



PaN-data ODI

WP2 2nd annual report

| | |
|------------------------------------|-----------------------------------|
| Grant Agreement Number | RI-283556 |
| Project Title | PaN-data Open Data Infrastructure |
| Lead Beneficiary | STFC |
| Deliverable Dissemination Level | Public |
| Deliverable Nature | Report |
| Contractual Delivery Date | none |
| Actual Delivery Date | 12 Nov 2013 |

The PaN-data ODI project is partly funded by the European Commission under the 7th Framework Programme, Information Society Technologies, Research Infrastructures.

Table of Contents

| | |
|--|----------|
| Report on Dissemination and Engagement Activities | 3 |
| 1.1. Generic activities..... | 3 |
| 1.2. Activities related to global data infrastructures..... | 3 |
| 1.3. Dissemination and Engagement with respect to the Umbrella AAI | 4 |
| 1.3.1. FIM4R workshops | 5 |
| 1.3.2. Terena AAA paper..... | 5 |
| 1.3.3. Further activities | 6 |
| 1.4. Dissemination and Engagement with respect to NeXus and Standards..... | 6 |
| 1.5. Contribution to and (co-)organization of community events | 7 |
| 1.5.1. Joint PNI-HDRI and PaNdata ODI workshop..... | 8 |
| 1.5.2. NOBUGS 2012..... | 9 |
| 1.5.3. Joint Photon Science and European XFEL User meeting | 9 |
| 1.5.4. Co-operations with industry | 9 |
| 1.5.5. Open Technical Forum | 10 |
| 1.6. Engagement with the general public | 10 |
| 1.7. Workshops and Events organized by PaNdata partners..... | 12 |

Report on Dissemination and Engagement Activities

1.1. Generic activities

PaNdata ODI reports, deliverables and documents including all presentations, poster and other materials are freely available from the PaNdata web-site:

- All PaNdata ODI deliverables: <http://pan-data.eu/ODIWP>
- All PaNdata ODI presentations: <http://pan-data.eu/presentations>
- All other documents: <http://pan-data.eu/Documents>

We of course aim to update all information in a timely manner. Occasional announcement on twitter and news entries on the PaNdata web-site serve as a basic tool to disseminate on-going activities.

1.2. Activities related to global data infrastructures

One of the crucial topics for any data infrastructure is the persistent identification (PID) of digital objects and the citability of the corresponding PIDs. Some PaNdata partners are actively working on this topic in co-operation or at least based on knowledge exchange with projects and initiatives like EUDAT, OpenAire¹ or DataCite². For example, ILL has been building up an infrastructure to assign DOIs to datasets and register them through DataCite. Likewise DESY has established contracts with DataCite and is using the DOI handle system to register instruments and resources; registration of actual scientific data will follow in due time.

The newly formed Research Data Alliance (RDA)³ also aims to establish a number of working groups dealing with PIDs and related meta-data. RDA is explicitly not a platform to promote projects, but rather a bottom-up community or experts approach to derive, implement and establish standards or recommendations in form of RFCs. As such RDA provides the prime platform to exchange knowledge and ideas with individuals involved in related projects and service providers, in particularly EUDAT, iCORDI or ORCID to name a few. One major goal will hence to take up the RDA recommendations and align the PaNdata ODI implementations accordingly (were feasible). To facilitate and guarantee the information flow from RDA to PaNdata ODI and vice versa, PaNdata partners are participating in the management board (J.Bicarregui, STFC) and three of the RDA working groups, in particular on PID Information Types and Type registries. In addition, PaNdata has been participating and contributing to all (open) RDA and EUDAT events so far, and provided substantial feedback to EUDAT in interviews as well as to the e-IRG blueprint.

