JOIN^2.

Invenio for Fraunhofer

Alexander Wagner, *for* ***JOIN***^{2}

Fraunhofer IRB

Stuttgart, 18.09.2017

HELMHOLTZ GEMEINSCHAFT

DESY

Overview

> Who is **JOIN**^{2}?

DESY

Overview

- Who is **JOIN**^{2}?
- Fraunhofer requirements

DESY

Overview

- Who is **JOIN**^{2}?
- Fraunhofer requirements
- What happened in 2016?

DESY

JOIN^2.

Part I: Who is 

Partner

Deutsches Elektronen-Synchrotron	$\approx 2000 + 3000$
Deutsches Krebsforschungszentrum	internal production
Forschungszentrum Jülich	$> \approx 5000 + 1000$
GSI Helmholtzzentrum für Schwerionenforschung	≈ 1050
Maier-Leibniz-Zentrum, Garching	$\approx 300 + 1000$
RWTH Aachen	≈ 9000
Museum Zitadelle Jülich	

Partner

Deutsches Elektronen-Synchrotron $\approx 2000 + 3000$

Deutsches Krebsforschungszentrum internal production

Forschungszentrum Jülich $> \approx 5000 + 1000$

50

0 + 1000

00

JOIN^2

Partner

Deutsches Elektronen-Synchrotron ≈ 2000 + 3000

Deutsches Krebsforschungszentrum internal production

Forschungszentrum Jülich > ≈ 5000 + 1000

50

0 + 1000

00

JOIN^2

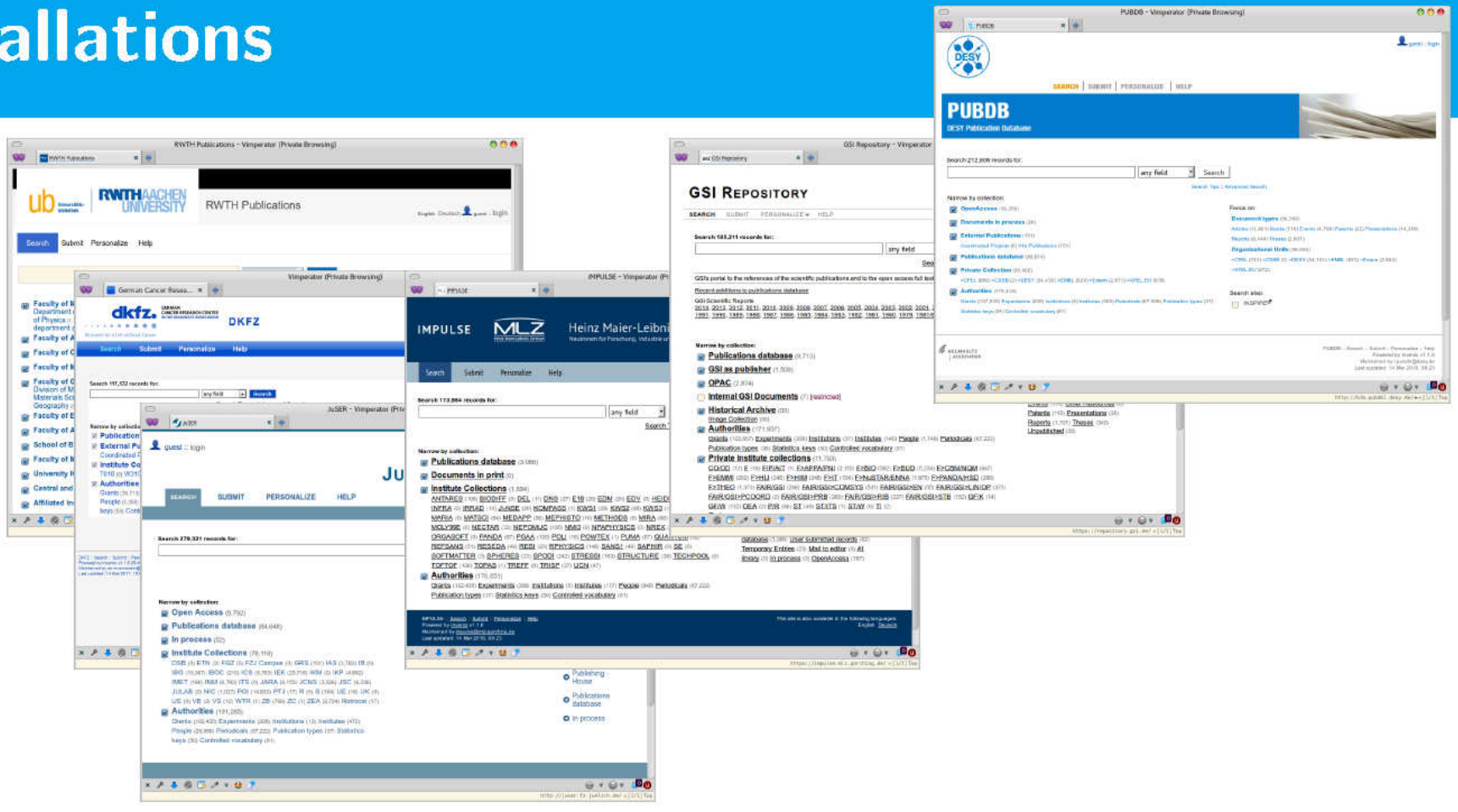
Online an 7 Installationsn for

ca. 20.000 + 5000 User

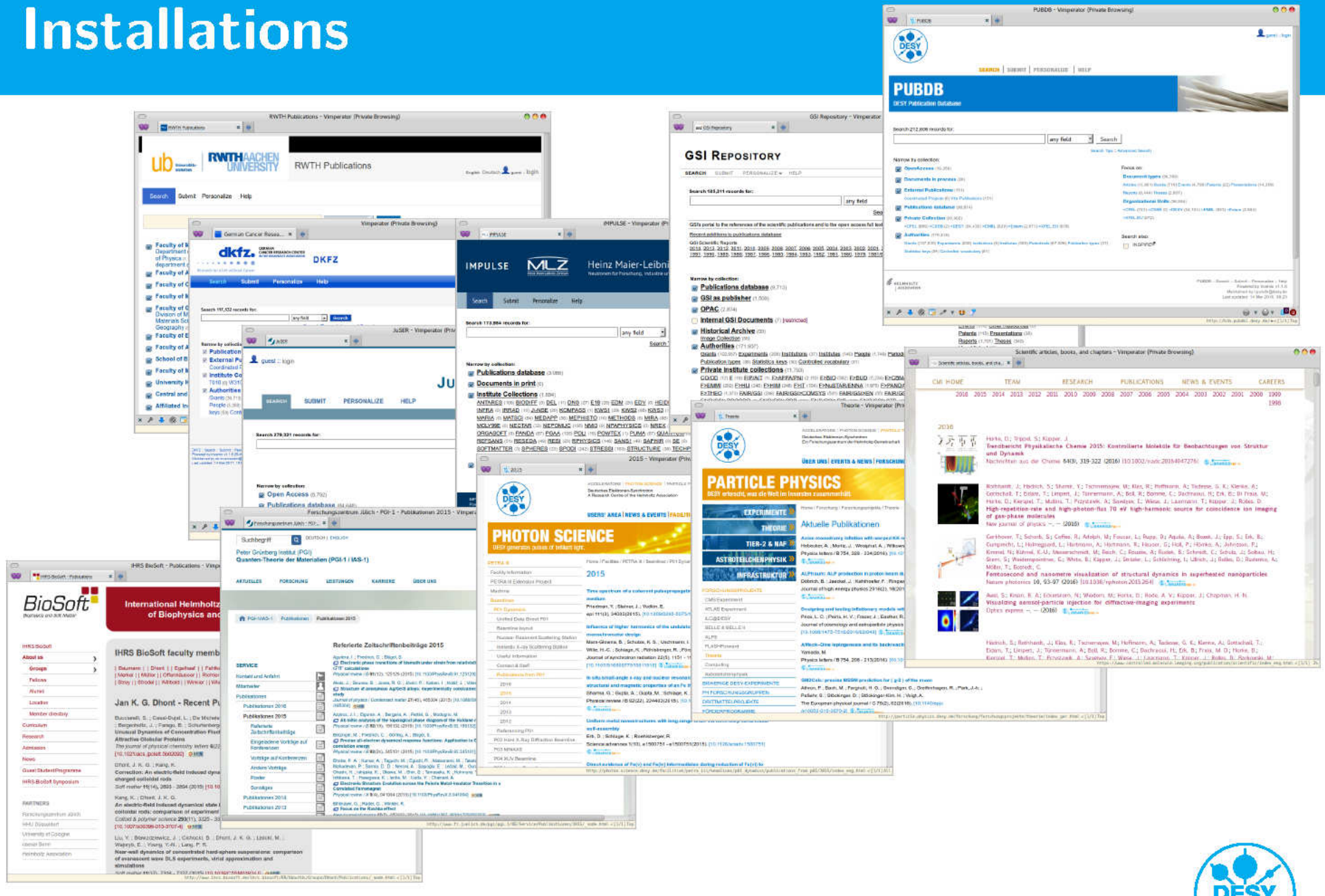
ca. 355.000 records, 43.800 OpenAccess,

≈ 117.000 authority records

Installations



