

Invenio: A Digital Document Repository Framework

Carlos Fernando Gamboa (cgamboa@bnl.gov),
Scientific Data and Computer Center, August 25 2020
SDCC Technical talks series

BROOKHAVEN
NATIONAL LABORATORY

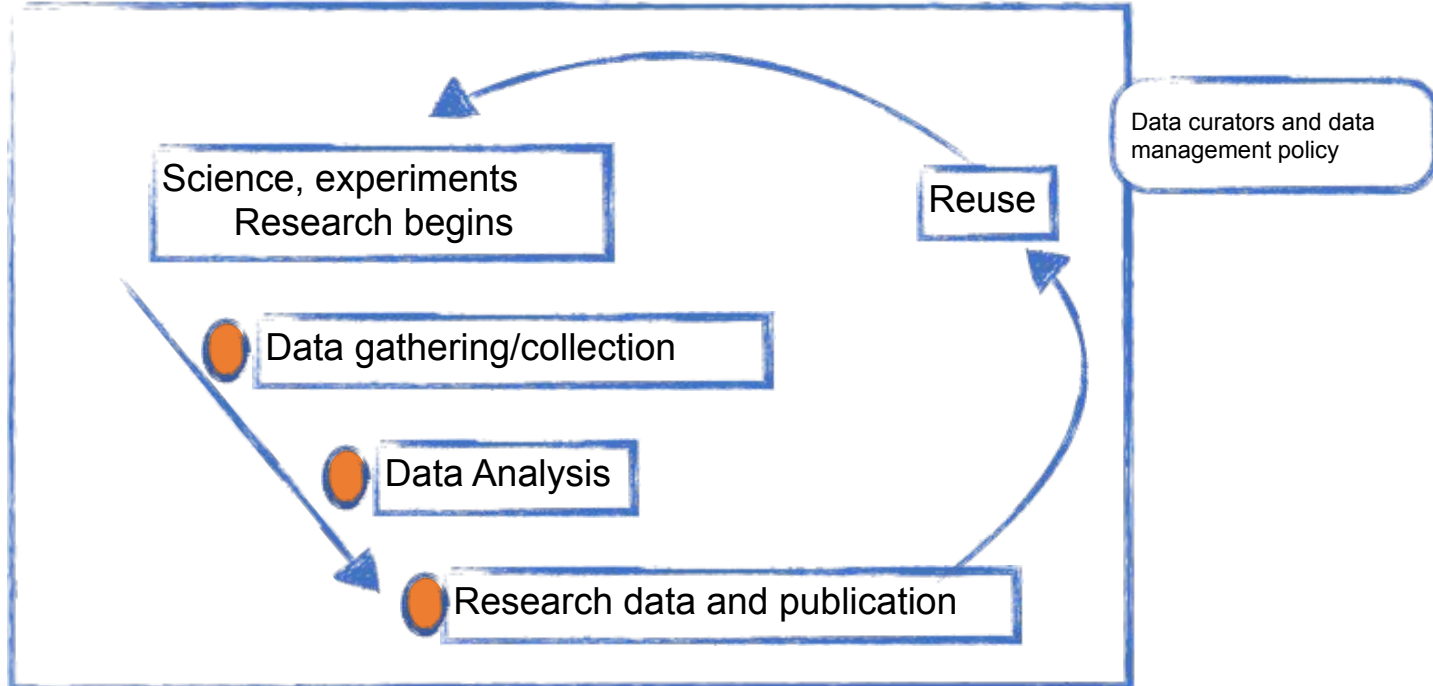


BROOKHAVEN SCIENCE ASSOCIATES

A scientific data workflow

Research data work

Research Digital Management repositories are key elements of infrastructure



Institutional infrastructure supports the repositories

Research Digital Management (RDM) repository

A web based service that provides a scientific community a means to share and preserve their scientific results enable reproducibility and empower reuse of datasets

In recent years, RDMs have been adopting **Open Science** and **FAIR** data policies. In general terms:

- **Open Science**: is the movement to make scientific research and data accessible to all
- **FAIR** refers to a digital record that is:
 - Findable**, metadata are assigned a globally unique and persistent identifier
 - Accessible**, metadata are retrievable by their identifier using a standardized communications protocol. Open or by providing Authentication/Authorization schemes
 - Interoperable**, data need to interoperate with applications or workflows for analysis, storage, and processing
 - Reusable**, metadata and data should be well-described so that they can be replicated and/or combined in different settings

Zenodo (example of a Research Data Management Repository)

<https://zenodo.org>

The screenshot shows the Zenodo repository landing page for COVID-19 related communities. The page features a search bar, navigation links for 'Upload', 'Communities', 'Log in', and 'Sign up'. The main content area is titled 'COVID-19 related communities' and includes a 'Chicago COVID-19 Response' community card with a 'Browse' button and a 'New upload' button. Below this, there are 'Featured uploads related to COVID-19' with several cards for datasets and software, including 'Japanese Sample Tweets, COVID-19 Keywords and Emotions from 2020-01-01 to 2020-06-30' and 'SARS-CoV-2 Infections: Nature Communications final version'. A 'Recent uploads' section is highlighted with a red box, showing a 'List of studies on Covid-19 identified in PubMed and excluded from our systematic review' dataset from August 14, 2020. A 'Need help?' section is also visible on the right side of the page.

Repository landing page

The screenshot shows the Zenodo record landing page for the dataset 'List of studies on Covid-19 identified in PubMed and excluded from our systematic review'. The page displays the dataset title, authors (Carolina Riveros, Olivier Pierre, Leopold Fezeu, Isabelle Boutron, Anna Chaimani, Declan Devane, Joerg J Meerpoth, David Tovey, Asbjørn Hróbjartsson, Philippe Ravaut), and a description of the dataset. A table lists the files, including 'Exclus PubMed Zenodo 14 ao0t.xlsx' (3.7 MB) with a 'Download' button. The page also shows the 'Open Access' status, 'Citations' (0), and 'Versions' (8). A 'Share' section is visible on the right side of the page. Red arrows labeled 'A', 'F', and 'R' point to the 'Communities' link, the DOI, and the 'Versions' section respectively.