A very recent initiative aims to establish a Photon and Neutron Science Interest Group (PaNSIG) with RDA. The corresponding case statement will be submitted in November 2013 and a PaNSIG working space has already been arranged with RDA. Very supporting feedback has been received from PaNdata partners and colleagues at Australian and American Neutron and Photon sources. We

¹ <http://www.openaire.eu/>

² <http://datacite.org/>

³ <http://rd-alliance.org/>

aim to co-locate the final PaNdata workshop with the upcoming RDA plenary meeting in Dublin end of March 2014 and have started discussions with RDA members recently.

1.3. Dissemination and Engagement with respect to the Umbrella AAI

There are a number of PaNdata ODI goals, which require a high level of harmonization or standardization with projects or communities outside the PaNdata consortium. The core elements of the PaNdata Open Data Infrastructure are hence intensively discussed with and disseminated to the communities and projects. One of these core elements is the pan-European identity system.

PaNdata ODI aims to establish a common, unique identity for all its users. This identity system, the so called Umbrella, is intended to provide additional services to both users and facilities, like a harmonized proposal submission procedure, educational elements or a common facility database. The facility database is a major undertaking, since it requires establishing a common, multi-lingual schema and merging of literally ten thousand of entries. However, once established maintenance of the database will become much easier and more efficient since the effort is divided through all partners.

Umbrella has been thoroughly tested not only by facility staff but also by a significant number of users providing valuable feedback to the developers. To promote implementation and deployment two teams have been organized, one concentrating more on management issue, the other on in-depth technical aspects. The teams hold regular telephone conferences and meetings. Since the topic is of high interest not only for the RIs organized in PaNdata, a number of additional RIs and projects are participating in the effort, such as EMBL⁴ as the Biostruct-X⁵ project leader, members of the NMI3 and CALIPSO projects, GSI/FAIR⁶ and the European XFEL⁷.

Umbrella has been presented at several meetings and conferences, in particular Biostruct-X and Calipso⁸ meetings, the IUCR⁹, the series of workshops on Federated Identity Management (FIM) for Research Collaborations¹⁰ initiated by EiroForum¹¹ and more general events of EUDAT¹² or e-IRG¹³. There has been a meeting between WP3 (Umbrella) and ESUO, the synchrotron user organization. This has led to an explicit recommendation by ESUO of Umbrella as the official identification tool at the European synchrotron facilities.

The main goal of the PaNdata ODI project is to establish a common, federated, open data infrastructure. This involves a number of issues ranging from federated identity management to meta-data standards or persistent identifier for data, instruments and related publications. Quite a number of projects and initiatives are working in closely related topics. Seeking for synergies, establishing co-operations and exchanging knowledge is therefore particularly important to arrive at solutions facilitating integration into a global data infrastructure.

⁴ European Molecular Biology Laboratory: <http://www.embl-hamburg.de>

⁵ <http://www.biostruct-x.eu/>

⁶ Facility for Antiproton and Ion Research: <http://www.fair-center.de/>

⁷ European X-Ray Free Electron Laser: <http://www.xfel.eu>

⁸ Calipso: Coordinated Access to Lightsources to Promote Standards and Optimization

⁹ International Union of Crystallography: <http://www.iucr.org/>

¹⁰ <https://cdsweb.cern.ch/record/1442597>

¹¹ <http://www.eiroforum.org/>

¹² <http://www.eudat.eu/>

¹³ e-Infrastructure Reflection Group: <http://www.e-irg.eu/>

Projects like CALIPSO or Biostruct-X are aim for example to ease the access to the facilities through transnational access support or harmonized proposal submission systems for protein crystallography (PX) beamlines. Implementation of Umbrella at the synchrotron facilities is now an explicit part of the CALIPSO project. Both projects consequently need to tackle identity management issues, and the Umbrella also intends to incorporate certain parts of the proposal submission, but not restricted to PX instrument. To arrive at common solution satisfying the needs of the user communities, intense consultations between these projects and PaNdata ODI have been established and Biostruct-X is meanwhile participating in the Umbrella management team to exploit synergies in this area of common interest.

