

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

Project Partners



Deutsches Elektronensynchrotron, Zentralbibliothek

≈ 2000 + 3000



Forschungszentrum Jülich, Zentralbibliothek

≈ 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

Project Partners



Deutsches Elektronensynchrotron, Zentralbibliothek

≈ 2000 + 3000



Forschungszentrum Jülich, Zentralbibliothek

≈ 5000 + 1000



GSI Helmholtzzentrum für Schwerionenforschung, Bibliothek + Kern-IT

≈ 1050



Maier-Leibniz-Zentrum, Garching

≈ 300



RWTH Aachen, Hochschulbibliothek

≈ 9000

Museum Zitadelle Jülich

Institut für Experimentelle Kernphysik, Karlsruhe

Open for new Partners!

Project Partners



Deutsches Elektronensynchrotron, Zentralbibliothek

≈ 2000 + 3000



Forschungszentrum Jülich, Zentralbibliothek

≈ 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 ≈ **17.000** people (+ visitors)

(≈ 260.000 documents + 67.000 Authorities)

Starting out

Goal

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

Starting out

Goal

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

- 1 “Learn Invenio” (thanks to CERN ☺)

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!)

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

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

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

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

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

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

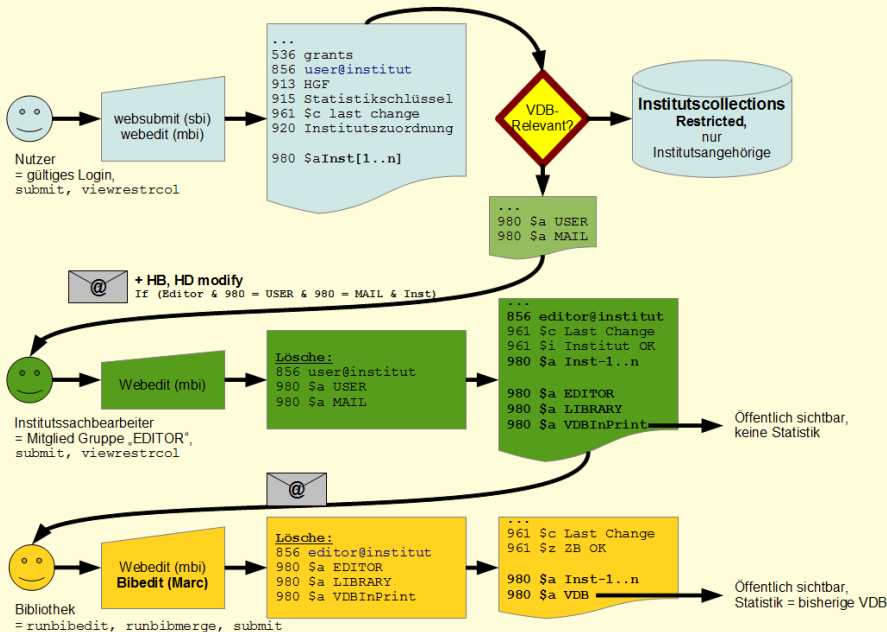
We need to roll out **10+** instances
with **different** data sets and **keep them consistent** on code level

Initial ToDo

- Design the system around web based literature management

Initial ToDo

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



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)

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. . .

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

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

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!)

Firefox

www.fz-juelich.de/pgi/pgi-1/DE/Service/Publikationen/2013/_node.html

☆

▼

↺

↻

➤

JuSER

Forschungszentrum Jülich - PGI-1 - Pu...

+

Suchbegriff

Q

MITARBEITERSUCHE | DEUTSCH | ENGLISH

Institutswegweiser

JÜLICH
FORSCHUNGSZENTRUM

AKTUELLES | FORSCHUNG | LEISTUNGEN | KARRIERE | ÜBER UNS

PGI-1/IAS-1

Publikationen

Publikationen 2013

Referierte Zeitschriftenbeiträge 2013

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 Bi₂Se₃, Bi₂Te₃, and Sb₂Te₃: 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] 

Callens, M. ; Caciuc, V. ; Kiselev, N. ; Atodiresel, 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. ; Atodiresel, 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] 

Kontakt und Anfahrt

Mitarbeiter

Publikationen

Publikationen 2013

Referierte Zeitschriftenbeiträge

Eingeladene Vorträge auf Konferenzen

Andere Vorträge

Poster

Sonstiges

Publikationen 2012

Publikationen 2011

Publikationen 2010

Find: publication

Next

Previous

Highlight all

Match case

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: ≈ 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)

Initial ToDo

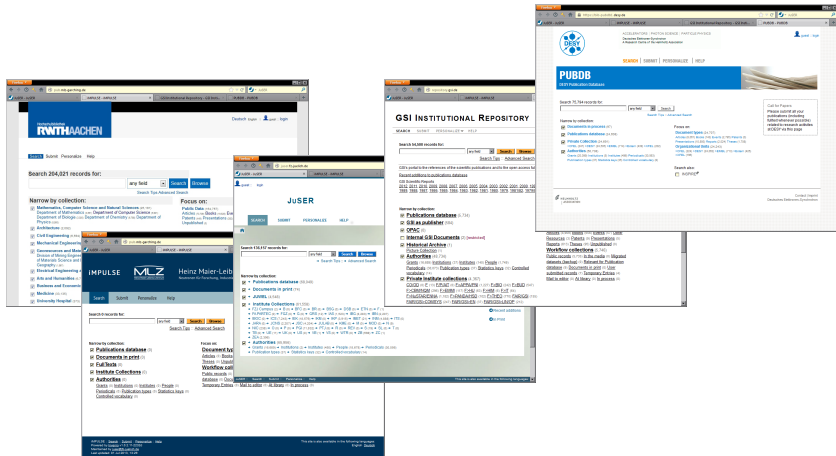
- Design a document workflow (3 steps with privilege escalation)
- Establish easy ingestion workflow (websubmit, imports, author disambiguation)

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)

Accomplishments and status



Accomplishments and status

- All partners have running systems (roll out works)

Accomplishments and status

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

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**)

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

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)

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

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**

Tools used

- Workflow

Tools used

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

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)

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)

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)

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)

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)

Lessons learned / Next steps

- CERN is way to fast to keep up with

Lessons learned / Next steps

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

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 ☺

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!

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...)

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

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

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)

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 pdfL^AT_EX