Mitglied der Helmholtz-Gemeinschaft

JÜLICH
FORSCHUNGSZENTRUM

Invenio@HGF – status and perspectives

[sic!] Jülich – 2nd Invenio User Group Workshop

Jülich, 18. November 2013 | Alexander Wagner, for the Collaboration

Overview

- Partner
- Initial TODO
- Accomplishments
- Lessons learned
- Project group

BY NC SA

Project Partners

	Deutsches Elektronensynchrotron, Zentralbibliothek	$\approx 2000 + 3000$
	Forschungszentrum Jülich, Zentralbibliothek	$\approx 5000 + 1000$
	GSI Helmholtzzentrum für Schwerionenforschung, Biblitohek + Kern-IT	≈ 1050
	Maier-Leibniz-Zentrum, Garching	≈ 300
	RWTH Aachen, Hochschulbibliothek	≈ 9000
	Museum Zitadelle Jülich	
	Institut für Experimentelle Kernphysik, Karlsruhe	

CC BY NC SA

Project Partners

	Deutsches Elektronensynchrotron, Zentralbibliothek	$\approx 2000 + 3000$
	Forschungszentrum Jülich, Zentralbibliothek	$\approx 5000 + 1000$
	GSI Helmholtzzentrum für Schwerionenforschung, Biblitohek + Kern-IT	≈ 1050
	Maier-Leibniz-Zentrum, Garching	≈ 300
	RWTH Aachen, Hochschulbibliothek	≈ 9000
	Museum Zitadelle Jülich	
	Institut für Experimentelle Kernphysik, Karlsruhe	

Open for new Partners!

CC BY NC SA

Project Partners

	Deutsches Elektronensynchrotron, Zentralbibliothek	$\approx 2000 + 3000$
	Forschungszentrum Jülich, Zentralbibliothek	$\approx 5000 + 1000$
	GSI Helmholtzzentrum für Schwerionenforschung, Biblitohek + Kern-IT	≈ 1050
	Maier-Leibniz-Zentrum, Garching	≈ 300
	RWTH Aachen, Hochschulbibliothek	≈ 9000
	Museum Zitadelle Jülich	
	Institut für Experimentelle Kernphysik, Karlsruhe	

Open for new Partners!

Serving now $\approx$ **17.000** people (+ visitors)

($\approx$ 260.000 documents + 67.000 Authorities)

CC BY NC SA

Starting out

Goal

Replace existing systems, at GSI build up from scratch.
User-centric design (users aka scientists)

CC BY NC SA

Starting out

Goal

Replace existing systems, at GSI build up from scratch.
User-centric design (users aka scientists)

1. “Learn Invenio” (thanks to CERN ☺)

CC BY NC SA

Starting out

Goal

Replace existing systems, at GSI build up from scratch.
User-centric design (users aka scientists)

1. “Learn Invenio” (thanks to CERN ☺)
2. Define wording. . . (different insitutions!)

CC BY NC SA

Starting out

Goal

Replace existing systems, at GSI build up from scratch.
User-centric design (users aka scientists)

1. “Learn Invenio” (thanks to CERN ☺)
2. Define wording. . . (different insitutions!)
3. Build infrastructure: git and friends

CC BY NC SA

Starting out

Goal

Replace existing systems, at GSI build up from scratch.
User-centric design (users aka scientists)

1. “Learn Invenio” (thanks to CERN ☺)
2. Define wording. . . (different insitutions!)
3. Build infrastructure: git and friends
4. Build more infrastructure: authorities and friends

CC BY NC SA

Starting out

Goal

Replace existing systems, at GSI build up from scratch.
User-centric design (users aka scientists)

1. “Learn Invenio” (thanks to CERN ☺)
2. Define wording. . . (different insitutions!)
3. Build infrastructure: git and friends
4. Build more infrastructure: authorities and friends
5. Build a **deployment scheme**: InstallInvenio and friends

CC BY NC SA

Starting out

Goal

Replace existing systems, at GSI build up from scratch.
User-centric design (users aka scientists)

1. "Learn Invenio" (thanks to CERN ☺)
2. Define wording. . . (different insitutions!)
3. Build infrastructure: git and friends
4. Build more infrastructure: authorities and friends
5. Build a **deployment scheme**: InstallInvenio and friends

We need to roll out 10+ instances

CC BY NC SA

Starting out

Goal

Replace existing systems, at GSI build up from scratch.
User-centric design (users aka scientists)

