

# JOIN<sup>2</sup> .

## Status Update

Alexander Wagner, for *JOIN<sup>2</sup>*

**CERN**

**Geneva, 12.09.2015**



> Who is **JOIN<sup>2</sup>**

# Overview



- > Who is **JOIN<sup>2</sup>**
- > Recent achievements





Deutsches Elektronensynchrotron, Zentralbibliothek



≈ 2000 + 3000

Forschungszentrum Jülich, Zentralbibliothek



≈ 5500 + 1000

GSI Helmholtzzentrum für Schwerionenforschung,  
Bibliothek + Base-IT



≈ 1350 + 1200

Deutsches Krebsforschungszentrum, Bibliothek

≈ 3000

Maier-Leibniz-Zentrum, Garching

≈ 300 + 1000

RWTH Aachen, Universitätsbibliothek



≈ 9000

Museum Zitadelle Jülich





Deutsches Elektronensynchrotron, Zentralbibliothek



≈ 2000 + 3000



Forschungszentrum Jülich, Zentralbibliothek



≈ 5500 + 1000



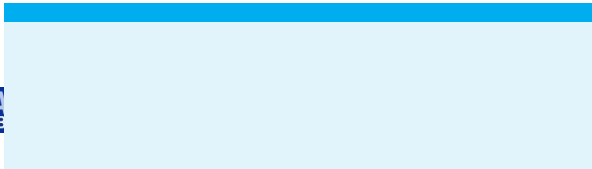
GSI Helmholtzzentrum für Schwerionenforschung,  
Bibliothek + Base-IT



≈ 1350 + 1200

dkfz.

GERMAN  
CANCER RESEARCH CENTER  
HELMHOLTZ-ZENTRUM  
FÜR STRAHLENTHERAPIE



≈ 3000



≈ 300 + 1000



≈ 9000





Deutsches Elektronensynchrotron, Zentralbibliothek



≈ 2000 + 3000



Forschungszentrum Jülich, Zentralbibliothek



≈ 5500 + 1000



GSI Helmholtzzentrum für Schwerionenforschung,  
Bibliothek + Base-IT



≈ 1350 + 1200

dkfz.

GERMAN  
CANCER RESEARCH CENTER  
HELMHOLTZ-ZENTRUM  
FÜR SCHRWERIONENFORSCHUNG

≈ 3000



Online for ≈ **20.000** staff members

≈ 300 + 1000



+ ≈ 6000 visiting scientists

≈ 9000

(≈ 310.000 records, ≈ 30.000 OpenAccess + 150.000 authorities)



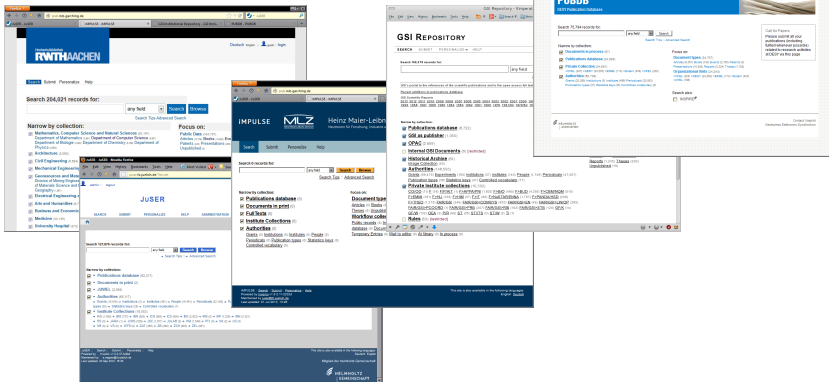
# Installations

The image displays five overlapping screenshots of academic search engines:

- RWTH Aachen:** Shows a search interface with a search bar and a list of results under 'Narrow by collection:'.
- IMPULSE:** Features a search bar and a list of results under 'Narrow by collection:'.
- JUSER:** Shows a search bar and a list of results under 'Narrow by collection:'.
- GSI Repository:** Displays a search bar and a list of results under 'Narrow by collection:'.
- PUBDB:** Shows a search bar and a list of results under 'Narrow by collection:'.