Record landing page

Zenodo is built using **Invenio 3** framework

The screenshot shows the 'Share' and 'Export' sections of the Zenodo record landing page. The 'Share' section includes social media icons for Twitter, Facebook, LinkedIn, and YouTube, along with a 'Share' button. The 'Export' section includes a 'Start typing a citation style...' input field and a list of export formats: BibTeX, CS, DataCite, Dublin Core, DCAT, JSON, JSON-LD, GeoJSON, MARCXML, and Mendley.

Invenio 3 is a open source framework to build scalable digital repositories

Integrated in a scalable software architecture

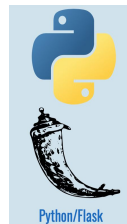
- Flexible record and persistent identifier store
- Record can use custom or standard metadata formats like JSON-LD, MARC21, Datacite
- Invenio can manage bibliographic records, authority records, grants among others
- DOI (Digital Object Identifier) to support records to be properly citable
- Elasticsearch is leveraged by Invenio to provide scalable and complex searching capability

Accessibility enabled for web UI or programmatically via a REST API

- Implemented for metadata and files
- Invenio supports different data transmission and storage protocols (e.g S3, XRootD, WebDAV, among others)

State of the art authentication/authorization implementation

- Single Sign On and Authentication OAuth allows integration with Github, ORCID out of the box



Invenio software is distributed as modular framework

Code is released in modules distributed in **bundles**

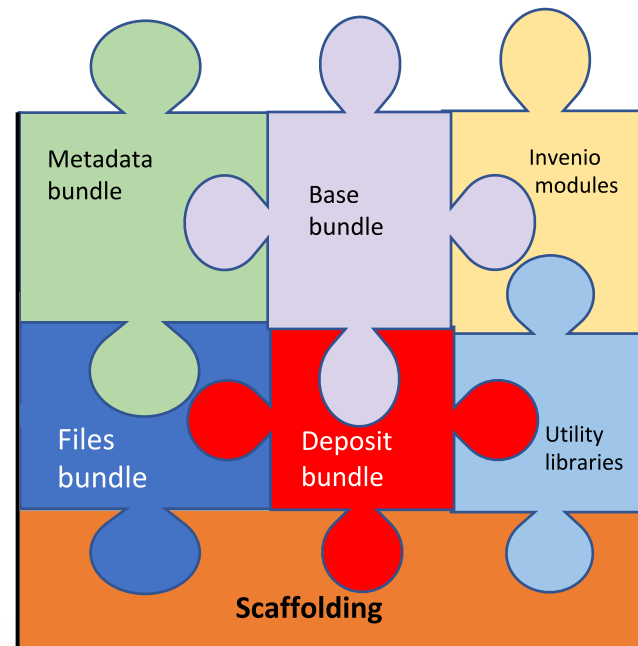
The Base bundle aggregates modules needed to create a generic web application, i.e:

- **Invenio-config:** Configuration loading pattern responsible for loading configuration from Python modules, instance folder and environment variables
- **Invenio-app:** Flask, WSGI, Celery and CLI applications for Invenio including security-related headers and rate limiting
- **invenio-admin:** Administration interface for Invenio based on Flask-Admin

Other bundles and sample modules are :

- **Auth bundle:** invenio-oauth
- **Files bundle:** invenio-files-rest
- **Statistics bundle (beta):** invenio-stats
- **Deposit bundle (alpha):** invenio-deposit
- **Invenio modules (alpha):** invenio-github
- **Utility libraries:** Datacite
- **Scaffolding:** cookiecutter to create base application template

An invenio base digital repository



Examples of digital repositories on Invenio 3 hosted at CERN

The image displays three screenshots of digital repositories hosted on Invenio 3 at CERN:

- Zenodo:** A screenshot of the Zenodo website showing recent uploads, including a historical climate model simulation and a reproducibility package for AGILE conference papers. The interface includes a search bar, navigation links, and a sidebar with various utility links.
- CERN Video Platform:** A screenshot of the CERN Video Platform showing a video titled "LINAC4 joins the CERN accelerator chain". The interface includes a search bar, navigation links, and a section for recent videos.
- CERN Open Data Portal:** A screenshot of the CERN Open Data Portal showing a search bar, navigation links, and a section for explore and focus on data.

Below the screenshots are three blue boxes with white text:

- Zenodo**
~100 TBs / 2M records
- CERN Video Platform**
~35 TBs / >5000 videos
- CERN Open Data Portal**
1.7 PBs / >620K files

Invenio 3 at BNL

Initially we investigated digital repository options for BNL's science programs

- After evaluation and testing Zenodo was implemented as a R&Ds testbed for NNSD and CSI users
- By interacting with the Invenio framework and testing its capabilities, these communities built their own digital repositories to meet their specific needs (SET and GENESIS)

Most recently new BNL scientific communities (NPP) and DOE Medical Therapeutics are interested in a RDM repository like Zenodo

Now Invenio based repositories is a service supported as part of SDCC mission

Digital repositories hosted at BNL

SDCC supports *custom* data repositories based on invenio for different scientific communities:

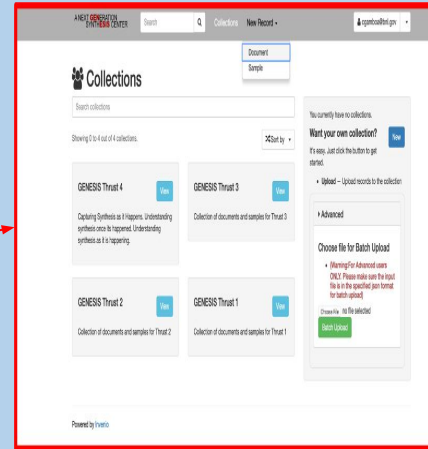
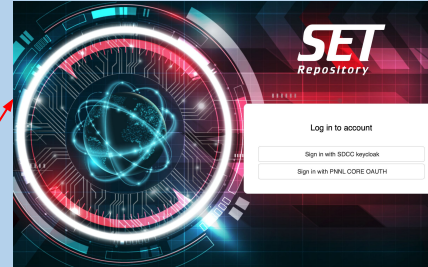
National Nuclear Security Administration (NNSA)

