

Forschungsinfrastruktur.

Workshop: Repositorien jenseits des Dublin Core

Alexander Wagner, für JOIN²

Open Access Tage 2014

Köln, 09.09.2014



Überblick



- > Ausgangslage
- > **JOIN²**: Wünsche und Partner
- > Implementation
- > Normsätze



Forschungsinfrastruktur.

Teil I: Ausgangslage



Repositorien, Bibliographien etc.

- > Helmholtz-Gemeinschaft = Großgeräte + X
(Schiffe, Flugzeuge, Forschungsstationen, Teilchenbeschleuniger...)
- > ständiges Reporting, permanente Evaluierung
 - Programmorientierte Förderung (POF $\approx$ "DFG für HGF")
 - Teilinstitute, Gruppen, Projekte...
 - Personen (z. B. Institutsleiter, Gruppenleiter...)
- > Meldepflicht für Publikationen

Zentrale Nachweissysteme, Repositorien sind oft nur ein Addon
bisher: verschiedene **Eigenentwicklungen**

Vergleichsweise **aufwändige** Datenbanken

ABER: für **Wissenschaftler** kaum nutzbar ☹

- > Wissenschaftler
 - 1 Schnittstellen (Im-/Export: BibTeX, EndNote, ...)
 - 2 Private Sammlungen (Einbindung in den Publikationsworkflow)
 - 3 Integration von "Bibliothek" (z. B. Linkresolver, Volltexte, Publikationslisten...)
- > Sichtbarkeit der Forschungsergebnisse (mehr Volltexte, zeitnaher Eintrag)
- > zukunftssicher (keine Eigenentwicklung)
- > Standardisierung und Normierung (Schnittstellen, Metadaten, ...)
- > detaillierte Auswertungen (komplexe Schlüssel, externe Quellen, Bibliometrie...)

flexibles aber **standardisiertes** Metadatenschema,
deutlich größer als Dublin Core



JOIN² Partner



Deutsches Elektronensynchrotron, Zentralbibliothek

≈ 2000 + 3000



Forschungszentrum Jülich, Zentralbibliothek

≈ 5000 + 1000



GSI Helmholtzzentrum für Schwerionenforschung, Bibliothek + Base-IT

≈ 1050



Deutsches Krebsforschungszentrum, Bibliothek, NEU 2014

≈ 3000



JOIN²: Offen für neue Partner!