One platform to discuss directions in identity management and the implementation of a common system are the series of so called harmonization meetings. Four of these meetings^{14 15 16} have been organized by PaNData ODI partners, the latest jointly between the European XFEL and PSI. The meeting took place in Hamburg at the Eur.XFEL site and was visited by 35 participants from various RIs¹⁷. At the meeting important security and organizational issues have been intensely discussed. As a result the participants agreed on major standardization issues and an essentially final draft of an Memorandum of Understanding aims to provide a solid basis for a sustainable operation of the IdM infrastructure.

1.3.1. FIM4R workshops

Federated identity management is also a crucial topic heavily discussed in the FIM4R workshops. Within the last two years there have been series of these workshops, a number of them have been organized by PaNdata partners, the latest at PSI in March 2013¹⁸. The most recent workshop was held in Helsinki in conjunction with Terena and VAMP2013 conferences¹⁹. An important outcome of these activities is a paper ‘Federated Identity Management for Research Collaborations’²⁰.

The FIM4R activities also promoted the very recent collaboration between PaNdata ODI and the GN3+ project to construct a bridge between UmbrellaID and the edugain federation²¹, which would also serve as a case studio relevant for a significant number of projects dealing with the implementation or operation of AA-Infrasctructures.

1.3.2. Terena AAA paper

An interesting study has been performed by Terena “Advancing Technologies and Federating Communities²²” with significant contributions from PaNdata, such as a use case “Accessing Experiments and Data” from the Photon and Neutron Community as well as The Umbrella Project as an example of a AAI.

¹⁴ <http://indico.psi.ch/conferenceDisplay.py?confId=1039>

¹⁵ <https://indico.desy.de/conferenceDisplay.py?confId=5061>

¹⁶ <http://indico.psi.ch/conferenceDisplay.py?confId=1752>

¹⁷ <http://indico.psi.ch/conferenceDisplay.py?confId=2159>

¹⁸ <http://indico.psi.ch/conferenceDisplay.py?confId=2230>

¹⁹ <https://refeds.org/meetings/oct13/index.html>

²⁰ <https://cdsweb.cern.ch/record/1442597>

²¹ <http://pan-data.eu/node/72>

²² <https://confluence.terena.org/download/attachments/30474266/2012-AAA-Study-report-final.pdf?version=1&modificationDate=1355503760046>

1.3.3. Further activities

Beyond these activities FIM has been intensely discussed with projects like Eur.XFEL, CRISP²³ and PNI-HDRI²⁴ or LSDMA²⁵ and the Umbrella developments are regularly tested and investigated by users from the scientific communities. For example, dcache.org²⁶ is developing a cloud-storage solution in the context of the LSDMA project. The dcache-cloud permits a simple yet powerful upload and sharing of data for example via webdav, a http-based protocol available for virtually any operating system including mobile devices. Current developments aim to support various authentication mechanisms including the Umbrella system.

Project moonshot²⁷ is developing tools which permit integration of the Umbrella in non-web based analysis workflows and could provide access to compute resources through the Umbrella credentials. PaNdata ODI is currently investigating the set of tools with Project Moonshot.

Identity management and AAI is also an important topic of the EUDAT project.²⁸ PaNdata has provided feedback on the requirements of the Photon and Neutron user communities to EUDAT through participation in interviews, workshops and conferences. Likewise, PaNdata has submitted very detailed comments on the e-IRG blueprint²⁹ on data infrastructures to e-IRG reflecting the particular view of our user communities and in the context of the EIROForum-position paper on e-infrastructure requirements.

1.4. Dissemination and Engagement with respect to NeXus and Standards

The standardization efforts in PaNdata focus mainly on the data formats, metadata schemes and defined vocabularies. HDF5 has been proposed by the EC to serve as the ISO standard for all binary data. PaNdata has adopted this proposal and selected NeXus as their community wide standard. NeXus is fully HDF5 compliant, but with a standardized meta-data scheme and controlled vocabularies. So NeXus is not really a new data format, but more of a convention how to organize data and meta-data in a HDF5 container, which greatly facilitates exchange and concurrent use of data from different scientific fields or experiments. The NeXus structure enables the automatic processing of the meta-data and ingestion of the data into data catalogues. For example, a NeXus ingestor is currently being developed for ICAT³⁰ and the ICAT scheme, which is fully compliant with the Dublin Core standard³¹, and will be further adopted to support community requirements and the NeXus data model.