1. "Learn Invenio" (thanks to CERN ☺)
2. Define wording. . . (different insitutions!)
3. Build infrastructure: git and friends
4. Build more infrastructure: authorities and friends
5. Build a **deployment scheme**: InstallInvenio and friends

We need to roll out 10+ instances
with different data sets

CC BY NC SA

Starting out

Goal

Replace existing systems, at GSI build up from scratch.
User-centric design (users aka scientists)

1. "Learn Invenio" (thanks to CERN ☺)
2. Define wording. . . (different insitutions!)
3. Build infrastructure: git and friends
4. Build more infrastructure: authorities and friends
5. Build a **deployment scheme**: InstallInvenio and friends

We need to roll out 10+ instances
with different data sets and keep them consistent on code level

CC BY NC SA

Initial ToDo

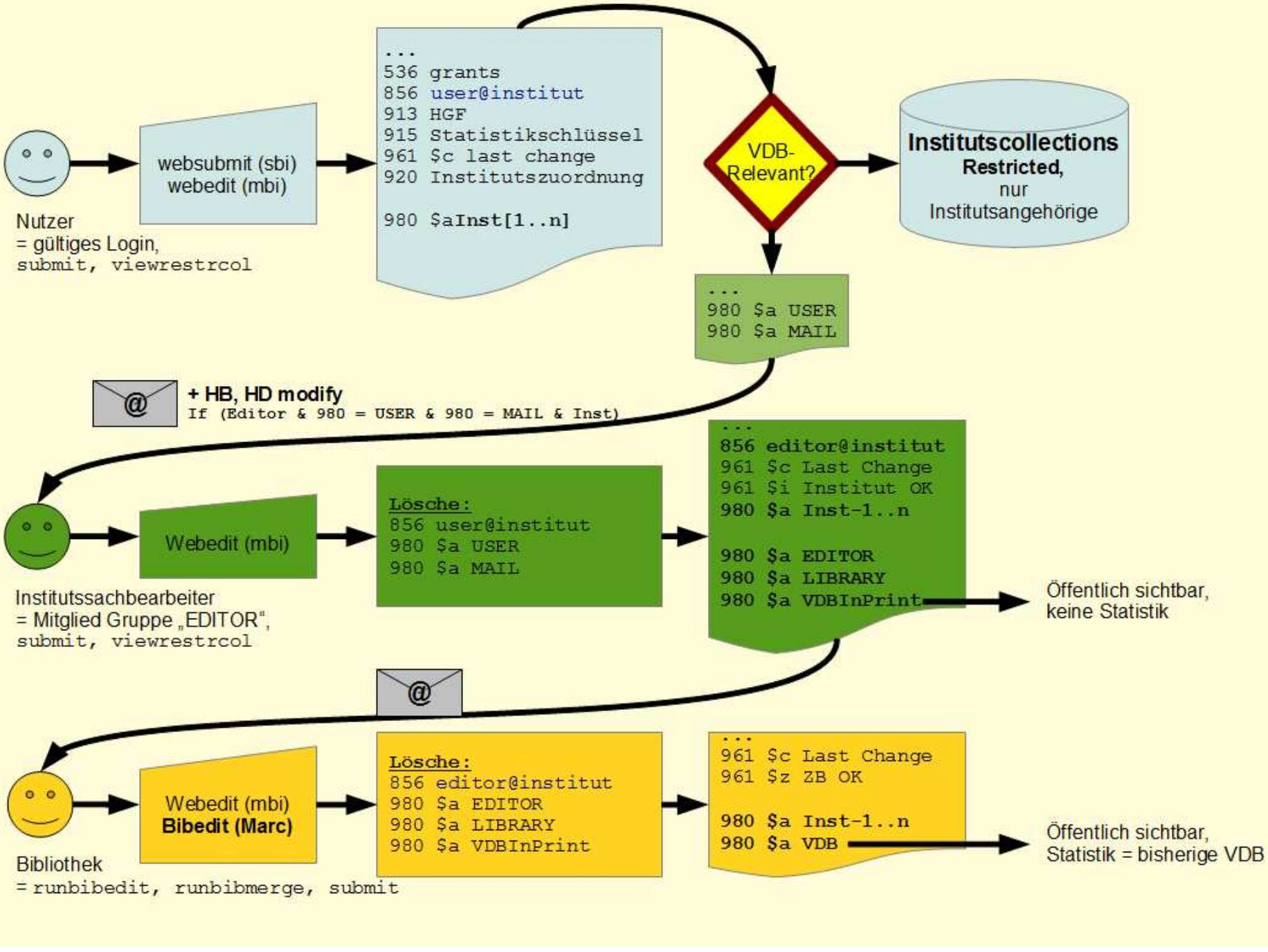
- Design the system around web based literature management

