

Invenio User Group Workshop 2019

Monday, June 10, 2019 - Monday, June 10, 2019

Universität Hamburg



Book of Abstracts

Contents

2 repositories in the NTK, Prague	1
WEKO3: an Invenio 3 based Institutional Repository Software - current development status and future plan	1
Invenio at TIND	2
Call for participation to implement a MARC flavor model into INVENIO 3	2
Invenio State of the Union	3
Introduction to JuSER repository and APC integration	4
Research Data Repository @ Universität Hamburg	4
Using Invenio RDM to power biomedical research	4
CERN Open Data portal: Invenio meets big data	4
Experience report: Developing the Repository for the CSMS (Centre for the Study of Manuscript Culture) of the Uni Hamburg with Invenio 3	5
Bye MARC, hello JSON	5
Invenio v3 at RERO	5
Citations and COUNTER-compliant Usage Statistics	5
Hasdai and Data Futures	6
CERN Search - enterprise search on Invenio	6
Invenio at CottageLabs	6
Inveniocrdm project	6
Invenio community building & engagement	7
Open Discussion	8
Closing	8

Invenio User Group Workshop / 5**2 repositories in the NTK, Prague****Author:** Petra Černošková¹¹ *National Library of Technology, Prague***Corresponding Author:** peetulee@seznam.cz

The National Library of Technology (NTK) runs two invenio repositories - National Repository of Grey Literature (NUSL; v1.1.2) and institutional repository of NTK (IDR; v1.0.0). The repositories have partly same typology and data models. We have just started on works on upgrade and creating new data model. For now we are discovering which parts are missing and we still need to keep them as e.g. collections and we try to figure out the best and the easiest solution for our repositories what includes revision of whole concept and their functionalities.

Invenio User Group Workshop / 6**WEKO3: an Invenio 3 based Institutional Repository Software - current development status and future plan****Author:** Masaharu Hayashi¹**Co-author:** Yamaji Kazutsuna¹¹ *National Institute of Informatics, Japan***Corresponding Authors:** yamaji@nii.ac.jp, mhaya@nii.ac.jp

“JAIRO Cloud” has been provided as the shared repository service in Japan since 2012 by National Institute of Informatics. In April 2019, the number of institutional repositories in Japanese universities based on JAIRO Cloud is exceeded 558, which covers 70% of the number of institutional repositories in Japan. Nowadays, it has become to be widely recognized as the institutional repository infrastructure in Japan.

JAIRO Cloud is built using a repository software “WEKO” developed from scratch by National Institute of Informatics. First version of WEKO is released in 2008. WEKO has been developed as a content management module and is characterized by its ease of use and ease of customization. It has been being developed while reflecting actual user voices of JAIRO Cloud.

On other hand, With the rise of the open science movement in recent years, the content handled by repository software is becoming more diverse and bigger than before. As a result, WEKO has been required scalability and extensibility for supporting diversity of research data. However, WEKO built on traditional WEB architectures, had difficulty supporting the requirement. We needed to develop a new generation of WEKO. Therefore, we compared major open source repository software such as DSpace, EPrints and Fedora and chose Invenio version 3 for our purpose and started development of new version of WEKO (hereinafter called WEKO3).

In this presentation, we will briefly introduce our background, and introduce the functions, features, problems, and future plans of WEKO3 under development.

Invenio User Group Workshop / 7**Invenio at TIND****Author:** Audun Bjørkøy¹¹ TIND**Corresponding Author:** audun@tind.io

TIND was started in 2013 as an official CERN spin-off providing Invenio as a service, now serving customers such as UC Berkeley, Caltech, Columbia and the United Nations. Over the past years, TIND has heavily developed on the Invenio framework, including the implementation of many new technologies and standards (e.g. Elasticsearch, IIF) as well as developing small and large modules completely from scratch. Based on these developments, the Invenio framework has been packaged into four different ‘products’ or ‘applications’ offered by TIND. These packages offer streamlined functionalities presented in a nice user interface on top of the flexible Invenio digital library framework:

TIND Institutional Repository: using many of the basic Invenio modules, enhanced with many improvements to the WebSubmit module (approval workflow, authority lookup).

TIND Integrated Library System: based on the BibCirculation module, enhanced with several new modules such as modules for Acquisitions, Serials Management, Loan Rules.

TIND Digital Archive: heavily based around IIF and image streaming in combination with traditional Invenio curation modules.

TIND Research Data Management: based on the Invenio 3 framework, enhanced with a self-developed submission workflow, file management and DOI minting.

This presentation will explain how TIND has worked with and developed on the Invenio framework. We will highlight and demonstrate some of the technologies and modules that TIND has developed over the past years, as well as demonstrate selected customer implementations. Furthermore, we will also give insights into building an infrastructure for deploying and maintaining Invenio as a service for +30 customers (using tools and services such as AWS, docker, Sentry, Grafana).