Installations



OpenAccess Repository Ranking 2015

- 181 repositories from D/AT/CH
- 160 German repositories

DESY

OpenAccess Repository Ranking 2015

- 181 repositories from D/AT/CH
- 160 German repositories
- 6. JuSER (76/100 points)
- 17. pubdb (72/100 points)

Many issues are non-technical, but organizational (policies etc.)
http://repositoryranking.org/

Systems background

As a community **JOIN**^{2} is slower then CERN.

DESY

Systems background

As a community **JOIN**^{2} is slower then CERN.

Current Status:

- Python 2.6 (while compatible with 2.4)
- MySQL or MariaDB
- Running on Debian, FreeBSD, Redhat, Scientific Linux

DESY

Systems background

As a community **JOIN**^{2} is slower then CERN.

Current Status:

- Python 2.6 (while compatible with 2.4)
- MySQL or MariaDB
- Running on Debian, FreeBSD, Redhat, Scientific Linux

Invenio versions:

- 1.1.7 all **JOIN**^{2}-instances

DESY

Systems background

As a community **JOIN**^{2} is slower then CERN.

Current Status:

- Python 2.6 (while compatible with 2.4)
- MySQL or MariaDB
- Running on Debian, FreeBSD, Redhat, Scientific Linux

Invenio versions:

- 1.1.7 all **JOIN**^{2}-instances
- 1.2 caused major performance issues
- 2.x is not recommended for use
- 3.0 is not ready for use (Complexity(**JOIN**^{2}) $\approx$ Complexity(CDS) $\gg$ Complexity(Zenodo))