# Installations



Individual instances, operated by the partners





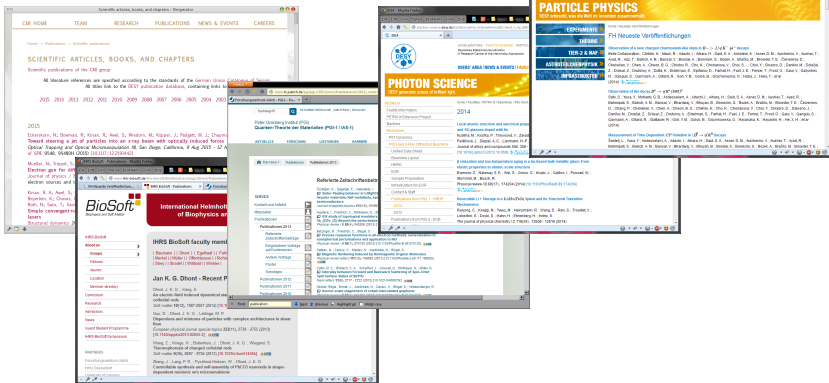
# Weboutput — CMS

The collage displays several web pages from the DESY CMS:

- Top Left:** A page titled "SCIENTIFIC ARTICLES, BOOKS, AND CHAPTERS" listing publications with abstracts and dates.
- Top Right:** A "PARTICLE PHYSICS" banner with a "NEWS" section listing recent articles and events.
- Middle Left:** A "BioSoft" page for the "International Helmholtz of Biophysics" with a navigation menu.
- Middle Right:** A "PHOTON SCIENCE" banner for the year 2014, listing various events and conferences.
- Bottom:** A "Referierte Zeitschriften" page listing scientific journals and articles.



# Weboutput — CMS



Integrate: think database + CMS

Publication lists need data as detailed as for evaluations...

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



# Just another Invenio iNstallation to join



# Just another Invenio iNstallation to join

**JOIN<sup>2</sup>** strives to deliver a publication database that

- > is centred around **scientists** needs (e. g. reusable citations, private areas)
- > makes **achievements visible** (e. g. by integrating with other systems)
- > is **easy to use**, while ensuring a **high data quality**
- > requires a **small footprint** (for operation and curation)
- > is adoptable to local needs, while enabling **data exchange**
- > allows progress reporting



# Just another Invenio iNstallation to join

**JOIN<sup>2</sup>** strives to deliver a publication database that

- > is centred around **scientists** needs (e. g. reusable citations, private areas)
- > makes **achievements visible** (e. g. by integrating with other systems)
- > is **easy to use**, while ensuring a **high data quality**
- > requires a **small footprint** (for operation and curation)
- > is adoptable to local needs, while enabling **data exchange**
- > allows progress reporting

Every **unwritten** line of **code** is a **good** line of code.

# Preface

As a **community** we are slower as all partners need to adopt.



As a **community** we are slower as all partners need to adopt.

Invenio Upgrades applied:

> 1.0 → 1.1 (checking for API changes)

As a **community** we are slower as all partners need to adopt.

Invenio Upgrades applied:

> 1.0 → 1.1 (checking for API changes)

> 1.1 → 1.1.3 → 1.1.4 → 1.1.5 → 1.1.6 (easy)



As a **community** we are slower as all partners need to adopt.

Invenio Upgrades applied:

- > 1.0 → 1.1 (checking for API changes)
- > 1.1 → 1.1.3 → 1.1.4 → 1.1.5 → 1.1.6 (easy)
- > starting preparation of 1.2 (API changes ⇒ will take long)

As a **community** we are slower as all partners need to adopt.

Invenio Upgrades applied:

- > 1.0 → 1.1 (checking for API changes)
- > 1.1 → 1.1.3 → 1.1.4 → 1.1.5 → 1.1.6 (easy)
- > starting preparation of 1.2 (API changes ⇒ will take long)

Complex requests require some **JOIN<sup>2</sup>** code.

As a **community** we are slower as all partners need to adopt.

Invenio Upgrades applied:

- > 1.0 → 1.1 (checking for API changes)
- > 1.1 → 1.1.3 → 1.1.4 → 1.1.5 → 1.1.6 (easy)
- > starting preparation of 1.2 (API changes ⇒ will take long)

Complex requests require some **JOIN<sup>2</sup>** code.  
API changes delay roll-out.

# JOIN<sup>2</sup> is complex — recent achievements

- > Validate all instances for BASE



# JOIN<sup>2</sup> is complex — recent achievements

- > Validate all instances for [BASE](#)
- > Updated all [OpenDOAR](#) records




# JOIN<sup>2</sup> is complex — recent achievements


- > Validate all instances for [BASE](#)
- > Updated all [OpenDOAR](#) records
- > Update grant harvesting to new [OpenAIRE API](#)



# JOIN<sup>2</sup> is complex — recent achievements


- > Validate all instances for **BASE**
- > Updated all **OpenDOAR** records
- > Update grant harvesting to new **OpenAIRE API**
- > Validate against **OpenAIRE**
  - implement new set specs
  - re-implemented OAI-DC from scratch
  - validate up to level 3 

# JOIN<sup>2</sup> is complex — recent achievements

- > Validate all instances for **BASE**
- > Updated all **OpenDOAR** records
- > Update grant harvesting to new **OpenAIRE API**
- > Validate against **OpenAIRE**
  - implement new set specs
  - re-implemented OAI-DC from scratch
  - validate up to level 3 
- > Brush up <http://join2.de>



