

Publication database.

EUCALL Workshop: User Access at Advanced Laser Light Source

Alexander Wagner

ELI-Beamlines

Dolní Břežany, 22.09.2017

HELMHOLTZ RESEARCH FOR
GRAND CHALLENGES



Overview



- > General considerations
- > Content
- > Data



Publication database.

Part I: General considerations



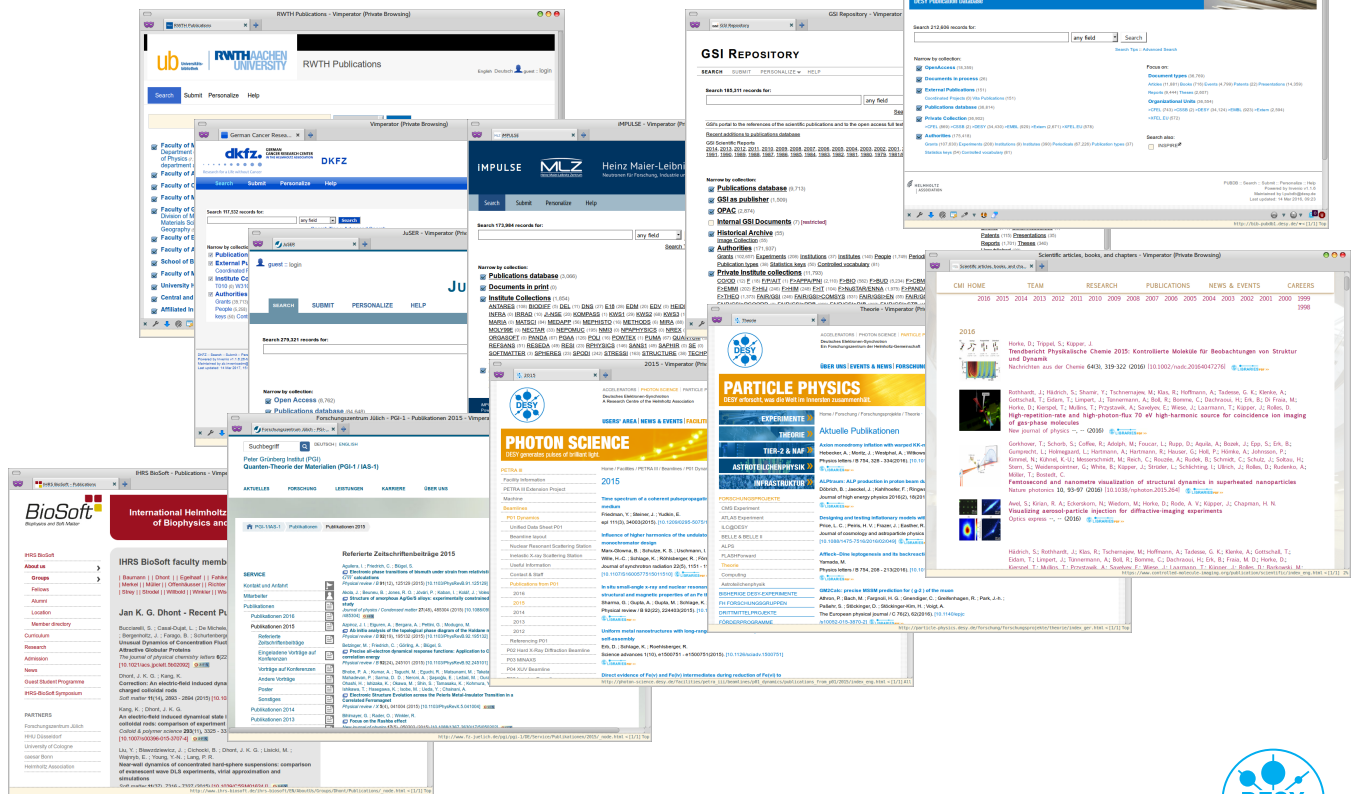
Stakeholders

Who is using your pubddb? How are they using it?

- > Scientists (e. g. publication lists)
- > non-scientific Staff (e. g. administration)
- > General public
- > Facility (e. g. web pages)
- > Funders (e. g. EU: OpenAIRE)

What services can you provide to them?
Are people using the database or only the data?

Example: JOIN²



Alexander Wagner | ELI-Beamlines | Dolní Břežany, 22.09.2017 | Page 5

Goals

What do you want to achieve?

- Visibility (think web, OpenAccess)
- OpenAccess (requires much more than a database; e. g. policies, copyright, work flows, interfaces)
- Repository and archive (e. g. What was our impact after n years of operation?)
- Reporting (also: application for new funding, new proposals)

Always think OpenAccess.

Never forget Green OpenAccess: go for 100%

Remember: "All that is gold does not glitter..." (J.R.R. Tolkien, *The Riddle of Strider* in "The Fellowship of the Rings")



Alexander Wagner | ELI-Beamlines | Dolní Břežany, 22.09.2017 | Page 6

Publication database.

