAMD
NAMD is a parallel molecular dynamics code designed for high-performance simulation of large biomolecular systems.

PORE3D
large software suite for filtering, segmentation and quantitative analysis of 3D images (CT, MRI, CLSM, ...).

Ominc
Logiciel d'acquisition des spectres infrarouge. Ce logiciel permet aussi un de créer des cartographies

IMOD
IMOD is a set of image processing, modeling and display programs used for tomographic reconstruction and for 3D reconstruction of EM serial sections and optical sections.

2.1.0-DEV PHP 5.3.5-1ubuntu7.2 [xdebug] [accel] app|dev|debug|4e95977de05fe HomeController::indexAction|pandata_index|200 111 ms 5376 KB anon. 4 en-US zotero

PaNdata Europe - Search - Mozilla Firefox

File Edit View History Bookmarks Title Tools Help

PaNdata Europe - Search

Back Forward Tile Tabs

pandata.ill.fr/app_dev.php/software/search/result/6de49fd742e3de0cf14a9fd29621f88

Reload Stop Google Home grabMyBooks

Les plus visités Débuter avec Firefox À la une English-French Diction... Thinkmap Visual Thes... Import to Mendeley Cambridge SIGAL OCS Inventory eGroupWare (Calendar...

PaNsoft Help & Support Software Catalogue Institutes Search software... Login Register

Search Results

Filter Results

Institutes

- ALBA Synchrotron Light Facility
- DESY (Deutsches Elektronen-Synchrotron)
- Diamond Light Source
- ELETTRA Synchrotron Light Facility
- European Synchrotron Radiation Facility**
- Helmholtz-Zentrum Berlin
- Institut Laue-Langevin**
- ISIS neutron source
- Laboratoire Léon Brillouin
- Paul Scherrer Institut
- Soleil Synchrotron

Categories

- Absorption
- Diffraction
- Imaging
- Instrument Simulation
- Large Scale Structures
- Muons
- Protein Crystallography
- Sample Simulations
- Spectroscopy

Data Formats

- ASCII
- FullProf
- HDF5
- IDL
- ILL TAS Data Format
- Instrument Raw Data
- NetCDF
- NetX

Your search for returned 11 results. Use the options on the left to filter the results.

- HILBERT++**
XMCD, and different spectroscopies
- LAMP (runtime)**
Large Array Manipulation Program. General Data Inspection, Reduction and Analysis software framework.
- MagDraw**
Magnetic Editor Cell Constructor
- Materials Studio**
Materials Studio is a software suite for atomistic simulations including DFT, force field methods and others.
- NAMD**
NAMD is a parallel molecular dynamics code designed for high-performance simulation of large biomolecular systems.
- nMoldyn**
nMOLDYN is an interactive analysis program for Molecular Dynamics simulations. It is especially designed for the computation and decomposition of neutron scattering spectra, but also computes other quantities.
- Phonon**
PHONON is a code to calculate lattice dynamics, mainly using input force constants from ab initio (DFT) codes like VASP. There are many kinds of output including spectroscopic scattering functions.
- RECUIT**
inverse problem of multilayers and design of supermirrors
- SHADOW**
X-ray tracing

en-US zotero

Currently only a prototype.

1st public release
Spring 2012

- Reduce need for local expertise
- Reduce number of conversion utilities
- Reduce redundant software development
- Increase metadata description – a prerequisite for data archiving/curation
- Increase cooperation in software development
- Increase functionality of generic software
- Create a community standard
- Create critical mass to influence detector manufactures
- Ease cross facility experiments
- ...

The PaN-data consortium has decided to adopt **HDF5/Nexus**



Slow progress in the deployment

DOIs (currently mainly used for journals) for the identification of the raw data sets.

- DOI = 10.5291/ILL-DATA.6-05-579
- <http://doi.ill.fr/10.5291/ILL-DATA.6-05-579>
- Creator = Investigators/PI
- Title = Proposal Title
- Publisher = Institut Max Von Laue – Paul Langevin
- Date = 3rd December 2010
- Data Format = NXS

Deployment phase at ISIS and ILL

Other projects with related activities :

[DATACITE](#)

[CRISP](#)

[ORCID](#)

Questions ?