CC BY NC SA

Initial ToDo

- Design the system around web based literature management
- Design a document workflow

CC BY NC SA


Nutzer
= gültiges Login,
submit, viewrestrcol
websubmit (sbi)
webedit (mbi)
...
536 grants
856 user@institut
913 HGF
915 Statistikschlüssel
961 $c last change
920 Institutszuordnung
980 $aInst[1..n]
VDB-Relevant?
Institutscollections
Restricted,
nur
Institutsangehörige
...
980 $a USER
980 $a MAIL
+ HB, HD modify
If (Editor & 980 = USER & 980 = MAIL & Inst)
Institutssachbearbeiter
= Mitglied Gruppe „EDITOR“,
submit, viewrestrcol
Webedit (mbi)
Lösche:
856 user@institut
980 $a USER
980 $a MAIL
...
856 editor@institut
961 $c Last Change
961 $i Institut OK
980 $a Inst-1..n
980 $a EDITOR
980 $a LIBRARY
980 $a VDBInPrint
Öffentlich sichtbar,
keine Statistik
Bibliothek
= runbibedit, runbibmerge, submit
Webedit (mbi)
Bibedit (Marc)
Lösche:
856 editor@institut
980 $a EDITOR
980 $a LIBRARY
980 $a VDBInPrint
...
961 $c Last Change
961 $z ZB OK
980 $a Inst-1..n
980 $a VDB
Öffentlich sichtbar,
Statistik = bisherige VDB

Initial ToDo

- Design the system around web based literature management
- Design document workflow (3 steps with privilege escalation)
- Design easy ingestion workflow (websubmit, imports, author disambiguation)

CC BY NC SA

Initial ToDo

- Design the system around web based literature management
- Design document workflow (3 steps with privilege escalation)
- Design easy ingestion workflow (websubmit, imports, author disambiguation)

Finally we wrote some code. . .

CC BY NC SA

Initial ToDo

- Design the system around web based literature management
- Design document workflow (3 steps with privilege escalation)
- Design easy ingestion workflow (websubmit, imports, author disambiguation)

Finally we wrote some code. . .

Every unwritten line is a good line

CC BY NC SA

Initial ToDo

- Design the system around web based literature management
- Design document workflow (3 steps with privilege escalation)
- Design easy ingestion workflow (websubmit, imports, author disambiguation)

Finally we wrote some code. . .

Every unwritten line is a good line, still: $\approx$ 55.000 lines

CC BY NC SA

Initial ToDo

- Design the system around web based literature management
- Design document workflow (3 steps with privilege escalation)
- Design easy ingestion workflow (websubmit, imports, author disambiguation)

Finally we wrote some code. . .

Every unwritten line is a good line, still: $\approx$ 55.000 lines

- Migrate old data (various, proprietary sources)
- Train the inputters and users (secretaries, scientists, librarians)
- Hook up with content management system(s) (visibility!)

CC BY NC SA

Referierte Zeitschriftenbeiträge 2013

SERVICE
Kontakt und Anfahrt
Mitarbeiter
Publikationen
Publikationen 2013
Referierte Zeitschriftenbeiträge
Eingeladene Vorträge auf Konferenzen
Andere Vorträge
Poster
Sonstiges
Publikationen 2012
Publikationen 2011
Publikationen 2010

Özdoğan, K. ; Şaşıoğlu, E. ; Galanakis, I.
Slater-Pauling behavior in LiMgPdSn-type multifunctional quaternary Heusler materials: Half-metallicity, spin-gapless and magnetic semiconductors
Journal of applied physics **113**(19), 193903 - (2013) [10.1063/1.4805063]

Aguilera, I. ; Friedrich, C. ; Bihlmayer, G. ; Blügel, S.
GW study of topological insulators Bi_{2}Se_{3}, Bi_{2}Te_{3}, and Sb_{2}Te_{3}: Beyond the perturbative one-shot approach
Physical review / B **88**(4), 045206 (2013) [10.1103/PhysRevB.88.045206]

Betzinger, M. ; Friedrich, C. ; Blügel, S.
Precise response functions in all-electron methods: Generalization to nonspherical perturbations and application to NiO
Physical review / B **88**(7), 075130 (2013) [10.1103/PhysRevB.88.075130]