≈ 300 + 1000



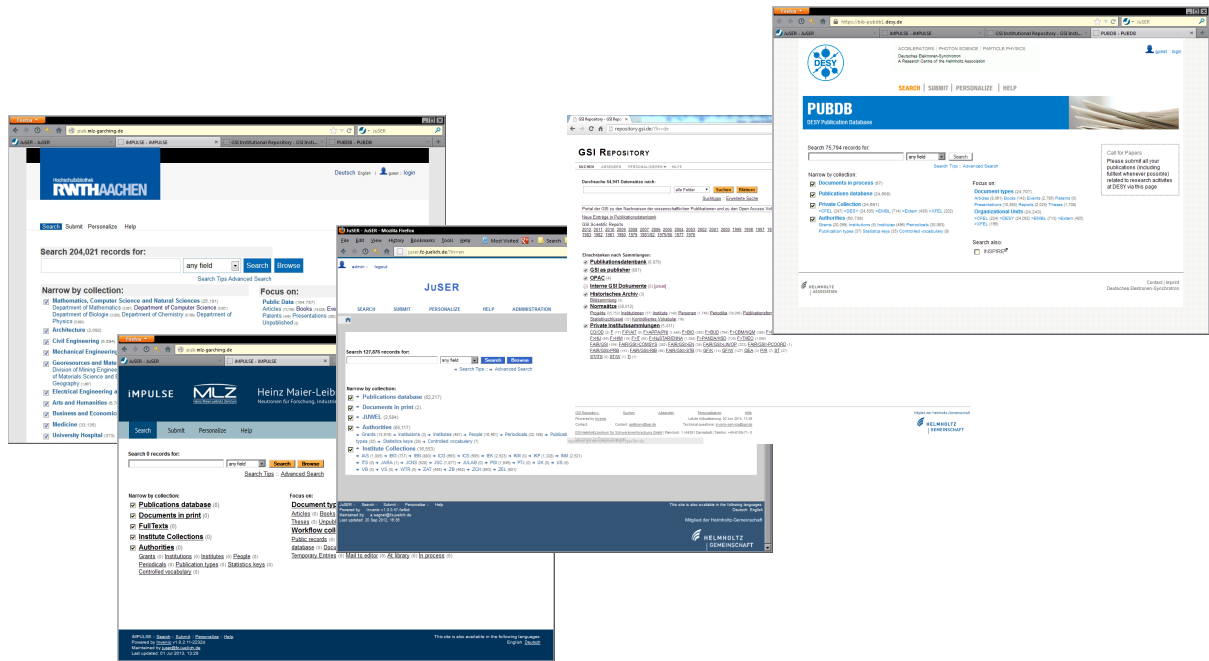
Online für ≈ **20.000** Mitarbeiter

≈ 9000

+ ≈ 5000 Gastwissenschaftler + Besucher

(≈ 260.000 Dokumente + 80.000 Normsätze)





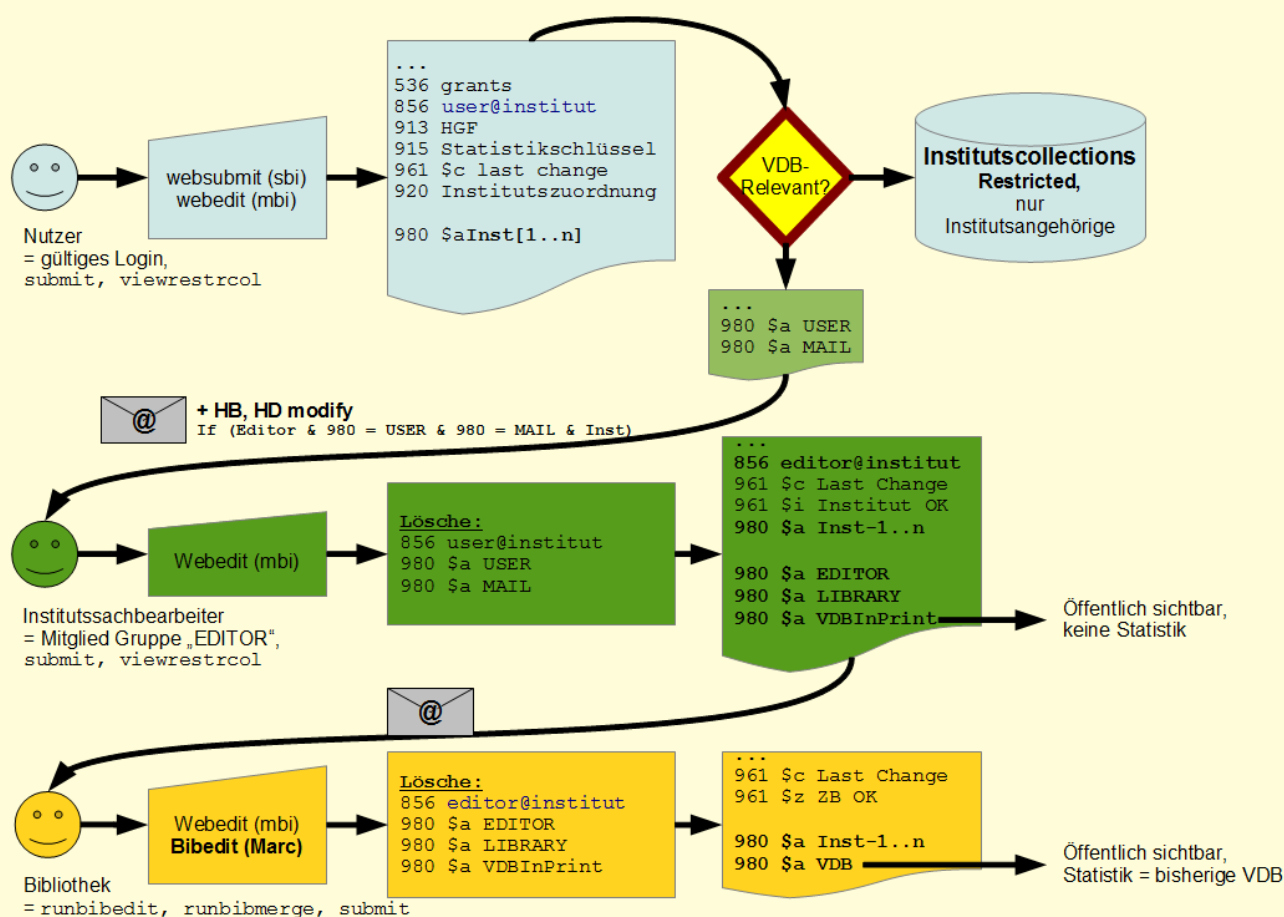
Forschungsinfrastruktur.