DESY

JOIN^2.

Part II: Fraunhofer Requirements

Technical requirements — met by JOIN^2

- Operating system (Ubuntu or Debian, both fine)
- Database (MySQL/MariaDB)
- Checksums for documents (md5, Invenio is largely OAIS compliant)
- Programming languages (Python, JavaScript, Perl)
- Authentication (LDAP, Shibboleth, local)
- **JOIN^2** release cycle: monthly (automatised roll out)
- Extensibility (python, but *cooperative*!)

DESY

Technical requirements — partly met by JOIN^2

- Community: 7 productive instances, 3 in preparation

  (Prod: DESY, DKFZ, FZJ, GSI, MLZ, MZJ, RWTH, Prep: DZNE(?), JINR, UWH(?))
- Software docs (mainly pydoc: RTFS)
- Templating (own system, **JOIN^2** implements the templates)
- Languages: DE, EN, but not in all areas (i. e. submission forms are english)
- PI Minting (DOI via DataCite, Handle, not (yet): URN)

DESY

Submission — met by JOIN^2

- Bibliographic records via web (every user, simplified by importers)
- Single item importer:
  - DOI (CrossRef and DataCite)
  - Inspire, WoS, PubMed
  - ISBN (GBV, LoC)
  - Data reuse (join, part, recid, link)
- Append full texts (No restriction on file formats)
- Grant assignment (= Funding-IDs)
- Collection structure (however, different concept!)