Standardization requires naturally the involvement of many disjoint communities, like developers, vendors, users and IT-expert. PaNdata partners aim to continuously contribute and drive the process. One core platform to engage with the NeXus community is the Nexus International Advisory Committee (NIAC)³². Some PaNdata partners have been participating in the NIAC almost since its existence. The PaNdata ODI efforts have led to a stronger and broader representation in the NIAC.

²³ Cluster of Research Infrastructures for Synergies in Physics: <http://www.crisp-fp7.eu>

²⁴ Photon and Neutron Infrastructures - High Data Rate Initiative: <http://www.pni-hdri.de>

²⁵ Large Scale Datamanagement and Analysis: <http://www.helmholtz-lsdma.de>

²⁶ <http://dcache.org>

²⁷ <http://www.project-moonshot.org/>

²⁸ <http://www.eudat.eu/>

²⁹ <http://www.e-irg.eu/publications/blue-papers.html>

³⁰ <http://www.icatproject.org/>

³¹ <http://dublincore.org/>

³² <http://wiki.nexusformat.org/NIAC>

At the latest NIAC meeting, H. Bernstein representing imgCIF has joined the NIAC. CIF is another well-established standard widely used in the field of protein crystallography. To achieve interoperability between NeXus and CIF is hence an important issue tackled by the NIAC and the IUCr, and is well progressing³³. More information can also be found on the IUCr forum on NeXus HDF5 CIF convergence³⁴.

At the ECM28³⁵ the issue has been further developed. The presentation of Herbert Bernstein on *Managing crystallographic data in facilities using integrated CIF, HDF5 and NeXus*³⁶ nicely summarizes the efforts for interoperability between well-established standards, stressing the importance to combine the best of CBF, NeXus and HDF5 for BigData management. A way towards concordance³⁷ has been outlined and intensively discussed between CIF and NeXus members at the CIF/NeXus/HDF5 meeting preceding the CIF workshop³⁸.

The aim has been expressed to achieve a full interoperability between CIF and NeXus, such that NeXus becomes a IUCr recognized standard data format, with a standardized, IUCr validated NeXus application definition. It's certainly a major step towards a widely adopted meta-data standard for a large scientific community.

High Speed data recording is becoming more and more important. Free Electron Laser facilities aiming to resolve processes at a femto-second scale strongly depend on the ability to record images at a MHz rate, but also synchrotrons, FAIR and even some Neutron facilities like ESS face the same problem with emerging new detector generations. To cope with these challenges PaNdata is co-operating with developers and vendors. There is for example an intense co-operation with Dectris³⁹, one of the leading detector companies, the NIAC and HDF5.org to enable detectors writing NeXus/HDF5 natively at the speed required. A workshop has been organized by PSI to support this process and DESY has developed in co-operation with the PNI-HDRI project an implementation of the NeXus API capable to deal with such data rates. Work on compression algorithms and methods to augment HDF5 with pluggable compression and image filtering modules has meanwhile successfully been concluded.

The NeXus developments have also been presented at the NIAC and NeXus code camp and the NOBUGS conference in September 2012.

1.5. Contribution to and (co-)organization of community events

A number of events have been organized by PaNdata partners which we considered particularly suited to promote PaNdata ODI. Each of these events has a major impact in user communities as well as Photon and Neutron RIs.