Callsen, M. ; Caciuc, V. ; Kiselev, N. ; Atodiresei, N. ; Blügel, S.
Magnetic Hardening Induced by Nonmagnetic Organic Molecules
Physical review letters **111**(10), 106805 (2013) [10.1103/PhysRevLett.111.106805]

Cottin, M. C. ; Bobisch, C. A. ; Schaffert, J. ; Jnawali, G. ; Bihlmayer, G. ; Möller, R.
Interplay between Forward and Backward Scattering of Spin–Orbit Split Surface States of Bi(111)
Nano letters **13**(6), 2717 - 2722 (2013) [10.1021/nl400878r]

Decker, Régis ; Brede, J. ; Atodiresei, N. ; Caciuc, V. ; Blügel, S. ; Wiesendanger, R.
Atomic-scale magnetism of cobalt-intercalated graphene
Physical review / B **87**(4), 041403 (2013) [10.1103/PhysRevB.87.041403]

Initial ToDo

- Design a document workflow (3 steps with privilege escalation)
- Establish easy ingestion workflow (websubmit, imports, author disambiguation)

Finally we wrote some code. . .

Every unwritten line is a good line, still: $\approx$ 55.000 lines

- Migrate old data (various, proprietary sources)
- Train the inputters and users (secretaries, scientists, librarians)
- Hook up with content management system(s) (visibility!)
- Derive necessary reporting (details are way beyond this intro)

CC BY NC SA

Initial ToDo

- Design a document workflow (3 steps with privilege escalation)
- Establish easy ingestion workflow (websubmit, imports, author disambiguation)

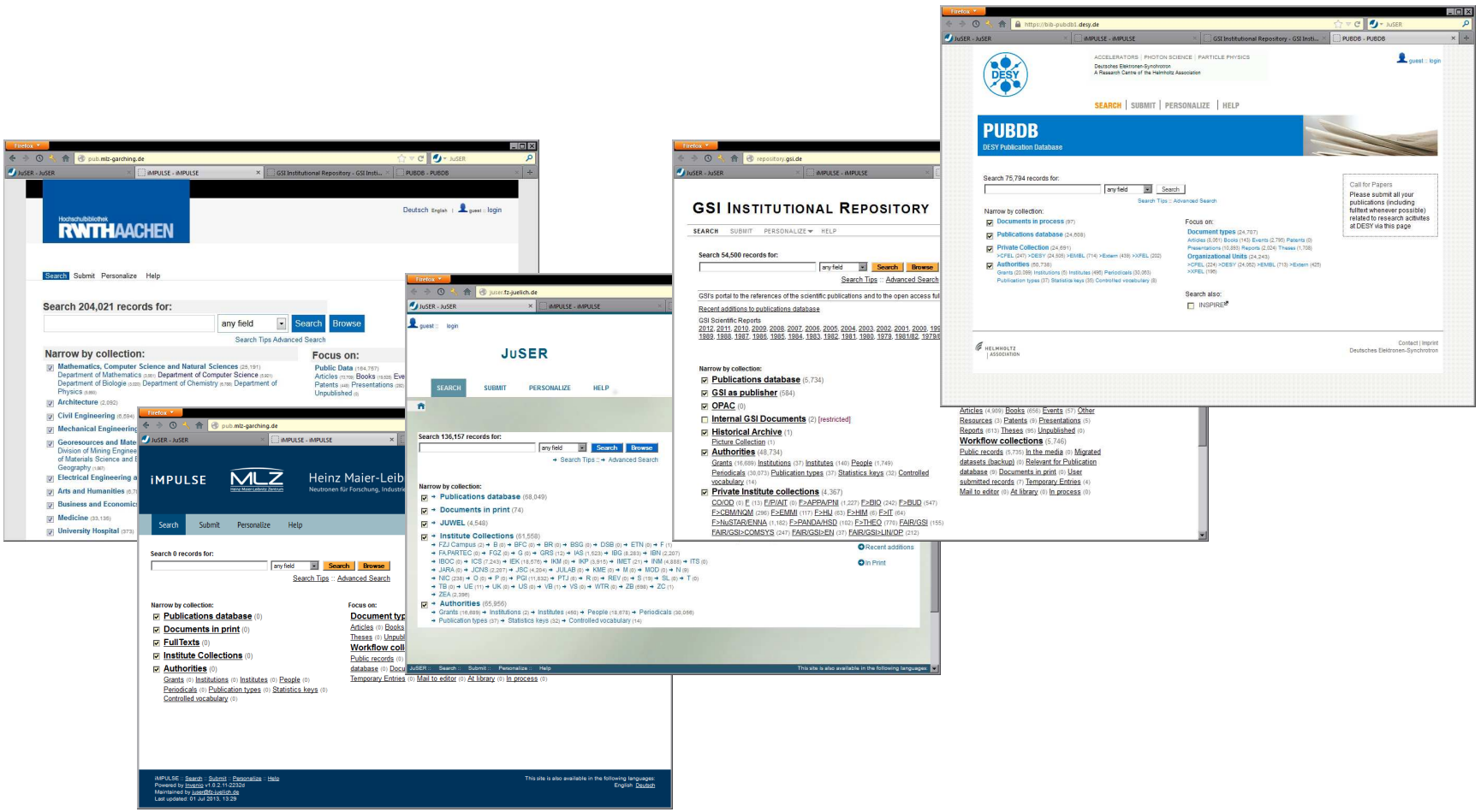
Finally we wrote some code. . .