DESY

Submission — not met JOIN^2

No Batch upload

DESY

Submission — not met JOIN^2

No Batch upload

- Poor data (missing normalizations, e. g. journals)
- Missing data (Funding, especially local schemes, experiments etc.)
- Need for manual cataloguing

  (bibedit requires STAFF rights, cumbersome for normalized entries)
- De-duping necessary
- Only one proof reader

DESY

Submission — not met JOIN^2

No Batch upload

- Poor data (missing normalizations, e. g. journals)
- Missing data (Funding, especially local schemes, experiments etc.)
- Need for manual cataloguing

  (bibedit requires STAFF rights, cumbersome for normalized entries)
- De-duping necessary
- Only one proof reader

web submission is much more efficient

DESY

Content — met by JOIN^2

- Grants
  - EU funding: automatically available (OAI harvesting via **JOIN^2**)
  - Grants are shared in **JOIN^2**
  - Local funding schemes possible
- Record linking (e. g.. Book and chapter, proceedings and article, article and data)
- DDC from journal (automatic, STAFF could change)
- Journal normalization (shared authority base of ≈ 75.000 journals, automatic update)
- Author identification (ORCiD ready)

DESY

Content — not met by JOIN^2

- Keyword suggestions (free and normalized)
- Series normalization (however, possible)

DESY

Output — met by JOIN^2

- Bibliographic formats (BibTeX, EndNote (XML, Text), RIS)
- DC Simple (BASE/OpenAIRE compliant DC v3 for repositories)
- DataCite v3 (exposes ORCiD, thus can feed ORCiD)
- ORCiD (requires membership, currently only Jülich)
- schema.org (GSBLST, HD, HB)
- Google Scholar (However, no way to submit to GoogleScholar any more)

Visibility ...

DESY

Output — met by JOIN^2

- Bibliographic formats (BibTeX, EndNote (XML, Text), RIS)
- DC Simple (BASE/OpenAIRE compliant DC v3 for repositories)
- DataCite v3 (exposes ORCiD, thus can feed ORCiD)
- ORCiD (requires membership, currently only Jülich)
- schema.org (GSBLST, HD, HB)
- Google Scholar (However, no way to submit to GoogleScholar any more)

Visibility ...

...is pretty good, especially in search engines.

DESY

Output — met by JOIN^2

- Bibliographic formats (BibTeX, EndNote (XML, Text), RIS)
- DC Simple (BASE/OpenAIRE compliant DC v3 for repositories)
- DataCite v3 (exposes ORCiD, thus can feed ORCiD)
- ORCiD (requires membership, currently only Jülich)
- schema.org (GSBLST, HD, HB)
- Google Scholar (However, no way to submit to GoogleScholar any more)

Visibility ...

...is pretty good, especially in search engines.

JOIN^2-instances feed the institutes websites directly.

DESY

Output — not met by JOIN^2

JOIN^2 partly meets:

- REST (access any format via http + query)
- xMetaDiss+ (in final preparation at RWTH together with DNB)

JOIN^2 does not meet:

- CERIF / CRIS (no use case, should not be too complex)
- SPARQL, SRU and friends (no use case, would be complex)

DESY

GUI — met by JOIN^2

- Browser compatibility (JOIN^2 currently tests FireFox, Chrome, Safari)
- Online help (co-operative via Wiki, German and real(!) English)
- Citation hint (HD displays like a journal citation)
- RSS
- Social Media uplinks
- Baskets and Alerts
- Storable complex searches

DESY

GUI — not met by JOIN^2

- Facetting
- HTML5 / CSS3 (Invenio 1.x and Browser restrictions)
- mobile view / responsive design (Invenio 1.x restriction)
- Internal PDF-Viewer (every browser has this anyway)

DESY

GUI — not met by JOIN^2

- Facetting
- HTML5 / CSS3 (Invenio 1.x and Browser restrictions)
- mobile view / responsive design (Invenio 1.x restriction)
- Internal PDF-Viewer (every browser has this anyway)

All is met in Invenio 3

DESY

GUI — not met by JOIN^2

- Facetting
- HTML5 / CSS3 (Invenio 1.x and Browser restrictions)
- mobile view / responsive design (Invenio 1.x restriction)
- Internal PDF-Viewer (every browser has this anyway)

All is met in Invenio 3

However, Invenio 3 *ready for production* may need another one or two years.

DESY

Part III: What happened recently?