Teil II: Implementation



The flexibility and performance of Invenio make it a comprehensive solution for [...] document repositories [...] to **large sizes** (several millions of records).

- > Basis: [Marc21](#) Bibliographic ☺
- > **JOIN²**: Workflow ([Kontrollschritte](#))



The flexibility and performance of Invenio make it a comprehensive solution for [...] document repositories [...] to **large sizes** (several millions of records).

- > Basis: Marc Bibliographic
- > **JOIN²**: Workflow (Kontrollschritte)
- > **JOIN²**: Marc-**Authority** zur Normierung
 - bibliographische Angaben (z. B. Journals)
 - Institutszuordnungen (Vorgänger, Nachfolger, Tops, virtuelle...)
 - Autorenidentifikation (auch ATLAS...)
 - Projektzuordnungen (HGF, EU, DFG, BMBF...)
 - Statistikschlüssel (z. B. für Evaluationen)
 - Vokabular (z. B. Lizenzen, Schlagworte...)



Normsätze: Verwendung

- > Vorschlagsfunktionen (Nutzer: websubmit)
 - Autoren
 - Institute
 - Projekte
 - ...



The screenshot shows the 'Modify Record' form in the PUBDB system. The form includes fields for 'Group(s) involved', 'Beamline/Experiment/Facility', 'Grant name / Proposal No.', 'Author(s) / Contributor(s)', 'Title', 'Language', 'Conference name', 'Acronym', 'City', 'Country', 'Begin', and 'End'. A blue overlay box titled 'websubmit^a' lists the following features:

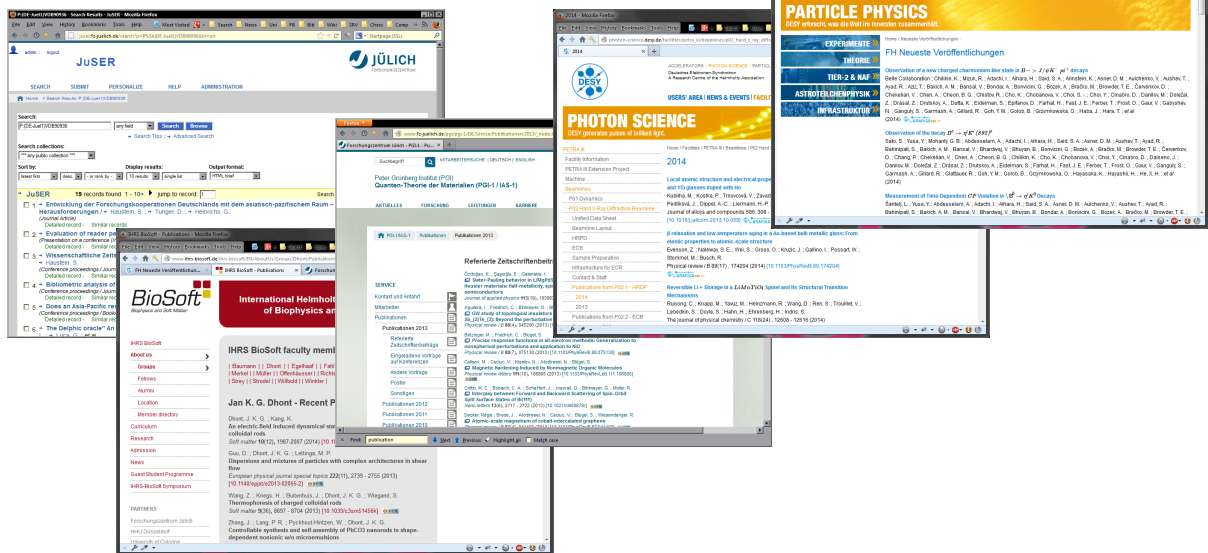
- > (fast) alle Eingaben normiert (Institute, Grants, Personen, ...)
- > Vorschlagsfunktionen^{b,c} (Ajax: Authorities + Suchen + jQuery)
- > semiautomatische Autorenindividualisierung^b
- > Importfunktion^b (doi, pmid, arXiv, Inspire...)
- > Backendfunktionen^d

Footnotes for the overlay: ^aDESY, ^bJülich, ^cAachen, ^dDarmstadt

Normsätze: Verwendung

- > Vorschlagsfunktionen (Nutzer: websubmit)
 - Autoren
 - Institute
 - Projekte
 - ...
- > Webausgaben (z. B. Institutsseiten, Projektseiten, Publikationslisten...)

Webausgaben – CMS



Auch eine Publikationsliste ist eine Auswertung

(aid, cid, fai pof, pub, sid, typ ...)



Normsätze: Verwendung

- > Vorschlagsfunktionen (Nutzer: websubmit)
 - Autoren
 - Institute
 - Projekte
 - ...
- > Webausgaben (z. B. Institutsseiten, Projektseiten, Publikationslisten...)
- > Statistiken (s. a. K. Große)
- > ...

Vorteile

Flexibel, einfach erweiterbar



- > alle Publikationen eines Journals
- > alle Publikationen von Journals in einer Datenbank
- > alle Publikationen eines bestimmten Feldes

Fragestellungen (Beispiele)

- > “Welche Artikel sind in WoS oder Scopus?” — “oder Pubmed?”
- > “Welche unserer OpenAccess Artikel sind in WoS-Journals?”
- > “Was publizierten wir in Nature (ohne Baby-Natures)?”
- > “Welche OA-Journals haben einen Impact Factor?”
- > ...



Normsätze, z. B. Institute

```
001__ 98676
0247_ $2I:(DE-Juel1) $aI:(DE-Juel1)VDB793
035__ $aI:(DE-Juel1)VDB793
1101_ $aAgrosphäre $y-31.10.2010
4101_ $aICG-4 $wd
5101_ $OI:(DE-Juel1)VDB1109 $2I:(DE-Juel1) $aICG $wt
5101_ $OI:(DE-Juel1)VDB50 $2I:(DE-Juel1) $aAgrosphäre $wa
5101_ $OI:(DE-Juel1)IBG-3-20101118 $2I:(DE-Juel1)
      $aAgrosphäre $wb
970__ $aI:(DE-Juel1)VDB793
980__ $aI
980__ $aAUTHORITY
```



Record #98676 - JuSER - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Search News Um FR Bib Wiki Inspire Invenio Dev SRV Chess

https://juser.fz-juelich.de/record/98676

Record #98676 - JuSER

a.wagner :: logout

JuSER

JÜLICH
FORSCHUNGSZENTRUM

SEARCH SUBMIT PERSONALIZE HELP ADMINISTRATION

Home > Authorities > Institutes > Record #98676

Information Discussion (0) Files Plots

Agrosphäre (ICG-4)