³³ http://wiki.nexusformat.org/NIAC2012#Meeting_Minutes

³⁴ <http://forums.iucr.org/viewforum.php?f=31>

³⁵ <http://ecm28.ecanews.org/>

³⁶ http://www.iucr.org/__data/assets/pdf_file/0007/80269/HJB_Managing_Crystallographic_Data.pdf

³⁷ http://www.iucr.org/__data/assets/pdf_file/0016/80152/Concordance_Summary_8Aug13.pdf

³⁸ <http://www.iucr.org/resources/cif/comcifs/workshop-2013>

³⁹ <https://www.dectris.com/>

1.5.1. Joint PNI-HDRI and PaNdata ODI workshop

PaNdata ODI and PNI-HDRI have jointly organized two workshops on more technical aspects. The first Joint PNI-HDRI and PaNdata workshop took place in Hamburg beginning of 2012⁴⁰ and was rather successful in knowledge exchange and creation of collaborations on specific aspects of application development.

The second workshop⁴¹ was held again at DESY in Hamburg in March 2013 and focusing more on (parallel) Computing in Photon and Neutron Science Applications emphasizing topics related to work packages 5 (Virtual Labs) and 7 (Scalability).

Slides and materials for both workshops are available from the PaNdata web.

SRI 2012 and ECM 27

In July 2012 the 11th International Conference on Synchrotron Radiation Instrumentation⁴² took place in Lyon, by far the largest event of this kind, organized by PaNata partners ESRF and Soleil with a number of satellite workshops at PSI, DESY or the European XFEL. One particularly noteworthy activity at the conference was the information gathering on data management.

Rudolf Dimper (ESRF) and Philippe Martinez (Synchrotron SOLEIL) organized a round table on large data volume management⁴³. The outcome of the round table discussion was presented⁴⁴ at the Workshop on Data Diffraction Deposition⁴⁵ (DDD) at ECM27⁴⁶, the European Crystallography Meeting in Bergen/Norway as a supplement to the talk by Heinz Weyer presenting PaNdata and related projects⁴⁷. Erica Yang and Brian Matthews were completing the PaNdata ODI presentations with a talk on *Linking raw experimental data with scientific workflow and software repository*⁴⁸. The presentations revived the discussion on data deposition and open access to scientific data in the IUCr, and the PaNdata ODI policy framework and the implementation with ICAT at ISIS was featured in the workshop report as “*An exemplar of good practice demonstrating access to raw data is at the ISIS UK neutron source*”⁴⁹. This workshop offered an ideal occasion to exchange ideas and requirements with between facility staff and users at the large facilities represented by PaNdata.

The workshop report lists some actions and recommendations. It suggests in particular “*to encourage and recommend to the IUCr Executive Committee that authors should provide a permanent and prominent link from an article to the raw data sets underpinning a journal publication with a view to making this a formal requirement on authors at such time as the community has adopted raw data deposition as a routine procedure*” and emphasized the “*urgent need to be clear about the metadata required for the various IUCr Commissions and their experimental raw data*”.

This is an important step towards an open data infrastructure for one of our particularly prominent user communities. The deposition of scientific data, linked to persistent identifier cited in a publication, standardization of meta-data with the idea to make data re-usable and accessible is exactly the type of infrastructure PaNdata ODI aims to provide.

⁴⁰ <https://indico.desy.de/conferenceDisplay.py?confId=5517>

⁴¹ <https://indico.desy.de/conferenceDisplay.py?confId=7333>

⁴² SRI2012: <http://www.sri2012.org>

⁴³ Programme: <http://www.lepublicsystemepco.com/files/modules/freezones/ProgrammeSRI2012-Web2.pdf>

⁴⁴ http://www.iucr.org/_data/assets/pdf_file/0008/69479/Data-Mgt-RT.pdf

⁴⁵ <http://www.iucr.org/resources/data/dddwg/bergen-workshop>

⁴⁶ <http://ecm27.ecanews.org/>

⁴⁷ <http://pan-data.eu/sites/pan-data.eu/files/03-BergenECM272012fb-Weyer.pdf>

⁴⁸ <http://pan-data.eu/sites/pan-data.eu/files/04-Linking-Raw-Data-with-Scientific-Workflow-PanData-ODI-Early-Experience.pdf>