Publishing and OpenAccess

DataCite integration (JOIN^2 Developers Meeting, 21. - 23. Mar 2016)

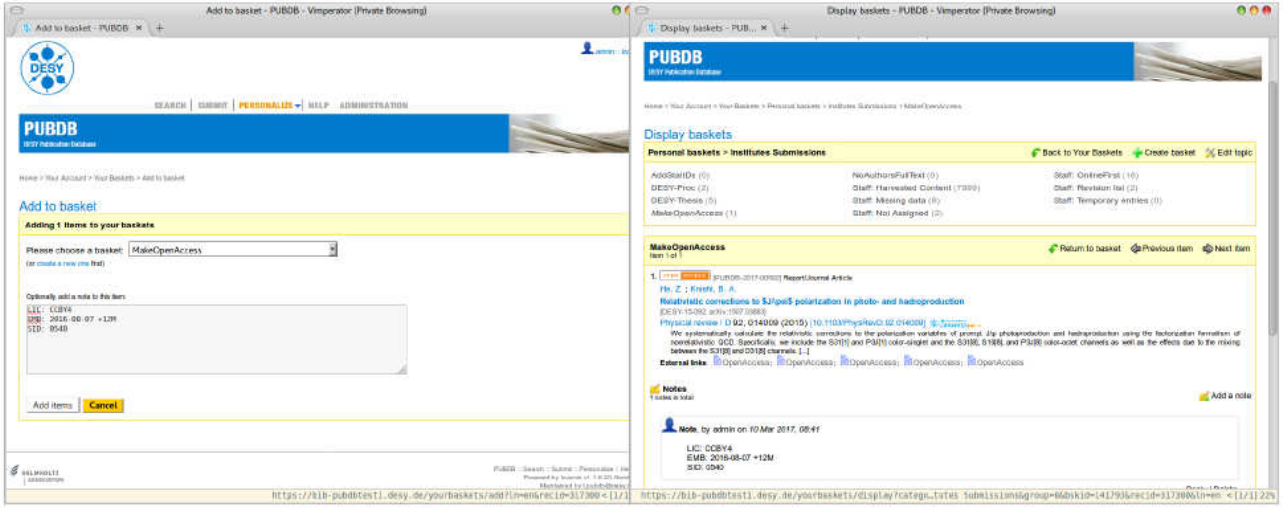
- define metadata format (currently v3, nothing really new in v4)
- allow update (metadata and URL)
- Automatic DOI minting (upon release to OpenAccess)

DESY

Publishing and OpenAccess

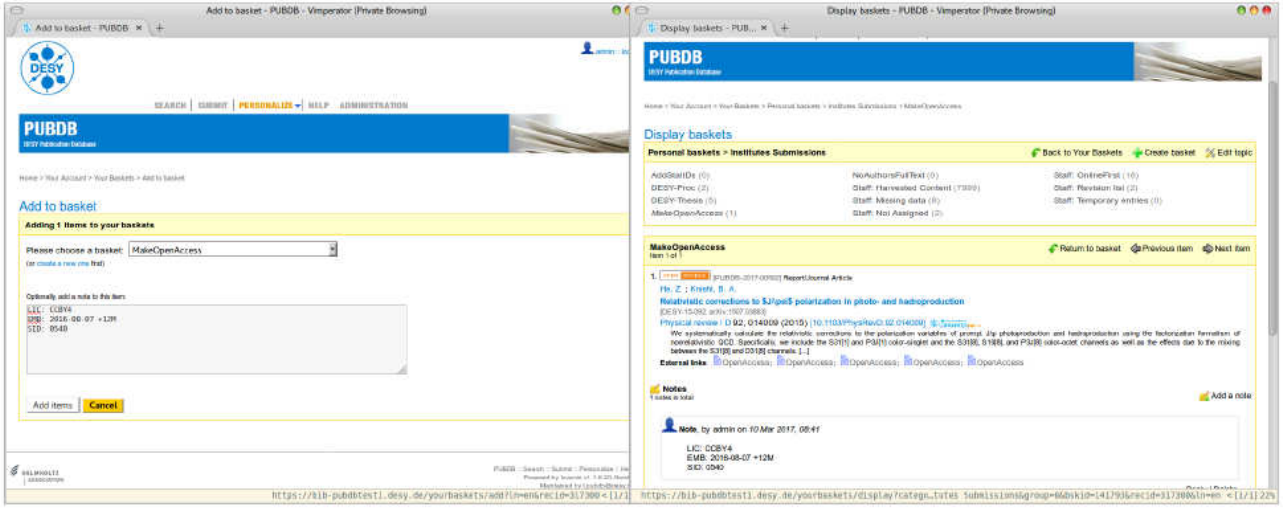
DataCite integration (JOIN² Developers Meeting, 21. - 23. Mar 2016)