Application **SET**, Smuggling Detection and Deterrence Science and Engineering Team

Materials Science community

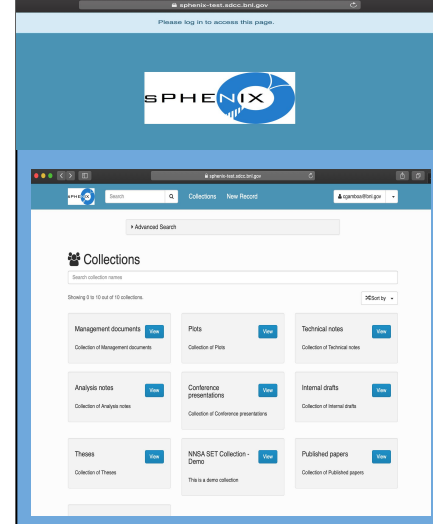
Application **GENESIS**, Next-Generation Synthesis Center

Repositories in operation



Repository in development and testing

Nuclear and Particle Physics (NPP)



SDCC supports infrastructure for Invenio based applications, along with customized network, storage and Authentication infrastructure enabled to host services (production, testing and developing)

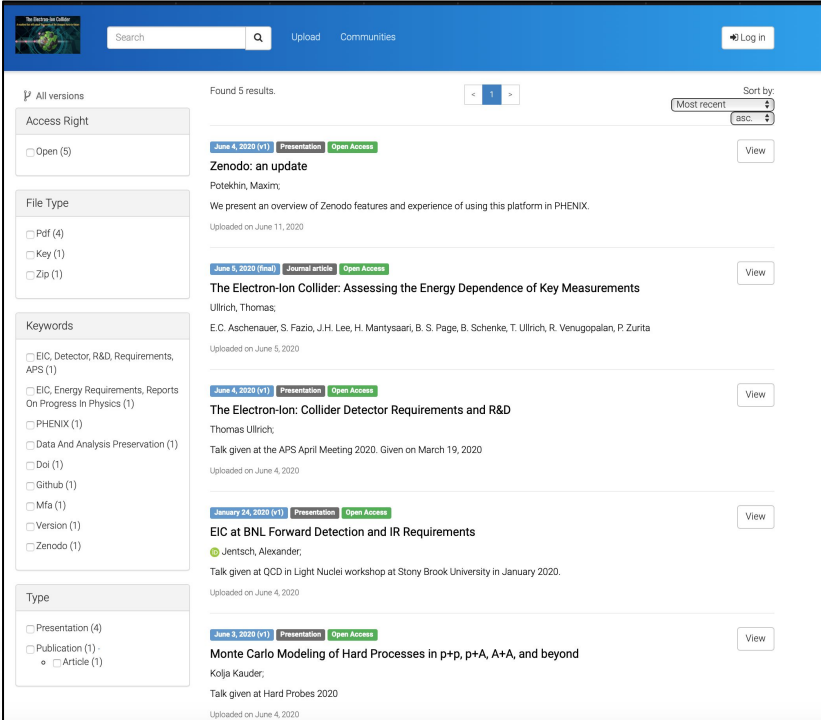
Invenio 3 at BNL: Zenodo based repositories

EIC-Zenodo

Learning about Zenodo:

- Community was able to experiment its features using a test instance
 - Helped identify requirements
- CILogon, Federated ID (InCommon / COMange) used for authorization (allows to use institutional credentials to login into the web application)
- Based on this experience the community requested a production instance

A production EIC-Zenodo instance being commissioned and is accessible to BNL



The screenshot displays the Zenodo web interface. At the top, there is a search bar, navigation links for 'Upload' and 'Communities', and a 'Log in' button. Below the search bar, a sidebar on the left offers filters for 'All versions', 'Access Right' (with 'Open (5)' selected), 'File Type' (with 'Pdf (4)', 'Key (1)', and 'Zip (1)' listed), 'Keywords' (with various physics-related terms like 'EIC, Detector, R&D, Requirements, APS (1)'), and 'Type' (with 'Presentation (4)', 'Publication (1)', and 'Article (1)' listed). The main content area shows search results for 'Zenodo: an update' by Potekhin, Maxim, with a 'View' button. Below it, another result 'The Electron-Ion Collider: Assessing the Energy Dependence of Key Measurements' by Ullrich, Thomas, is shown. Further down, 'The Electron-Ion: Collider Detector Requirements and R&D' by Thomas Ullrich is listed. At the bottom, 'EIC at BNL Forward Detection and IR Requirements' by Jentsch, Alexander, and 'Monte Carlo Modeling of Hard Processes in p+p, p+A, A+A, and beyond' by Kolja Kauder are visible. Each result includes a date, version, and 'Open Access' status, along with a 'View' button.

Covid-19-archive, a BNL custom digital repository based on Zenodo

Being commissioned to host COVID-19 related digital documents as a part of DOE COVID - Medical Therapeutics project based on Zenodo software DOE project for which BNL receives funding

A selected group of researchers uploads and curates the documents in the repository:

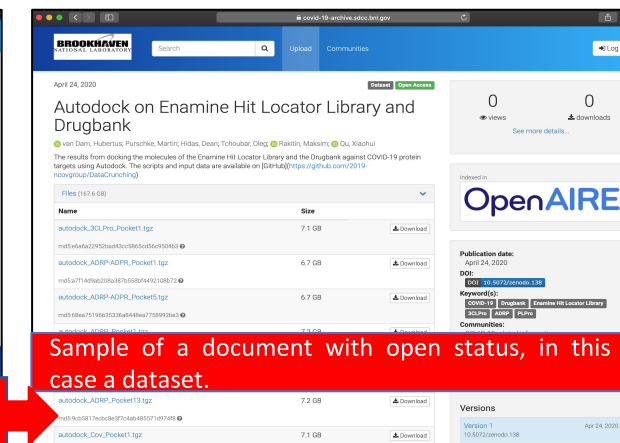
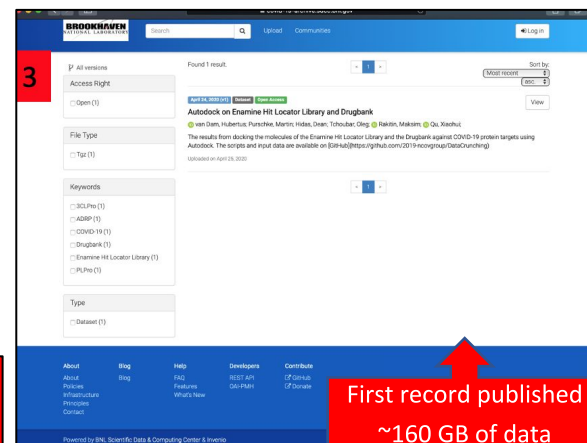
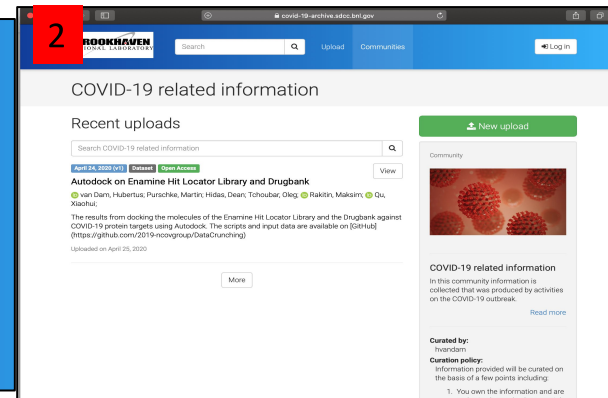
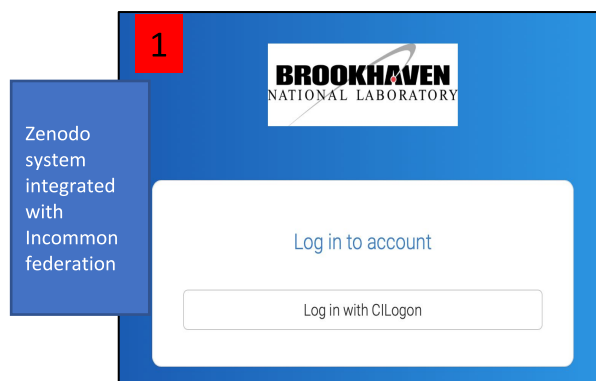
1 The selected researches will be able to use their institution's (ANL, ORNL ,.., BNL) login and passwords to authenticate to the system

2 A *community* can be created to collect and curate topic/theme centric aggregation of documents

3 General users will be able to download data (files) from the repository based on **document status**:

- **Open**, can read and download
- **Restricted**, can request access
- **Embargo**, once the embargo period ends the document is publicly available
- **Closed**, not permitted

Zenodo will be migrated to invenioRDM



Why InvenioRDM?

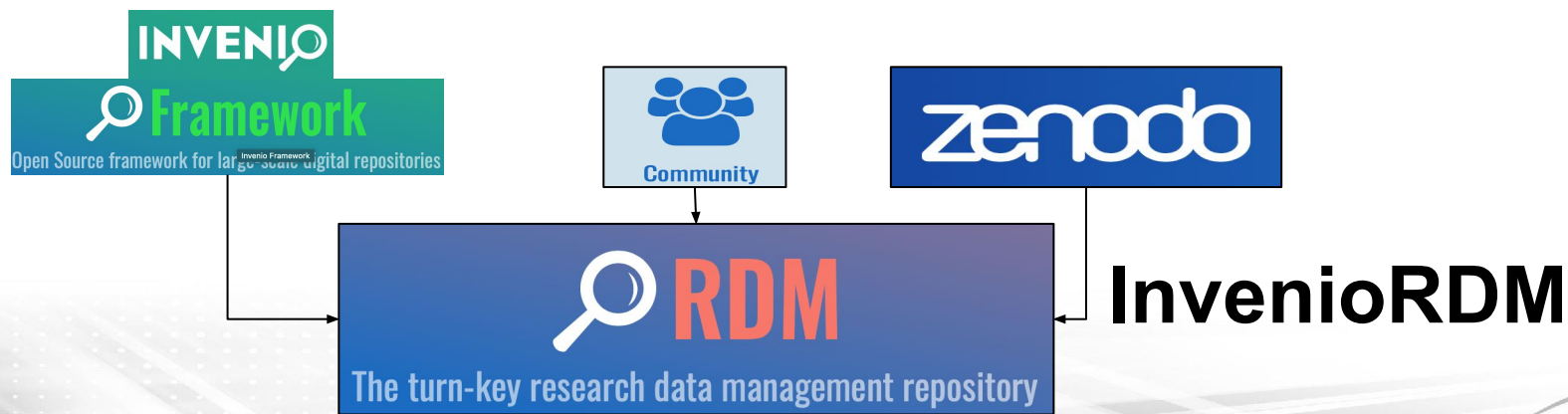
Different institutions were looking for a solution for a RDM for their communities

Zenodo was seen as a model to create their local RDM. However, while Zenodo code is released as open source it is not designed to be deployed outside of CERN

Other institutions tried to build a RDM using Invenio 3 framework. However, Invenio 3 *is a code library* used to build an application from scratch (e.g CERN Open Data or Zenodo)

Many organizations tried to share and reuse code modifications with not an easy portability

These interested multidisciplinary institutions gathered to create a collaborative open source (invenioRDM project) and grow a sustainable community. This project will provide a platform for institutions to be able to install their own RDM.



What is InvenioRDM?