Every unwritten line is a good line, still: $\approx$ 55.000 lines

- Migrate old data (various, proprietary sources)
- Train the inputters and users (secretaries, scientists, librarians)
- Hook up with content management system(s) (visibility!)
- Derive necessary reporting (details are way beyond this intro)
- Get it up and running (First Light: 11/19/2012)

CC BY NC SA

Accomplishments and status



CC BY NC SA

Accomplishments and status

- All partners have running systems (roll out works)

CC BY NC SA

Accomplishments and status

- All partners have running systems (roll out works)
- Almost all partners are online

CC BY NC SA

Accomplishments and status

- All partners have running systems (roll out works)
- Almost all partners are online
- Rich websubmit (including **repeatable field** handling)
- Importer routines (doi, pmid, arXiv, inspire, ISBN, own recs,. . . in **websubmit**)

CC BY NC SA

Accomplishments and status

- All partners have running systems (roll out works)
- Almost all partners are online
- Rich websubmit (including **repeatable field** handling)
- Importer routines (doi, pmid, arXiv, inspire, ISBN, own recs,. . . in **websubmit**)
- Authorities

CC BY NC SA

Accomplishments and status

- All partners have running systems (roll out works)
- Almost all partners are online
- Rich websubmit (including **repeatable field** handling)
- Importer routines (doi, pmid, arXiv, inspire, ISBN, own recs,. . . in **websubmit**)
- Authorities
  - Generate (≈ 67.000 recs)
  - Use (e. g. JSON returns, statistics. . .)
  - Share (MarcXML OAI-PMH)

CC BY NC SA

Accomplishments and status

- All partners have running systems (roll out works)
- Almost all partners are online
- Rich websubmit (including **repeatable field** handling)
- Importer routines (doi, pmid, arXiv, inspire, ISBN, own recs,. . . in **websubmit**)
- Authorities
  - Generate (≈ 67.000 recs)
  - Use (e. g. JSON returns, statistics. . .)
  - Share (MarcXML OAI-PMH)
- Implement

CC BY NC SA

Accomplishments and status

- All partners have running systems (roll out works)
- Almost all partners are online
- Rich websubmit (including **repeatable field** handling)
- Importer routines (doi, pmid, arXiv, inspire, ISBN, own recs,... in **websubmit**)
- Authorities
  - Generate (≈ 67.000 recs)
  - Use (e. g. JSON returns, statistics...)
  - Share (MarcXML OAI-PMH)
- Implement
  - Author identification (ORCiD ready!)
  - Output formats (JSON, BibTEX, EndNote...)
  - Reporting (publication statistics)
  - Delivery to content management systems

CC BY NC SA

Tools used

- Workflow

CC BY NC SA

Tools used

- Workflow
  - Webbaskets (e. g. revision lists)
  - Alerts (e. g. revision lists)
  - Collections (e. g. private for institutes)
  - Webmessage (e. g. correction requests)

CC BY NC SA

Tools used

- Workflow
  - Webbaskets (e. g. revision lists)
  - Alerts (e. g. revision lists)
  - Collections (e. g. private for institutes)
  - Webmessage (e. g. correction requests)
- Authority records (almost everywhere)

CC BY NC SA

Tools used

- Workflow
  - Webbaskets (e. g. revision lists)
  - Alerts (e. g. revision lists)
  - Collections (e. g. private for institutes)
  - Webmessage (e. g. correction requests)
- Authority records (almost everywhere)
- OAI-PMH (authority exchange)