- define metadata format (currently v3, nothing really new in v4)
- allow update (metadata and URL)
- Automatic DOI minting (upon release to OpenAccess)



Publishing and OpenAccess

- NEW **DataCite integration** (JOIN² Developers Meeting, 21. - 23. Mar 2016)
  - define metadata format (currently v3, nothing really new in v4)
  - allow update (metadata and URL)
  - Automatic DOI minting (upon release to OpenAccess)



- NEW **ORCID upload** (cf. "ORCID implementation in Invenio 1.1")
- NEW **xMetaDiss+** (DNB harvesting)
- **Rework Online help** (> 60 twiki pages, to be translated)

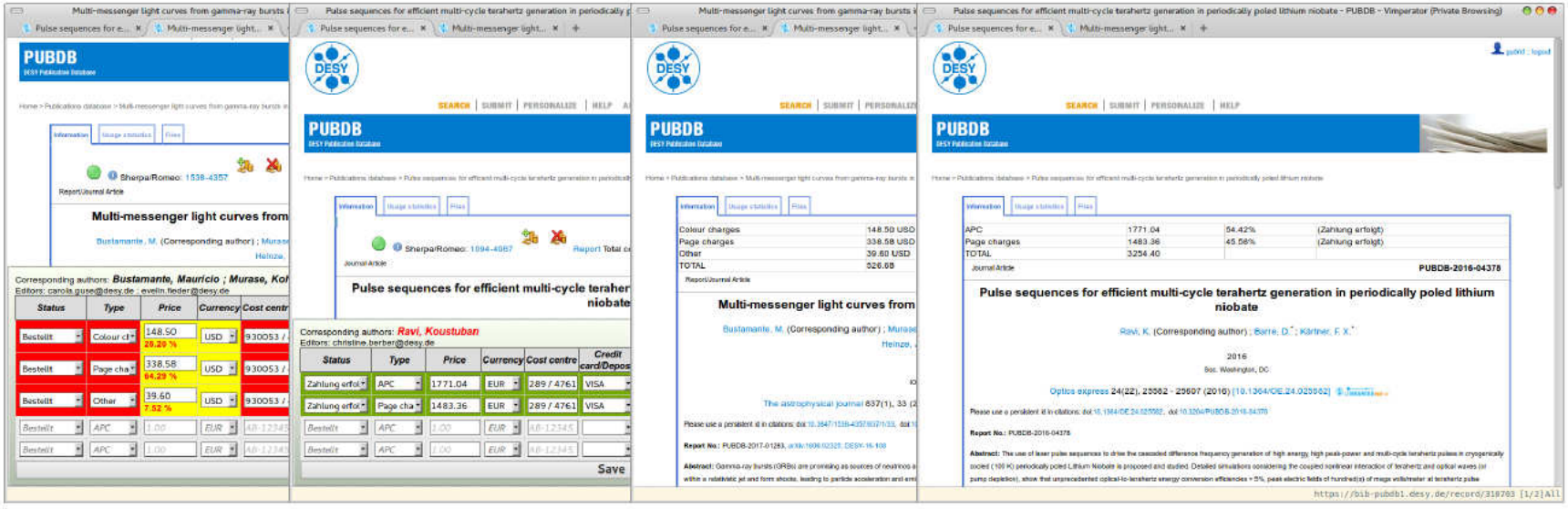
Handling of Article Processing Charges

- easy fee addition (similar to book pruchases)
- Payment status and accounting
- OpenAPC-de (APC data delivery)

DESY

NEW Handling of Article Processing Charges

- easy fee addition (similar to book pruchases)
- Payment status and accounting
- OpenAPC-de (APC data delivery)



cf. “Article Processing Charges and OpenAPC”