InvenioRDM *is an application* built using invenio 3 aiming to be:

- A turn-key research data management repository: minimize the amount of developing work and code support to customize the RDM to the institution's needs
- Community supported: it is envisioned that a community of research institutions, private companies and individuals will be growing and help sustain this platform

It is planned that Zenodo will be migrated to InvenioRDM once it is released

The screenshot shows the InvenioRDM website interface. At the top, there's a navigation bar with the InvenioRDM logo, 'Uploads', 'Communities', and a 'Log in' button. Below the navigation bar is a search bar with the text 'july-release-rdm-new' and a search icon. The main content area features a large heading: 'You've successfully installed InvenioRDM! What is ahead?'. Below this heading are three numbered steps: 1. Configure, 2. Customize, and 3. Deploy it!. Each step has a corresponding icon and a brief description. Step 1: 'Configure' - 'Configure the datamodel, resource types, permissions and much more to make it suit your needs! For more information on how to do it visit the configuration section in the docs.' Step 2: 'Customize' - 'Customize the layout and adapt the styling to match your institution's. See how to do so here. You can also add new functionality using supported plugins, or writing your own.' Step 3: 'Deploy it!' - 'Use the Helm charts to deploy your InvenioRDM instance. You can use the Helm chart to deploy in OpenShift.' Below the steps are two buttons: 'Docs' and 'Code'. At the bottom of the page, there's a section titled 'Brought to you by' with logos of various institutions: Brookhaven National Laboratory, Caltech Library, CERN, National Center for Data to Health, data futures, GEO GROUP ON EARTH OBSERVATIONS, HZDR HELMHOLTZ ZENTRUM DRESDEN-ROSENDORF, INFN, JRC, NORTHWESTERN UNIVERSITY, OpenAIRE, and TIND.

InvenioRDM: Benefits and Features

Benefits

- **Research safely shared:** Permits share and preserve records with collaborators
- **Communities:** Allows user the creation and management of its community (e.g. journal, project, workshop).
- **Implements FAIR like policies for data deposited:** is achieved by a collection of a robust metadata in conjunction of an open API and powerful search index.
- **DOI persistent Identifier:** is available for citation and compliance with data sharing requirements
- **Simplicity:** Turn-key research data management platform can be installed in the local environment or by a service provider

Features

- **Class UX:** enhanced with user experience in mind
 - End-users, curators, sys admins and developers
- **Repository Profiles:** Comes with pre-configured repositories, Institutional Repositories (IRs), Research Data Management repository (RDM) and domain-specific repositories for health and biomedical sciences
- **Other features includes Resilience, Scalable and institutional integration**

InvenioRDM collaboration

Github is used to host:

Repository: <https://github.com/inveniosoftware/invenio-app-rdm>

Documentation: <https://invenio-app-rdm.readthedocs.io/en/latest>

Effort is coordinated by project boards:

- Priorities definition and its documentation
- Allows to identify or trace issues of the monthly developing sprints
- Example for this month

<https://github.com/orgs/inveniosoftware/projects/4>
7

Invenio Request For Comments (RFC)

- Coordinate the design process
- Creates consensus with the parties involved
- Helps document invenioRDM development

 <https://github.com/inveniosoftware/rfcs>

Brought to you by

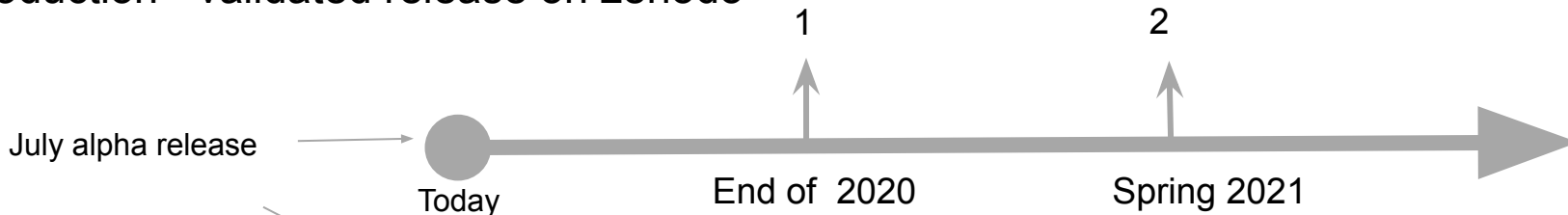


Collaborators

InvenioRDM (status)

Milestones

- 1) First minimal release - bare bones
- 2) Production - validated release on zenodo



The screenshot shows a web page titled "INVENIO" with a sub-header "InvenioRDM Alpha 9 (July Release)". The page content includes a post by "fenekeu" dated "4 Aug" with "1" comment and "14c" views. The post text reads: "We are happy to announce InvenioRDM Alpha 9 (July release)". Under "What's new?", it states: "The July release adds translation support and starts integrating the major backend development of last release into InvenioRDM. In particular, draft functionality has been added at the API level and is used under the hood for deposits. To achieve this integration, the hard problems like responsibility separation, error handling, pagination, linking and internal PID management were solved. Other improvements and more details follow." It also includes an update to "Invenio-CLI to version 0.15.0" and a "Module translation" section thanking "TUBITAK" for enabling internationalization. At the bottom, there is a small image of the InvenioRDM interface with the text "InvenioRDM başkanı ile yükleddiniz!" below it.

invenioRDM at BNL

Electron Ion Collider (EIC), the DOE covid-19-archive Therapeutics and PHENIX interested in using it

- Zenodo based deployments will be migrated to invenioRDM once released

Other invenio based repositories (sPhenix) on development at BNL are looking at invenioRDM

OSTI (Office of Scientific and Technical Information) DOI integration with invenioRDM begun:

- DOE mandates that work funded by or performed at DOE labs should be registered with OSTI
 - OSTI DOIs are free
- Invenio-based repositories register DOIs directly through DataCite
- Project initiated by SULI summer student

https://indico.bnl.gov/event/9091/contributions/40194/attachments/29865/46597/InvenioRDM_UI_and_DOI_Enhancement_Summary.pdf

invenioRDM as is (July alpha release)

The screenshot shows the record page for a publication titled "A Romans story" by Julio Cesar. The page includes a header with the INVENIO RDM logo, a search bar, and navigation links for "Uploads" and "Communities". The record details include the date "August 10, 2020", version "v0.0.1", and a "Publication / Journal article" type. A "Metrics" section displays 1,175 views, 52 downloads, and 8 citations. A "Versions" section shows the current version (v0.0.1) and a previous version (10.9999/rdm.9999999) dated Aug 10, 2020. The page also features a "Contact" section with a "Send message" button, a "Citation" section, and a "Keywords" section. The DOI is 10.9999/rdm.9999999. The page is labeled "Record page" at the bottom.