# JOIN<sup>2</sup> is complex — recent achievements

- > Validate all instances for **BASE**
- > Updated all **OpenDOAR** records
- > Update grant harvesting to new **OpenAIRE API**
- > Validate against **OpenAIRE**
  - implement new set specs
  - re-implemented OAI-DC from scratch
  - validate up to level 3 
- > Brush up <http://join2.de>
- > implement **steering group** (9 members)

# ...more achievements

> Move to  **GitHub**

# ...more achievements

## > Move to **GitHub**

- include our librarians
- 14 teams (one per instance, devs, librarians, curators etc.)
- 29 members where 18 are librarians

# ...more achievements

- > Move to  **GitHub**
  - include our librarians
  - 14 teams (one per instance, devs, librarians, curators etc.)
  - 29 members where 18 are librarians
- > implement issue tracking:

# ...more achievements

- > Move to  **GitHub**
  - include our librarians
  - 14 teams (one per instance, devs, librarians, curators etc.)
  - 29 members where 18 are librarians
- > implement issue tracking:
  - `join2` (technical issues)
  - `join2-data` (librarians)

# ...more achievements

- > Move to  **GitHub**
  - include our librarians
  - 14 teams (one per instance, devs, librarians, curators etc.)
  - 29 members where 18 are librarians
- > implement issue tracking:
  - join2 (technical issues)
  - join2-data (librarians)
- > reorganize git structure
  - join2 (code)
  - join2-private (configs, hosted at DESY)
  - invenio

# Technical improvements

> libRelease2OpenAccess



# Technical improvements

- > `libRelease2OpenAccess`
- > `implement bibtasklets`
  - `alertengine` (*JOIN<sup>2</sup> edition*)
  - `Release2OpenAccess`
  - `libStatistics` (*complex counters*)
  - `hgfDupes` (*based on ids*)
  - `redlist`
  - `Workflow checker`



# Technical improvements

- > `libRelease2OpenAccess`
- > `implement bibtasklets`
  - `alertengine` (*JOIN<sup>2</sup> edition*)
  - `Release2OpenAccess`
  - `libStatistics` (*complex counters*)
  - `hgfDupes` (*based on ids*)
  - `redlist`
  - `Workflow checker`
- > `implement inter-recordlinking` (*7877\_ + imports*)

# Technical improvements

- > libRelease2OpenAccess
- > implement bibtasklets
  - alertengine (JOIN<sup>2</sup> edition)
  - Release2OpenAccess
  - libStatistics (complex counters)
  - hgfDupes (based on ids)
  - redlist
  - Workflow checker
- > implement inter-recordlinking (7877\_ + imports)
- > add mathml → L<sup>A</sup>T<sub>E</sub>X conversion

# Technical improvements

- > libRelease2OpenAccess
- > implement bibtasklets
  - alertengine (JOIN<sup>2</sup> edition)
  - Release2OpenAccess
  - libStatistics (complex counters)
  - hgfDupes (based on ids)
  - redlist
  - Workflow checker
- > implement inter-recordlinking (7877\_ + imports)
- > add mathml → L<sup>A</sup>T<sub>E</sub>X conversion
- > improve schema.org
- > improve search engine metadata
- > improve importers (e. g. iNSPIRE)

# Technical improvements

- > `libRelease2OpenAccess`
- > `implement bibtasklets`
  - `alertengine` (*JOIN<sup>2</sup> edition*)
  - `Release2OpenAccess`
  - `libStatistics` (*complex counters*)
  - `hgfDuples` (*based on ids*)
  - `redlist`
  - `Workflow checker`
- > `implement inter-recordlinking` (*7877\_ + imports*)
- > `add mathml` → `LATEX conversion`
- > `improve schema.org`
- > `improve search engine metadata`
- > `improve importers` (*e. g. iNSPIRE*)

## 2015 OPEN ACCESS REPOSITORY RANKING



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

## 2015 OPEN ACCESS REPOSITORY RANKING



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

Many issues are non-technical.  
<http://repositoryranking.org/>

# People behind JOIN<sup>2</sup>

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

<sup>a</sup> DKFZ Zentralbibliothek, <sup>b</sup> DESY Library and Documentation, <sup>c</sup> Project Inspire, <sup>d</sup> GSI Library, <sup>e</sup> GSI Core IT, <sup>f</sup> Forschungszentrum Jülich, Zentralbibliothek, <sup>g</sup> MLZ Garching, <sup>h</sup> RWTH Aachen, Hochschulbibliothek, <sup>i</sup> European XFEL, <sup>1</sup> (retired)

plus a little help from our friends

CERN Library, CERN IT, iNSPIRE



# Thank you!



Alexander Wagner  
Deutsches Elektronen-Synchrotron  
Central Library

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

<http://library.desy.de>