Part II: Content



Items

Which items to whom and how to access?

- > Bibliographic data (OpenAccess mandates: will no longer be sufficient)
- > Full text documents (easy future access, build an archive)
- > Access rights (restrictions, embargoes)

Document types:

- > Publications (really only "own" results?)
- > Data(?) (will pose different problems)
- > Other documents (theses, instructions, guidelines, documentations)

Do and Don't

- > Build tools scientists **want to** use and non-scientists **can** use
- > Do **not** build a **lighthouse**
- > **Minimize** necessary manual **work** (importer, but don't fall for batch ingestion!)
- > **Maximize** re-use (structured exports, web integration)
- > **Integrate** with your **library** (catalogue, document delivery, repository, publishing house...)
- > Imagine **workflows**:
 - Who knows what at when?
 - What can be derived, what is only internal knowledge?
 - How can available data be reused? (What derives from the proposal e. g.?)

...give me an universal, unique identifier and ...

Archimedes?



Alexander Wagner | ELI-Beamlines | Dolní Břežany, 22.09.2017 | Page 9



Import data ⓘ e.g. DOI, arXiv, PUBMED...

Relevant for Reporting * ⓘ ☒ yes ☐ no

Group(s) directly involved * ⓘ
Select or type in name, shortcut (e.g. ATLAS, FS-PE, MK)

POF III: Topic/Research Theme/Facility * ⓘ
Select from list or type ID, Name of POF-Topic...

Grant name / Proposal No. ⓘ
e.g. EU project, FS proposal number (e.g. I-20120768)

Author(s) / Editor(s) * ⓘ

Abicht, F. [Extern] Author	
Braenzel, J. [Extern] Author	
Priebe, G. -> Priebe, Gerd (XFEL.EU: gerd.priebe@desy.de / Eur.XFEL) Author	✕
Koschitzki, Ch. -> Koschitzki, Christian (DESY: christian.koschitzki@desy.de / ZEU-PITZ) Author	✓ ✕
Andreev, A. A. [Extern] Author	✓ ✕
Nickles, P. V. [Extern] Author	✓ ✕
Sander, W. [Extern] Author	✓ ✕
Schnürer, M. [Extern] Author	✓ ✕

Start typing lastname and select...

Beamline/Experiment/Facility * ⓘ

xfel
Experiments at XFEL (XFEL)
FXE: Femtosecond X-ray Experiments (FXE)
Facility (machine) XFEL (XFEL)
HED: High-Density matter Experiments (HED)
MID: Materials Imaging and Dynamics (MID)
SCS: Spectroscopy & Coherent Scattering (SCS)
SPB: Single Particles, clusters & Biomolecules (SPB)

Title * ⓘ

Tracing dynamics of laser-induced fields on ultrathin foils using complementary imaging with streak deflectometry

edit

Title preview:

Tracing dynamics of laser-induced fields on ultrathin foils using complementary imaging with streak deflectometry

Journal * ⓘ Physical review accelerators and beams

DOI ⓘ 10.1103/PhysRevAccelBeams.19.091302

Volume * ⓘ 19

Issue ⓘ 9

Pages * ⓘ 091302

Publication Year * ⓘ 2016

Model data, not answers to today's questions

- > Think big (e. g. JOIN²: 7 institutions, serving 25.000 people, about 355.000 records)
- > Define publicly available values (e. g. bibliographic values)
- > Define specific values (data nobody else has/knows, e. g. proposal id)

The more details, the more work for **Scientists**



- > Author disambiguation:  ORCID (login, proposals, submissions)
- > Completeness and correctness (review work flow, web output)
- > Derive from proposals e. g.
 - funding information
 - scientific subject
 - coupling to submission

Rethink "Call for Submission"

- > Not fast enough (e. g. EU OpenAccess Mandate)
- > Asymmetric work load (usually everything in a few weeks)
- > Incomplete (something happened a year ago, even worse 12 months...)

Submission should happen right after publication, better before.

(approval forms?)



Storage

- > Use established standards (“established” may be old)
- > “Authority controlled ” data (normalize semantically)
- > Persistent, common, globally unique IDs (e. g. ORCID, DOI, arXiv etc.)
- > Run archival safe (OAIS)

Involve your Library

...but also believe them if something is impossible.



Output

- > Think in structured outputs (BibTeX, EndNote, RIS...)
- > Your web pages are *life online reports* (will require caching)
- > Enhance visibility for *machines* (schema.org, microformats, semantic markup...)
- > Facilitate item interchange (e. g. OAI-PMH, especially beyond Dublin Core)
- > Report only based on your publication database
- > Minimize manual efforts for common reports
 - Structured exports (e. g. BibTeX)
 - Batch processing (e. g. L^AT_EX or pandoc)
 - Common formats (e. g. BibL^AT_EX, CSL)



Publication database.