⁴⁹ <http://forums.iucr.org/viewtopic.php?f=21&t=102>

1.5.2. NOBUGS 2012

Another particularly important event was NOBUGS 2012⁵⁰ which was organized by STFC and DLS and took place at RAL Sept. 2012. NOBUGS is the only event for photon science applications where developers and users from various scientific areas come together. It was hence an obvious choice to attach a significant part of the PaNdata ODI activities to the NOBUGS conference, there is hardly a better opportunity to disseminate and exchange knowledge with developers and user communities.

The NOBUGS conference was hence accompanied by a number of PaNdata satellites, like Umbrella and ICAT workshops as well as a NeXus Code camp and NIAC meeting. PaNdata work was presented in several of the NOBUGS talk like the data management reports from DESY⁵¹ and ALBA⁵², a status report for NeXus⁵³, the data catalogue requirements analysis⁵⁴ and the practice of data citation at ISIS⁵⁵. Several of the talks were again focusing on data analysis frameworks and workflows further emphasizing the need for an integrated data analysis framework. More information about the satellite workshop can be found on the PaNdata web and the NOBUS2012 site.

1.5.3. Joint Photon Science and European XFEL User meeting

In January the joint Photon Science and European XFEL User meeting in Hamburg was organized by the European XFEL and PaNdata partner DESY. With more than 800 registered participants it is one of the largest user community events. PaNdata ODI presented a poster (jointly with LSDMA) and pursued a survey to investigate users' needs, requirements and awareness of data management issues. We in particular approached user groups asking for the availability of open access data. So far the feedback was surprisingly positive, and just hindered by the problem of users to actually locate experimental data, which nicely illustrated the need for a sustainable data infrastructure. These efforts will be continued and summarized at a later stage.

On the occasion of the users meeting we gave some interviews to a master student in science sociology, who intends to use PaNdata ODI as a case study to illustrate the pathways for creation of virtual environments.

1.5.4. Co-operations with industry

As mentioned above we are seeking co-operations or exchange of technical knowledge with industrial partners. The focus is here the promotion of the standard data format and the improvement of system for high-speed data collection and recording.

The co-operation with a number of PaNdata partners, in particular DESY, STFC and PSI, and Dectris and the HDFgroup has already led to significant results beneficial not only to this particular community but for all scientific communities making use of HDF5.

Dectris has enabled their next-generation detectors to natively write NeXus/HDF5. To achieve the desired performance/bandwidth some adjustment to NeXus/HDF5 were required. The adjustments of the NeXus header has been discussed and supported by the NIAC. This has led to the development of external filters/plugins in HDF5, which allow performing parallelized data manipulations in a

⁵⁰ <http://nobugs2012.org/>

⁵¹ T Kracht, T Nunez, A Rothkirch, E Wintersberger: [Experiment Control and Data Acquisition at PETRA III, DESY](#)

⁵² C Pascual-Izarra et al: [Data management for the Beamlines at Alba](#)

⁵³ M Koennecke et al: [The State of NeXus](#)

⁵⁴ M Prica: [Requirements for data catalogues within facilities](#)

⁵⁵ MD Wilson, BM Matthews, S Nagella, AJ Wilson: [Using DataCite DOIs for ISIS Neutron Source Data](#)

transparent manner without affecting the native HDF5 API. The development has been funded by PaNdata partners and DECTRIS and is freely available in the main source tree of HDF5. See also the DECTRIS “success story”⁵⁶.