The screenshot shows the community page in invenioRDM. The header is identical to the record page. The main content area displays search results for "4 result(s) found". The results are sorted by "Best match" and include a "New community" button. The results list includes "Electron Ion Collider" (Test EIC community), "sPhenix" (sPhenix test community), "Phenix" (Test PHENIX research archive), and "test" (test community). The page is labeled "Community page" at the bottom.

Customization of different User Interfaces (repository landing, record and community pages) will need to be done to reflect BNL's communities needs.

To facilitate the customization of these components invenioRDM uses:

- Semantic UI: Web design framework for responsive layouts with intuitive/human-friendly HTML
- ReactJS: JavaScript framework for real-time rendering/updating of pages

invenioRDM at BNL near future work

A invenioRDM (“bare bones release”) testbed installation is expected to be available for interested BNL communities

- Definition of the record model
- Possibility to identify communities that can use a multi-disciplinary research repository

Expected to start transition of BNL invenio repositories to invenioRDM in spring 2021

Summary

An overview of invenio related projects at BNL was presented

SDCC has a solid expertise in hosting and operating digital document repositories based on Invenio

SDCC is supporting digital repositories for different scientific communities within BNL and DOE in the US

SDCC is active within international Invenio community to ensure local community interests are well represented

BNL scientific communities will use invenioRDM to host their digital research records

References

Invenio, <https://invenio-software.org>

InvenioRDM, <https://invenio-software.org/products/rdm/>

invenioRDM user docs, <https://inveniordm.docs.cern.ch>

DOE OSTI, <https://www.osti.gov/>

DataCite, <https://datacite.org>

Backup slides

FAIR data policy

To be Findable	To be Accessible	To be Interoperable	To be Reusable
<p>F1: (meta)data are assigned a globally unique and persistent identifier. A DOI is issued to every published record on InvenioRDM.</p>	<p>A1: (meta)data are retrievable by their identifier using a standardized communications protocol Metadata for individual records as well as record collections are harvestable using the OAI-PMH protocol by the record identifier and the collection name. Metadata is also retrievable through the public REST API.</p>	<p>I1: (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation. InvenioRDM uses JSON Schema as internal representation of metadata and offers export to other popular formats such as Dublin Core or MARC-XML.</p>	<p>R1: (meta)data are richly described with a plurality of accurate and relevant attributes Each record contains a minimum of DataCite's mandatory terms, with optionally additional DataCite recommended terms and InvenioRDM's enrichments.</p>
<p>F2: data are described with rich metadata (defined by R1 below). InvenioRDM's metadata is compliant with DataCite's Metadata Schema minimum and recommended terms, with a few additional enrichments.</p>	<p>A1.1: the protocol is open, free, and universally implementable See point A1. OAI-PMH and REST are open, free and universal protocols for information retrieval on the web.</p>	<p>I2: (meta)data use vocabularies that follow FAIR principles For certain terms we refer to open, external vocabularies, e.g.: license (Open Definition), funders (FundRef) and grants (OpenAIRE).</p>	<p>R1.1: (meta)data are released with a clear and accessible data usage license License is one of the mandatory terms in InvenioRDM's metadata, and is referring to an Open Definition license. Data downloaded by the users is subject to the license specified in the metadata by the uploader.</p>
<p>F3: metadata clearly and explicitly include the identifier of the data it describes. The DOI is a top-level and a mandatory field in the metadata of each record.</p>	<p>A1.2: the protocol allows for an authentication and authorization procedure, where necessary Metadata are publicly accessible and licensed under public domain. No authorization is ever necessary to retrieve it.</p>	<p>I3: (meta)data include qualified references to other (meta)data Each referenced external piece of metadata is qualified by a resolvable URL.</p>	<p>R1.2: (meta)data are associated with detailed provenance All data and metadata uploaded is traceable to a registered InvenioRDM user. Metadata can optionally describe the original authors of the published work.</p>
<p>F4: (meta)data are registered or indexed in a searchable resource Metadata of each record is indexed and searchable directly in InvenioRDM's search engine immediately after publishing. Metadata of each record is sent to DataCite servers during DOI registration and indexed there.</p>	<p>A2: metadata are accessible, even when the data are no longer available. Data and metadata will be retained for the lifetime of the repository. Metadata are stored in high-availability database servers which are separate to the data itself. (note: recommendations for local implementations should be established here)</p>		<p>R1.3: (meta)data meet domain-relevant community standards InvenioRDM is not a domain-specific repository, yet through compliance with DataCite's Metadata Schema, metadata meets one of the broadest cross-domain standards available.</p>