Invenio User Group Workshop / 8**Call for participation to implement a MARC flavor model into INVENIO 3****Author:** Martin Koehler¹¹ DESY**Corresponding Author:** martin.koehler@desy.de

Currently, JOIN2 uses INVENIO 1.1.7 as base software. JOIN2 is implemented at eight instances and is in use

- as a repository
- as an integrated library system (ILS)
- as a current research information system (CRIS)
- as Article Processing Charges (APC) handling system
- and for Campus literature management (private collection)

INVENIO 1.x was preferred over the other systems on the market thanks to its flexibility and general (internal) standardized data model which enables JOIN2 to tackle its broad scope of needs and allows the required easy extensibility. In particular, the huge potential of the implementation of Authority control unearthed entirely new use cases without any major rewrite.

Among the principal characteristics that led to the selection INVENIO 1.x over its competitors: A library system does not allow users to store their private data. CRIS systems are usually not usable as a repository. APC management is not really implemented in either of these.

Given the library background of the JOIN2 members, the MARC-21 data model of INVENIO 1, even with its imperfections, marked a key point in the system selection.

Using the MARC-21 standard as internal format is in our opinion a good basis for our transition to INVENIO 3 and provides a good link between the different functionalities and use cases:

Using MARC-21-Authority based records allows JOIN2 to identify people, grants, experiments, facilities, affiliations and funding information using interchangeable, shared records. Additionally, a shared set of journal records was derived and is maintained. This set profits from the MARC-21 format, which allows reuse and enrichment of data obtained from other library systems using common MARC-21 and RDA concepts. The resulting data can be reused even outside of JOIN2. Adding MARC-21 Holdings enabled JOIN2 to almost seamlessly integrate the library catalogues at DESY and GSI and all functionalities required to handle Article Processing Charges (APCs), a rapidly growing part of our services.

Given the broad scope of research in the JOIN2 instances, JOIN2 profits a lot from bibliographic expertise at the libraries. The knowledge how to model things appropriately can almost immediately be transferred from the librarian to our internal structure using the same language on both ends: MARC-21 and RDA.

To this end **JOIN2 proposes the implementation of an INVENIO flavor with a MARC-21 data model as “first class citizen”**. This would form the base for the new JOIN2 flavor of INVENIO 3.

As formal procedure we suggest to set up an agreement with all interested parties and CERN, from which stems the INVENIO 3 technology.

Besides the obvious content such as title and description of the project, such an agreement needs to specify the deliverables, work packages and milestones with a schedule, together with a description of the organization including roles and responsibilities of all parties involved.

We believe that the unique combination of repository, ILS, CRIS, APC handling and campus literature management targets academic/research institutions and thus has a large market potential. It enables their libraries to run and maintain this crucial infrastructure, hugely benefiting from the synergies we already rely on at JOIN2 today. The field of application is not restricted to research institutes and libraries: the INVENIO MARC-21 flavor and later the JOIN2 flavor of INVENIO 3 will be useful for governmental organizations and companies, too.

We want to contribute to the to-be-defined work packages of this MARC 21 INVENIO 3 flavor.

Please contact us at library@desy.de, if you are interested in joining our efforts to establish first an INVENIO 3 MARC 21 flavor and later a full INVENIO 3 JOIN2 flavor.

Invenio User Group Workshop / 9

Invenio State of the Union

Author: Lars Holm Nielsen¹

¹ CERN

Corresponding Author: lars.holm.nielsen@cern.ch

An overview of the current status of Invenio, the community, ongoing projects and future projects.

Invenio User Group Workshop / 10**Introduction to JuSER repository and APC integration**

Authors: Zöe Hester¹; Claudia Frick¹

¹ *FZ Jülich*

Corresponding Authors: c.frick@fz-juelich.de, z.hester@fz-juelich.de

An introduction to JuSER repository as a publication database, and open access repository including authorities for reporting. Includes an overview of APC integration in JuSER reporting.

Invenio User Group Workshop / 11**Research Data Repository @ Universität Hamburg**

Author: Kai Wörner¹

¹ *Universität Hamburg*

Corresponding Author: kai.woerner@uni-hamburg.de

Research Data Repository @ Universität Hamburg

Invenio User Group Workshop / 12**Using Invenio RDM to power biomedical research**

Author: Kristi Holmes¹

¹ *Northwestern University*

Corresponding Author: kristi.holmes@northwestern.edu

Learn how the Invenio RDM will be used to support clinical and translational research through work happening in the CTSA Program National Center for Data to Health and Northwestern University.