Part III: Data



Data

There is no *best practice* for “Data” in 2017

However, note:

- > Only **published Data** is (re-)usable
- > Publishing implies **availability for ever** (“citation safe”)
- > The **publisher is responsible** for eternal storage and access
- > There is no fall back (no National Libraries, Portico or the like)
- > Publishing is not (only) an IT problem:
 - Metadata (findability)
 - Consistency (reliability)
 - Formats (availability)
 - Long term preservation and migration (a bit stream is usually not enough)



- > In your **Repository**: Small records (non-IT-issues still apply)
- > **Global data centres**: Large records

Major headache

Intermediary sized data

- > Samples: , ,  ...
- > Registry:  (if there is something suitable, use it)
- > General & Software:  (up to 50GB per dataset, more on request)



Data @ PubDB

Report/Journal Article

PUBDB-2016-05430

Performance of the EUDET-type beam telescopes

Jansen, H. (Corresponding author)*; Spannagel, S.*; Bulgheroni, A.; Claus, G.*; Corrin, E.; Cussans, D. G.; Dreyling-Eschweiler, J.; Eckstein, D.*; Eichhorn, T.*; Goffe, M.*; Gregor, I. M.*; Haas, D.*; Muhl, C.*; Perrey, H.*; Peschke, R.*; Roloff, P.*; Rubinsky, I.*; Winter, M.

EPJ Techniques and Instru

This record in other databases: A

Please use a persistent id in cita

Report No.: DESY-16-055; a

zenodo, (2016) [10.5281/zeno

Abstract: Test beam measurem

the EUDET project. The beam te

time stamp information on partic

acquisition framework via predel

two centre pixel planes using all

Lines are performed to estimate

electron/positron beam, the track

measured intrinsic resolution. To

MIMOSA26 sensors at a sensor

the material traversed and allow

Keyword(s): detector; design; e

programming; track data analysi

Classification:

• ddc:540

Contributing Institute(s):

1. LHC/CMS Experim

2. LHC/ATLAS Experi

Research Program(s):

1. 632 - Detector tech

Experiment(s):

1. LHC: CMS

2. LHC: ATLAS

Dataset

PUBDB-2017-01391

Dataset For The 'Performance Of The Eudet-Type Beam Telescopes' Publication (Epj Ti)

zenodo

August 2, 2016

Dataset Open Access

Indexed in OpenAIRE

Publication date: August 2, 2016

DOI: DOI: 10.5281/zenodo.59255

Keyword(s): reference dataset, beam telescope, EUDET, DATURA

Published in: EPJ Techniques and Instrumentation 3 pp. 7.

Related identifiers: Supplementary material: 10.5281/zenodo.49065

License (for files): CC Creative Commons Attribution 4.0

Share Cite as Start typing a citation style...

Export BibTeX CSL DataCite Dublin Core JSON MARCKM CiteMendley

Files (24.8 GB)

Name	Size	Download
readme.txt	1.7 kB	Download
md5-3609f9c6044b7a2b7b64a58417767b8	869.2 MB	Download
run000060.raw	324.4 MB	Download
md5-85e1cc48b736f9e79c823e9f0e571	140.8 MB	Download
run000061.raw	112.9 MB	Download
md5-049138f8d53257844d0c8e0e2e7		
run000062.raw		
md5-0b6a0a3d55f7ab9a9f6a2770b0f9		
run000063.raw		
md5-3d7e69a2b9bacc1c47f9662d7534ed		

This work is supported by the Commission of the European Communities under the 6th Framework Programme 'Structuring the European Research Area', contract number RI3-G26126. Furthermore, strong support from the Helmholtz Association and the BMBF is acknowledged.

Contributing Institute(s):

Research Program(s):

1. 632 - Detector technology and systems

Experiment(s):

1. LHC: CMS

2. LHC: ATLAS

Preprint/Internal Report

Jansen, H. (Corresponding author)*; Spannagel, S.*; Bulgheroni, A.; Claus, G.*; Corrin, E.; Cussans, D. G.; Dreyling-Eschweiler, J.; Eckstein, D.*; Eichhorn, T.*; Goffe, M.*; Gregor, I. M.*; Haas, D.*; Muhl, C.*; Perrey, H.*; Peschke, R.*; Roloff, P.*; Rubinsky, I.*; Winter, M.

Performance of the EUDET-type beam telescopes

OPEN ACCESS Files Fulltext by arXiv.org



Thank you!



Alexander Wagner
Deutsches Elektronen-Synchrotron
Central Library

Tel.: +49-40-8998-1758

alexander.wagner@desy.de

 0000-0001-9846-5516

<http://library.desy.de>

This document is available as <https://doi.org/10.3204/PUBDB-2017-10906>

Typeset by lua^AT_EX



Alexander Wagner | ELI-Beamlines | Dolní Břežany, 22.09.2017 | Page 19