CC BY NC SA

Tools used

- Workflow
  - Webbaskets (e. g. revision lists)
  - Alerts (e. g. revision lists)
  - Collections (e. g. private for institutes)
  - Webmessage (e. g. correction requests)
- Authority records (almost everywhere)
- OAI-PMH (authority exchange)
- High-level API (setup: e. g. collections, roles, groups, baskets. . . ; **no db-dump** sharing)

CC BY NC SA

Tools used

- Workflow
  - Webbaskets (e. g. revision lists)
  - Alerts (e. g. revision lists)
  - Collections (e. g. private for institutes)
  - Webmessage (e. g. correction requests)
- Authority records (almost everywhere)
- OAI-PMH (authority exchange)
- High-level API (setup: e. g. collections, roles, groups, baskets. . . ; **no db-dump** sharing)
- jQuery/jQueryUI (websubmit)

CC BY NC SA

Tools used

- Workflow
  - Webbaskets (e. g. revision lists)
  - Alerts (e. g. revision lists)
  - Collections (e. g. private for institutes)
  - Webmessage (e. g. correction requests)
- Authority records (almost everywhere)
- OAI-PMH (authority exchange)
- High-level API (setup: e. g. collections, roles, groups, baskets. . . ; **no db-dump** sharing)
- jQuery/jQueryUI (websubmit)
- intbitsets (e. g. statistics)

CC BY NC SA

Lessons learned / Next steps

- CERN is way to fast to keep up with

CC BY NC SA

Lessons learned / Next steps

- CERN is way to fast to keep up with
- Never use Dublin Core again (complex migration, to few data fields. . .)

CC BY NC SA

Lessons learned / Next steps

- CERN is way to fast to keep up with
- Never use Dublin Core again (complex migration, to few data fields. . .)
- All libraries are the same ☺

CC BY NC SA

Lessons learned / Next steps

- CERN is way to fast to keep up with
- Never use Dublin Core again (complex migration, to few data fields...)
- All libraries are the same ☺

- Upgrade to 1.2: get OAI-Server fixed!

CC BY NC SA

Lessons learned / Next steps

- CERN is way to fast to keep up with
- Never use Dublin Core again (complex migration, to few data fields...)
- All libraries are the same ☺

- Upgrade to 1.2: get OAI-Server fixed!

However...

In our use case switching of the base system is non-trivial

(Remember: 10+ instances...)

CC BY NC SA

Lessons learned / Next steps

- CERN is way to fast to keep up with
- Never use Dublin Core again (complex migration, to few data fields...)
- All libraries are the same ☺

- Upgrade to 1.2: get OAI-Server fixed!

However...

In our use case switching of the base system is non-trivial

(Remember: 10+ instances...)

- Open up for new partners
- Clean up our code and give it back

CC BY NC SA

Contributors

- *Martin Köhler*[a]
- Zaven Akopov[a,b]
- *Tomasz Pazera*[a]
- Katrin Große[c]
- *Stefan Hesselbach*[d]
- Bernhard Mittermaier[e]
- Anna Fründ[e]
- Heike Lexis[e]
- *Cornelia Plott*[e]
- Christoph Holzke[e]
- *Alexander Wagner*[e]
- Jürgen Neuhaus[f]
- Connie Hesse[f]
- *Björn Pedersen*[f]
- Ulrike Eich[g]
- *Louai Barake*[g]
- *Abdoulaye Diallo*[g]
- Roland Rappmann[g]
- Dominik Schmitz[g]
- Edmund Wollgarten[g]

[a] DESY Library and Documentation; [b] Project Inspire; [c] GSI Library; [d] GSI Core IT;

[e] Forschungszentrum Jülich, Zentralbibliothek; [f] MLZ, Garching; [g] RWTH Aachen, Hochschulbibliothek

CC BY NC SA

Further “reading”

- Invenio @ HGF - Technical background

  Talk at Invenio Developer Forum

- Collaborative tools for an institutional repository

  Talk at Helmholtz OA Workshop

- JuSER – Publications Database

  Introductory course at Jülich

- JuSER - Autorenhandling

  Talk at HGF-ORCiD Meeting, Berlin (in german)

CC BY NC SA

Thanks!

Alexander Wagner
Zentralbibliothek

Scientific Services /
Scientific Publishing

Tel.: +49-2461-61-1586
a.wagner@fz-juelich.de

This document is available as FZJ-2013-05410

Typeset by pdfLaTeX

CC BY NC SA