Invenio User Group Workshop / 13**CERN Open Data portal: Invenio meets big data**

Author: Diego Rodriguez Rodriguez¹

¹ *CERN*

Corresponding Author: diego.rodriguez.rodriguez@cern.ch

The CERN Open Data repository platform manages and disseminates close to two petabytes of open data from particle physics. This includes collision and simulated datasets together with software code examples, associated documentation, configuration files, and virtual machines enabling to explore

the released data for both educational and research purposes. In this talk, we will address how Invenio technology is used in running the service, from managing big research data sets, through keeping the system highly available during high user demand periods, up to providing customised interactive data visualisation and exploration of research data.

<https://gitlab.cern.ch/invenio/template-openshift>

Invenio User Group Workshop / 14

Experience report: Developing the Repository for the CSMS (Centre for the Study of Manuscript Culture) of the Uni Hamburg with Invenio 3

Corresponding Author: cristina.mossoni@uni-hamburg.de

Short presentation (from a software developer's point of view) of the process of developing the new Repository starting from zero.

Invenio User Group Workshop / 16

Bye MARC, hello JSON

Author: Karolina Przerwa¹

¹ *CERN*

Corresponding Author: karolina.zdzislawa.przerwa@cern.ch

The presentation will show some insights about the migration from Invenio 1 to Invenio 3. It will include some real examples with focusing on data management and architecture challenges.

Invenio User Group Workshop / 17

Invenio v3 at RERO

Author: Sébastien Délèze¹

¹ *RERO*

Corresponding Author: sebastien.deleze@rero.ch

There are two main projects based on Invenio v3 at RERO currently, both under active development: RERO ILS and SONAR. RERO ILS is an Integrated Library System, and it includes two sub-projects, RERO-EBOOKS and MEF (Multilingual Entity File). The second main project, SONAR, is a Swiss-wide institutional repository that intends to collect, promote and preserve scholarly publications of authors affiliated with Swiss public research institutions. It aggregates publications from various sources, including existing institutional and subject repositories. SONAR also provides IRaaS (Institutional Repositories as a Service).

What are the realizations and lessons learned?

Invenio User Group Workshop / 18

Citations and COUNTER-compliant Usage Statistics

Author: Chiara Bigarella¹

¹ *CERN*

Corresponding Author: chiara.bigarella@cern.ch

Citations and COUNTER-compliant Usage Statistics

Invenio User Group Workshop / 19

Hasdai and Data Futures

Author: Neil Jefferies¹

¹ *Data Futures*

Hasdai

Invenio User Group Workshop / 20

CERN Search - enterprise search on Invenio

Author: Pablo Panero¹

¹ *Universidad de Oviedo (ES)*

Corresponding Author: pablo.panero@cern.ch

An enterprise search as a service solution based on Invenio 3

Invenio User Group Workshop / 21

Invenio at CottageLabs

Author: Richard Jones¹

¹ *CottageLabs*

Corresponding Author: richard@cottagelabs.com

Invenio at CottageLabs

Invenio User Group Workshop / 22

InveniocrDM project

Author: Lars Holm Nielsen¹

¹ *CERN*

Corresponding Author: lars.holm.nielsen@cern.ch

CERN has partnered with 10 multidisciplinary institutions and companies to build a turn-key open source research data management platform called InvenioRDM, and grow a diverse community to sustain the platform.

The InvenioRDM project is funded by the CERN Knowledge Transfer Fund, as well as all the participating partners, including:

- Brookhaven National Laboratory (US)
- Caltech Library (US)
- Data Futures (UK)
- Helmholtz Zentrum Dresden-Rossendorf (DE)
- Northwestern University (US)
- OpenAIRE (GR)
- TIND (NO)
- Tubitak (TK)
- University of Hamburg (DE)
- University of Münster (DE)

The project has an ambitious one year schedule in which it will deliver:

- InvenioRDM - A research data management platform based on Zenodo and Invenio v3 Framework.
- A community of public and private institutions to sustain InvenioRDM.
- Minimum two existing repositories migrated to InvenioRDM, with Zenodo being one of them.

The key to successfully achieving the ambitious schedule is that InvenioRDM will be based on Zenodo that have already been successfully validated over the past 5 years.

Our vision in the next five years, is to make InvenioRDM a world-leading extensible research data management platform used by research institutions all around the world and with businesses providing services, support and customizations on top of InvenioRDM.

Invenio User Group Workshop / 23

Invenio community building & engagement

Author: Kristi Holmes¹

¹ *Northwestern University*

Corresponding Author: kristi.holmes@northwestern.edu

Join us for an open discussion on open source community building and how to empower your local engagement efforts.

Invenio User Group Workshop / 24

Open Discussion

Authors: Lars Holm Nielsen¹; Kristi Holmes²

¹ *CERN*

² *Northwestern University*

Corresponding Authors: kristi.holmes@northwestern.edu, lars.holm.nielsen@cern.ch

Open discussion about how to empower the Invenio community.

Invenio User Group Workshop / 25

Closing

Author: Lars Holm Nielsen¹

¹ *CERN*

Corresponding Author: lars.holm.nielsen@cern.ch

Closing notes