Publishing workflow

Automatic publication workflow (DESY-PROC, DESY-THESIS)

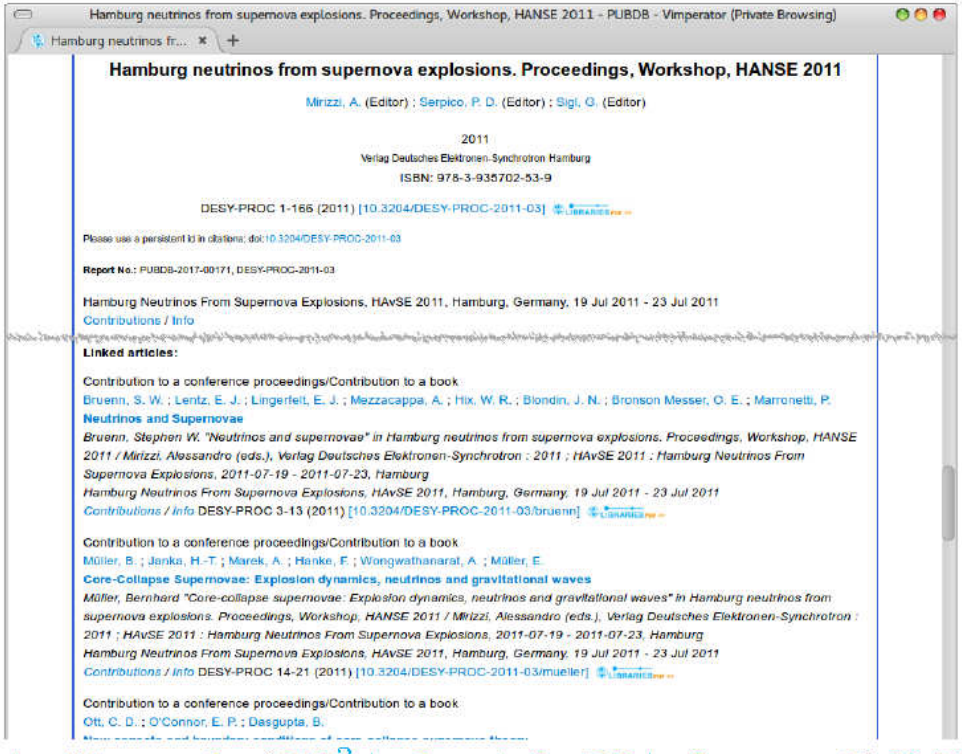
- volume creation
- article publication and DOI minting
- data delivery to INSPIRE (via DESY ingest work flow)

DESY

NEW Publishing workflow

Automatic publication workflow (DESY-PROC, DESY-THESIS)

- volume creation
- article publication and DOI minting
- data delivery to INSPIRE (via DESY ingest work flow)


Hamburg neutrinos from supernova explosions. Proceedings, Workshop, HANSE 2011 - PUBDB - Vimperator (Private Browsing)

Hamburg neutrinos fr...

Hamburg neutrinos from supernova explosions. Proceedings, Workshop, HANSE 2011

Mirizzi, A. (Editor) ; Serpico, P. D. (Editor) ; Sigl, G. (Editor)

2011

Verlag Deutsches Elektronen-Synchrotron Hamburg

ISBN: 978-3-935702-53-9

DESY-PROC 1-166 (2011) [10.3204/DESY-PROC-2011-03]

Please use a persistent id in citations: doi:10.3204/DESY-PROC-2011-03

Report No.: PUBDB-2017-00171, DESY-PROC-2011-03

Hamburg Neutrinos From Supernova Explosions, HAvSE 2011, Hamburg, Germany, 19 Jul 2011 - 23 Jul 2011

Contributions / Info

Linked articles:

Contribution to a conference proceedings/Contribution to a book

Bruenn, S. W. ; Lentz, E. J. ; Lingerfelt, E. J. ; Mezzacappa, A. ; Hix, W. R. ; Blondin, J. N. ; Bronson Messer, O. E. ; Marronetti, P.