ID I:(DE-Juel1)VDB793

→ Agrosphäre → ICG → Agrosphäre

RECENT PUBLICATIONS

→ All known publications

Journal Article

→ Hinnell, A. C. ; → Ferre, T. P. A. ; → Vrugt, J. A. ; → Huisman, J. A. ; → Moysey, S. ; → Rings, J. ; → Kowalsky, M. B.
→ **Improved extraction of hydrologic information from geophysical data through coupled hydrogeophysical inversion**
Water resources research **46**, W00D30 (2010) [10.1029/2008WR007060] 

Journal Article

→ Dutta, S. ; → Mallick, M. ; → Mathews, R. P. ; → Mann, U. ; → Greenwood, P.
→ **Chemical composition and palaeobotanical origin of Miocene resins from Kerala-Konkan coast, Western India**

Normsätze, z. B. Journals

```
0247_ $2ISSN$a0163-1829
0247_ $2ISSN$a1095-3795
...
035__ $aPERI:(DE-600)1473011-x
041__ $aeng
082__ $a530
...
245__ $aPhysical review / B ...
8564_ $uhttp://ezb.uni-regensburg.de/1473011$yEZB
8564_ $uhttp://prb.aps.org$yVerlag
85672 $2RSS$uhttp://feeds.aps.org/rss/accepted/prb.xml...
...
915__ $0StatID:(DE-HGF)0100$2StatID$aJCR
915__ $0StatID:(DE-HGF)0200$2StatID$aDBCoverage$bSCOPUS
915__ $0StatID:(DE-HGF)0300$2StatID$aDBCoverage$bMedline
915__ $0StatID:(DE-HGF)0110$2StatID$aWoS$bScience Citation Index
915__ ...
```

- > Repositoriensystem $\Rightarrow$ OAI-Server
- > OAI-Sets ([authority:*](#)):
 - Periodika
 - Projekte ([POF](#), [EU](#), [+X](#))
 - Experimente
 - Statistikschlüssel
 - Dokumenttypen
 - Vokabular
- > Exponiere MarcXML
- > Zentraler Server als Master ([derzeit JuSER](#))
- > Harvesting der relevanten Sets via OAI-PMH
- > Code: git



Projektgruppe

- | | |
|-------------------------------------|----------------------------------|
| > Dagmar Sitek ^a | > Cornelia Plott ^f |
| > Gudrun Friedburg ^a | > Christoph Holzke ^f |
| > Martin Köhler ^b | > Jürgen Neuhaus ^g |
| > Zaven Akopov ^{b,c,1} | > Connie Hesse ^g |
| > Tomasz Pazera ^{b,1} | > Björn Pedersen ^g |
| > Robert Thiele ^b | > Jörg Pulz ^g |
| > Alexander Wagner ^{f,b} | > Ulrike Eich ^h |
| > Katrin Große ^d | > Louai Barake ^h |
| > Stefan Hesselbach ^e | > Abdoulaye Diallo ^h |
| > Bernhard Mittermaier ^f | > Roland Rappmann ^h |
| > Anna Fründ ^f | > Dominik Schmitz ^h |
| > Heike Lexis ^f | > Edmund Wollgarten ^h |

^a DKFZ Zentralbibliothek, ^b DESY Library and Documentation, ^c Project Inspire, ^d GSI Library, ^e GSI Core IT,
^f Forschungszentrum Jülich, Zentralbibliothek, ^g MLZ Garching, ^h RWTH Aachen, Hochschulbibliothek
¹(retired)

With a little help from our friends

CERN IT, CERN Library, Project Inspire





Alexander Wagner
Deutsches Elektronen-Synchrotron
Central Library

Tel.: +49-40-8998-1758
alexander.wagner@desy.de

<http://library.desy.de>