sPHENIX Document store

Invenio custom application



Log in to account

Sign in with SDCC keycloak

Integrated with SSO
using SDCC keycloak
infrastructure

Filters for search

Collections

Beta release

Collections view

InvenioRDM: August project board

<https://github.com/orgs/inveniosoftware/projects/47>

Invenio digital repository framework

Repositories 134 Packages People 154 Teams 139 Projects 9

InvenioRDM August Board Updated 3 hours ago

Filter cards + Add cards Fullscreen Menu

- 3 Invenio Sprint Cross-over**
 - Todo
 - SUI: review remaining bootstrap related css classes in semantic-ui templates
 - 0 of 7
 - invenio-frontend#124 opened by zzacharo
 - settings: applications missing styling
 - invenio-app-rdm#186 opened by fenekku
 - Move InvenioRDM modules to use pytest in run-tests.sh
 - 6 of 8
 - invenio-app-rdm#191 opened by fenekku
- 22 Todo**
 - invenio-rdm-records#151 opened by fenekku
 - command pattern to hook actions at service level
 - invenio-drafts-resources#36 opened by ppanero
 - second round to the pid manager (config + move of package)
 - invenio-drafts-resources#33 opened by ppanero
 - verify links of docs
 - invenio-app-rdm#97 opened by ppanero
 - good first is...
 - service: filtering of record metadata according to permissions
 - invenio-records-resources#41 opened by lnleisen
 - enhancement
 - Implement search for versions
 - invenio-drafts-resources#17 opened by fenekku
 - Document translation VSCode extension
 - invenio-app-rdm#223 opened by fenekku
 - schema: internal fields
 - invenio-rdm-records#143 opened by ppanero
 - Integrate versioning.
 - invenio-records-resources#51 opened by fenekku
 - [Suggestion] ResourceConfig refactor
 - flask-resources#73 opened by fenekku
 - Document setup from a new developer point of view
 - invenio-app-rdm#231 opened by fenekku
 - good first is...
 - Indexing: use `schemas` field
 - invenio-drafts-resources#27
- 5 In progress**
 - 18h workflow: translation files (.mo) are not being created
 - invenio-app-rdm#230 opened by ppanero
 - destroy command
 - invenio-cli#120 opened by tmorrell
 - 1 linked pull request
 - Use response.data.links.self_html when configured properly
 - react-invenio-deposit#50 opened by fenekku
 - Get edit link from response (when generated there)
 - react-invenio-deposit#49 opened by fenekku
 - naming: resources
 - invenio-rdm-records#142 opened by ppanero
- 2 Blocked**
 - frontpage: add eko-connect logo
 - invenio-app-rdm#178 opened by ppanero
 - 1 linked pull request
 - August release checklist
 - 2 of 7
 - invenio-app-rdm#233 opened by fenekku
- 7 In Review**
 - Command pattern for DraftActionResource
 - invenio-drafts-resources#35 opened by ppanero
 - Changes approved
 - adapt to indexing split
 - invenio-rdm-records#147 opened by ppanero
 - Draft
 - errors: remove unused classes
 - invenio-records-resources#65 opened by ppanero
 - indexer cls: model should not be a cls attr
 - invenio-indexer#119 opened by ppanero
 - 1 linked pull request
 - Ensure links are generated
 - 10 of 12
 - invenio-rdm-records#149 opened by fenekku
 - 1 linked pull request
 - config: links
 - invenio-drafts-resources#8 opened by ppanero
 - service: args vs ctx
 - invenio-drafts-resources#26 opened by ppanero
 - question
 - 1 linked pull request
 - 19 Done**
 - Pass UI routes to Service
 - invenio-records-resources#55 opened by fenekku
 - Implement single draft "self" link generation
 - invenio-drafts-resources#19 opened by fenekku
 - fix titled name
 - invenio-drafts-resources#40 opened by fenekku
 - Changes approved
 - clean up: remove unused exceptions
 - invenio-records-resources#64 opened by ppanero
 - 1 linked pull request
 - use cmd patter for actions
 - invenio-drafts-resources#31 opened by ppanero
 - 1 linked pull request
 - [EPIC] Publish new version record
 - invenio-app-rdm#219 opened by ppanero
 - 1 linked pull request
 - Set RecordDraftSchema.JSONV1 as the default schema of RecordDraftJSONSerializer
 - invenio-drafts-resources#37 opened by fenekku
 - 1 linked pull request
 - data validation: use a single one for drafts and records
 - invenio-drafts-resources#30 opened by ppanero
 - 1 linked pull request
 - improvements
 - 6 of 6
 - invenio-drafts-resources#28 opened by ppanero

Automated as To do Manage

Automated as Done Manage

invenioRDM at BNL

Integration with CiLogon

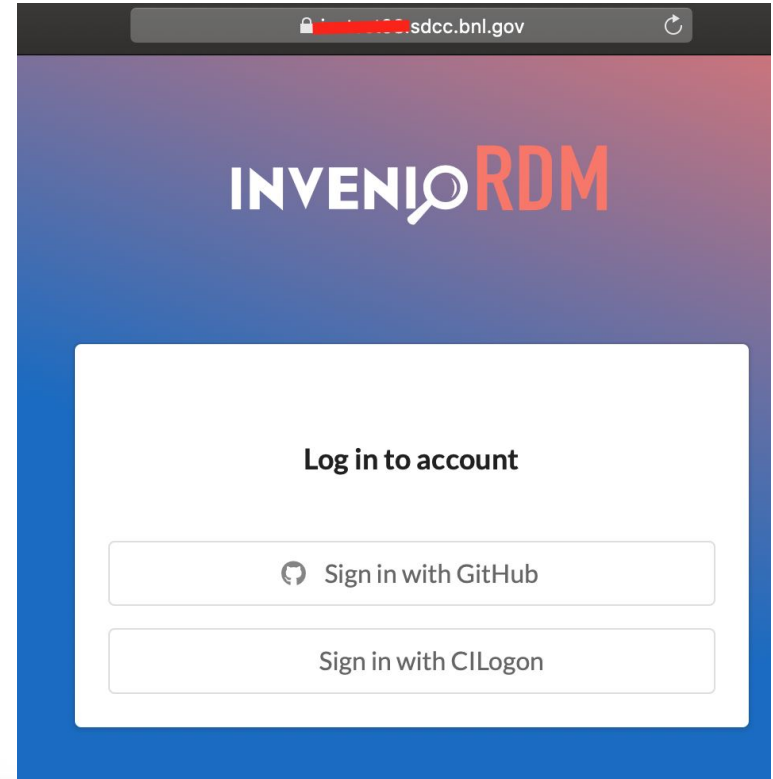
Out of box invenioRDM is integrated with Github OAuth authentication