Neutrinos and Supernovae

Bruenn, Stephen W. "Neutrinos and supernovae" in Hamburg neutrinos from supernova explosions. Proceedings, Workshop, HANSE 2011 / Mirizzi, Alessandro (eds.), Verlag Deutsches Elektronen-Synchrotron : 2011 ; HAvSE 2011 : Hamburg Neutrinos From Supernova Explosions, 2011-07-19 - 2011-07-23, Hamburg

Hamburg Neutrinos From Supernova Explosions, HAvSE 2011, Hamburg, Germany, 19 Jul 2011 - 23 Jul 2011

Contributions / Info DESY-PROC 3-13 (2011) [10.3204/DESY-PROC-2011-03/bruenn]

Contribution to a conference proceedings/Contribution to a book

Müller, B. ; Janka, H.-T ; Marek, A. ; Hanke, F. ; Wongwathanarat, A. ; Müller, E.

Core-Collapse Supernovae: Explosion dynamics, neutrinos and gravitational waves

Müller, Bernhard "Core-collapse supernovae: Explosion dynamics, neutrinos and gravitational waves" in Hamburg neutrinos from supernova explosions. Proceedings, Workshop, HANSE 2011 / Mirizzi, Alessandro (eds.), Verlag Deutsches Elektronen-Synchrotron : 2011 ; HAvSE 2011 : Hamburg Neutrinos From Supernova Explosions, 2011-07-19 - 2011-07-23, Hamburg

Hamburg Neutrinos From Supernova Explosions, HAvSE 2011, Hamburg, Germany, 19 Jul 2011 - 23 Jul 2011

Contributions / Info DESY-PROC 14-21 (2011) [10.3204/DESY-PROC-2011-03/mueller]

Contribution to a conference proceedings/Contribution to a book

Ott, C. D. ; O'Connor, E. P. ; Dasgupta, B.

DESY

Reporting

- Reporting (cf. “The usages of JOIN2 authority records”)
  - HGF reporting display improvements
  - LaTeX output improvements (Biber, BibLaTeX, LuaLaTeX)
  - NEW APC cost reporting (cf. “Article Processing Charges and OpenAPC”)
- Update of Journal records
  - Minor software updates (e. g. merging algorithm, RSS linkage, RDA changes)
  - DOAJ cleanup (remove all old DOAJ records and only populate the new ones)
  - NEW Reuse of data by KIT Karlsruhe

DESY

Invenio as a library system

- NEW Hook up login with authority records:
  - Allow multiple email addresses per login
  - Allow multiple login methods to one account
  - Allow library cards
- NEW SIP based self service terminals (reuse code from Evergreen, `Inline::Python`)
- NEW add holdings to bibliographic (Marc 8527_ and 8767_)
- > Migrate ExLibris Aleph:
  - bibliographic
  - holdings
  - circulation history
  - patron data
  - automatic dumping
  - semi-automatic loading

(cf. "Invenio as a library system")

DESY

People behind JOIN^2

- Dagmar Sitek[a]
- Gudrun Friedburg[a]
- Martin Köhler[b]
- Zaven Akopov[b,j,1]
- Tomasz Pazera[b,1]
- Robert Thiele[b]
- Alexander Wagner[g,b]
- Katrin Große[d]
- Stefan Hesselbach[e]
- Bernhard Mittermaier[g]
- Torsten Bronger[g]
- Claudia Frick[g]
- Anna Fründ[g]
- Christoph Holzke[g]
- Heike Lexis[g]
- Cornelia Plott[g]
- Jürgen Neuhaus[h]
- Connie Hesse[h]
- Björn Pedersen[h]
- Jörg Pulz[h]
- Ulrike Eich[i]
- Louai Barake[i]
- Corinna Brückener[i]
- Abdoulaye Diallo[i]
- Roland Rappmann[i]
- Dominik Schmitz[i]
- Edmund Wollgarten[i]

[a] DKFZ Central Library, [b] DESY Central Library, [c] Project Inspire, [d] GSI Library, [e] GSI Core IT, [g] Forschungszentrum Jülich, Central Library, [h] MLZ Garching, [i] RWTH Aachen, University Library, [j] European XFEL, [1] (retired)

with a little help from our friends

CERN Library, CERN IT, iNSPIRE

Vielen Dank!

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-2017-10826

Typeset by lualaTeX