Another topic are low level co-operations with some storage system providers. DESY for example is testing new platforms from Netapp and IBM. The API, in particular NFS 4.1 are thoroughly tested against the current needs of the Photon and Neutron facilities, while keeping an idea on future requirements originating for example from Eur.XFEL, ESS⁵⁷ or SKA^{58 59}. Similar activities are pursued at other facilities also testing products like dCache⁶⁰, pvfs⁶¹, Lustre⁶², fhgfs⁶³ or hadoop⁶⁴. Although these activities are mostly bilateral, experiences and knowledge are exchanged between all partners and the vendors hopefully accelerating the process to arrive at stable high performance systems tailored for the needs of current and future data infrastructures.

1.5.5. Open Technical Forum

Colleagues in the US have recently initiated an Open Technical Forum (OTF) to provide a platform for technical presentations and discussions on Neutron and Photon Science related issues⁶⁵. The Open Technical Forum is scheduled so that participation becomes possible for scientists located in Europe as well as in the US. The Open Technical Forum is to our knowledge the first and most desirable attempt for a cross-continental forum. PaNdata ODI has been trying to support and advertise the Open Technical Forum as well as possible and contributed a number of talks. Talks by PaNdata members are available on the PaNdata web; where available talks and announcement can also be found on an Indico conferencing system⁶⁶. We hope that the RDA Interest Group will promote the OTF and contribute to its sustainability.

1.6. Engagement with the general public

We made some attempts to advertise some of the PaNdata ODI developments to the general public, in particular to educators and teachers. At DESY we had a number of informal meetings with members of the Helmholtz Centre Geesthacht (HZG), scientists and education departments in northern Germany responsible for the further education of teachers at a college or high-school level. Aim of the meetings was to identify topics which could be interesting for higher education. One of the topics identified are the tomographic investigation of biological samples.

We are currently working on the implementation of an open access data catalogue for tomographic reconstructions of biological specimen. The data catalogue is based on an ICAT instance with a Drupal web-system and suitable long-term storage to host the scientific data and related information. In an early prototype the easy integration of the system with the Umbrella IdM could be

⁵⁶ <https://www.dectris.com/successstories.html>

⁵⁷ <http://ess-scandinavia.eu/>

⁵⁸ <http://www.ska.ac.za/>

⁵⁹ <http://www.ska.gov.au/>

⁶⁰ <http://www.dcache.org/>

⁶¹ <http://www.pvfs.org/>

⁶² http://wiki.lustre.org/index.php/Main_Page

⁶³ <http://www.fhgfs.com/cms/>

⁶⁴ <http://hadoop.apache.org/>

⁶⁵ <https://confluence.slac.stanford.edu/display/CLNSF/Home>

⁶⁶ <https://indico.desy.de/categoryDisplay.py?categId=333>

demonstrated. Since the system integrates essentially all components of a virtually laboratory, the open access tomography database will serve as one of the virtual laboratories and will become available under science3d.org. The technical implementation is still in an early phase. However, similar efforts are currently pursued at LBNL based on the NERSC webtoolkit⁶⁷ and a django backend. We are in communication with US colleagues hoping to mutually benefit from seemingly similar developments.

Usage of tomographic data and reconstructions for educational purposes requires a more detailed and appropriate description of the experimental technique and the scientific background, a suitable visualization and possibly preparation of printable 3D-models. This will not be possible without significant contributions by the research community. So far, feedback from the scientists has been positive both on making data and information publicly available as well as on the support of the effort. To intensify the collaboration between scientists, educators and infrastructure providers, we intend to hold a workshop with participation of these parties towards the end of the project (~02/2014).

We also presented these efforts (and the data management initiatives) at the “Nacht des Wissens”⁶⁸ in Hamburg, 11/2013 in form of posters⁶⁹ and 3-D printed extinct insects dating back more than 100 million years. The event had – alone at DESY – more than 18.000 visitors, many of which came along the tomography samples.

⁶⁷ <https://newt.nersc.gov/>

⁶⁸ Night of knowledge

⁶⁹ Posters are available from the PaNdata web.