A CiLogon/Incommon OAuth plugging is being customized for invenioRDM

Module: invenio-oauthclient-1.4.0a1

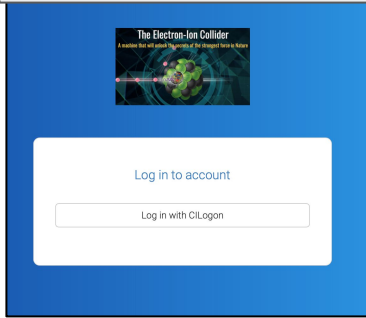
CiLogon.py plugging locally supported

First known use case of CiLogon/Incommon for invenioRDM

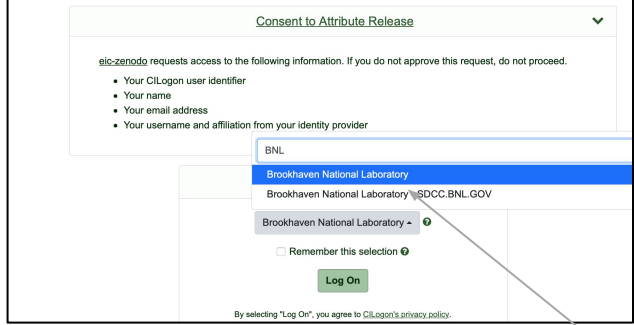


Invenio 3 authentication flow using Incommon Federated Id/COMange

1 Zenodo login page

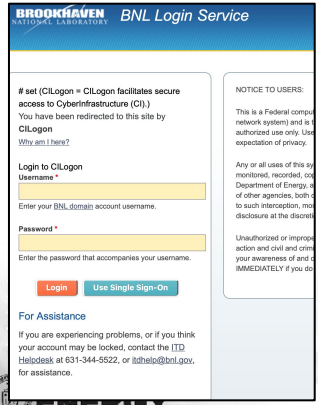


2 CILogon consent screen

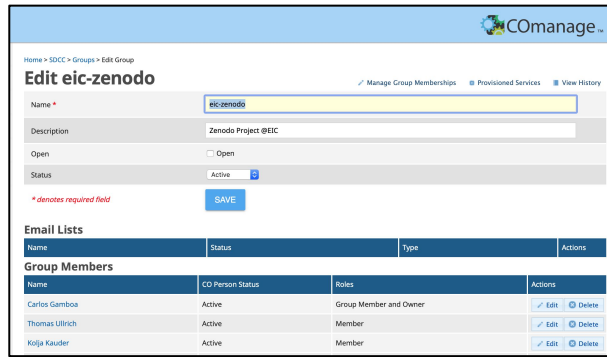
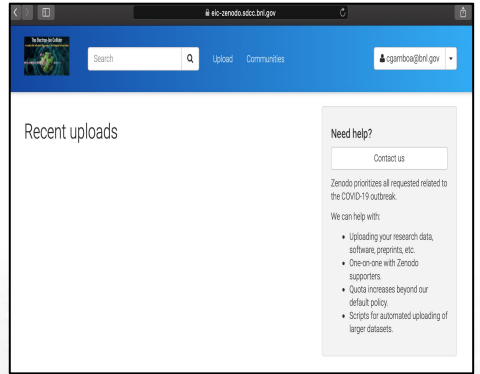


Incommon Fed ID/COMange integrated and used to restrict write access to the Zenodo instance

3 Brookhaven BNL Login Service



4 Zenodo interface



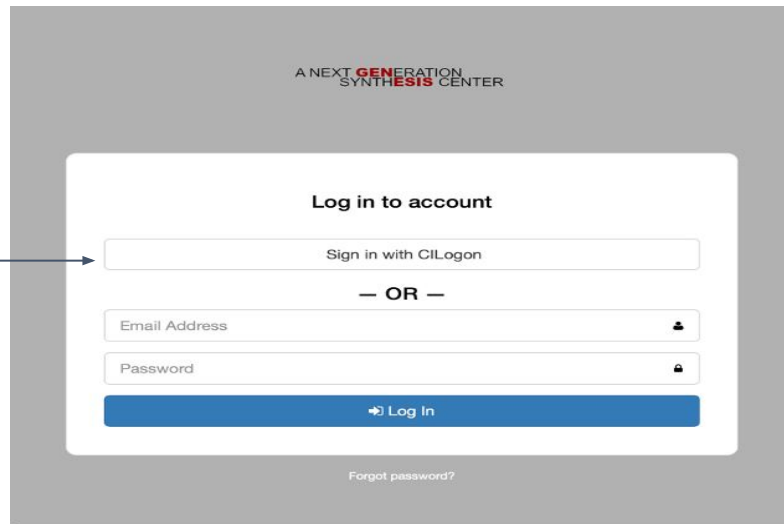
For now users allowed to login must:

- Belong to the COMange eic-zenodo group
- Use BNL Incommon IDPs to login

Versatile authentication mechanisms supported and integrated with the applications to provide SSO accessibility to users

GENESIS, Materials Science community

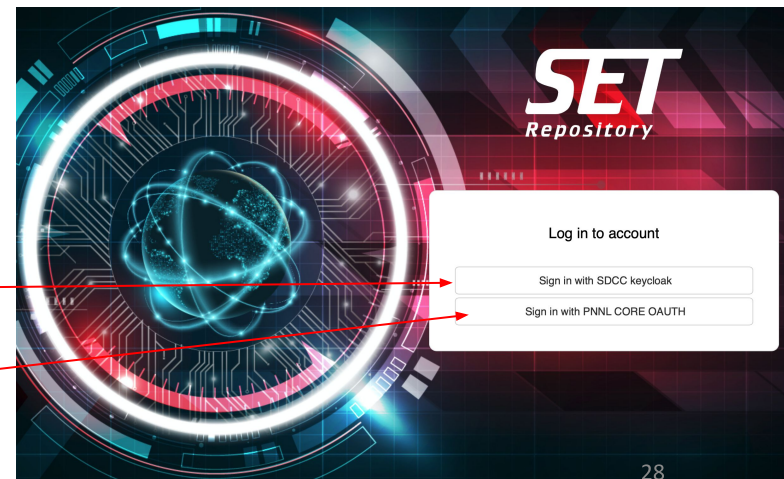
CILogon, Federated ID (InCommon / COMmange) support for authentication using Invenio's oauth native module.



SET, National Nuclear Security Administration (NNSA)

Due to the type of data been managed (OUO) provisioning the production environment required the deployment of isolated resources.

- 2FA DUO capable for authentication using Keycloak and Invenio's oauth native module. (Integrates ITD Active Directory + DUO)
- 2FA using PNNL's IDP
- DOE OneID federation to be integrated in May 2020



Contacts

SDCC, Eric Lancon, Tony Wong

EIC Zenodo, Jerome Lauret

Invenio, **CERN**, Jose Benito Gonzalez, Lars Holm Nielsen

sPHENIX, John Haggerty, Chris Pikenburg

SET, NNSD Warren Stern, Maia Gemmill, Yonggang Cui, Heather Orr (adjunct developer)

GENESIS, Line Pouchard

COVID-19-ARCHIVE, Kerstin Kleese Van Dam

CSI (SET, GENESIS, sPhenix), Developer, Uma Ganapathy