1.7. Workshops and Events organized by PaNdata partners

The following table summarizes the most important events (co-)organized by PaNdata ODI partners.

| Date | Event | Org. | Location | Type | Materials |
|---------------|---|----------------|-------------|-------------------|--|
| 14-15.11.2013 | ESRF Umbrella workshop | ESRF | Grenoble,FR | Open | Agenda |
| 06-07.11.2013 | WP4 ICAT Workshop | ELETTRA | Trieste, IT | Open | |
| 10-11.09.2013 | PaNdata ODI meeting & Umbrella workshop | ELETTRA | Trieste, IT | Internal/ Open | |
| 20-21.06.2013 | 5th PaNdata & CRISP harmonization meeting | HZB | Berlin,DE | Open | Agenda, Slides |
| 20-21.03.2013 | 5th Meeting on Federated Identity for Research Collaborations | PSI | Villigen,CH | Open | Agenda, Participants |
| 13-14.03.2013 | 5th PaNdata ODI Project Meeting | MAX IV | Lund,SE | Internal | Participants, Selected Slides |
| 12-13.03.2013 | ICAT workshop in Lund | MAX IV | Lund,SE | Open | Participants, Slides (Part 1), Slides (Part 2) |
| 04-05.03.2013 | (parallel) Computing in Photon and Neutron Science Applications | DESY | Hamburg,DE | Open | Agenda |
| 21-22.01.2013 | 4th PaNdata & CRISP Harmonisation meeting | Eur.XFEL & PSI | Hamburg,DE | Open | Agenda, Slides Participants |
| 10-11.10.2012 | European Common Affiliation Database Workshop | ESRF | Grenoble,FR | Open | Agenda |
| 28.09.2012 | 4th PaNdata ODI Project Meeting | STFC | Didcot,UK | Internal | Agenda |

PaNdata ODI open workshops and NOBUGS 2012

| Date | Event | Org. | Location | Type | Materials |
|------------|--|------------|-----------|------|--|
| 27.09.2012 | Dawn satellite workshop | DLS & STFC | Didcot,UK | Open | Calendar, Programme |
| 27.09.2012 | Pan European Authentication workshop | DLS & STFC | Didcot,UK | Open | Programme |
| 27.09.2012 | ICAT satellite workshop | DLS & STFC | Didcot,UK | Open | Agenda & Materials, All slides (zip) |

| 24-26.09.2012 | NOBUGS 2012 | DLS & STFC | Didcot,UK | Open | Programme |
|---------------|---|---------------|-------------|----------|---|
| 20-21.09.2012 | NIAC2012 | DLS & STFC | Didcot,UK | Open | Participants & Minutes |
| 18-19.09.2012 | NeXus code camp | DLS & STFC | Didcot,UK | Open | Participants & Minutes |
| Date | Event | Org. | Location | Type | Materials |
| 09-13.07.2012 | SRI 2012 | ESRF & Soleil | Lyon,FR | Open | Agenda |
| 27-28.06.2012 | 3rd PaNdata ODI Project Meeting | Elettra | Trieste,IT | Internal | Slides & Agenda, Participants |
| 13.06.2012 | 3rd PaN-Data & CRISP Harmonisation Meeting | PSI | Zurich,CH | Open | Participants, Slides & Agenda, all slides (zip) |
| 20.03.2012 | ICAT workshop | ILL | Grenoble,FR | Open | Slides, Participants & Agenda |
| 27-28.02.2012 | 1st Joint PNI-HDRI & PaNdata workshop and 2nd PaNdata ODI Project Meeting | DESY | Hamburg,DE | Open | Agenda, Participants, Slides, Report |
| 08.12.2012 | 2nd PaN-Data & CRISP Harmonisation Meeting | DESY | Hamburg,DE | Open | Agenda, Minutes, Participants, Slides |
| 03.11.2011 | PaNdata ODI kickoff meeting | STFC | Didcot,UK | Open | Slides, Participants |
| 02-03.11.2011 | 2nd workshop on Federated identity system for scientific collaborations | STFC | Didcot,UK | Open | Agenda & Slides |