

# INVENIO RDM

*A Next-Generation Repository for Research Data  
Management*

**Matt Carson**

Head, Digital Systems

**Sara Gonzales**

Data Librarian

**Guillaume Viger**

Senior Developer

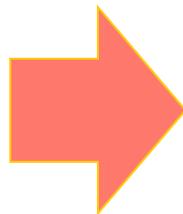
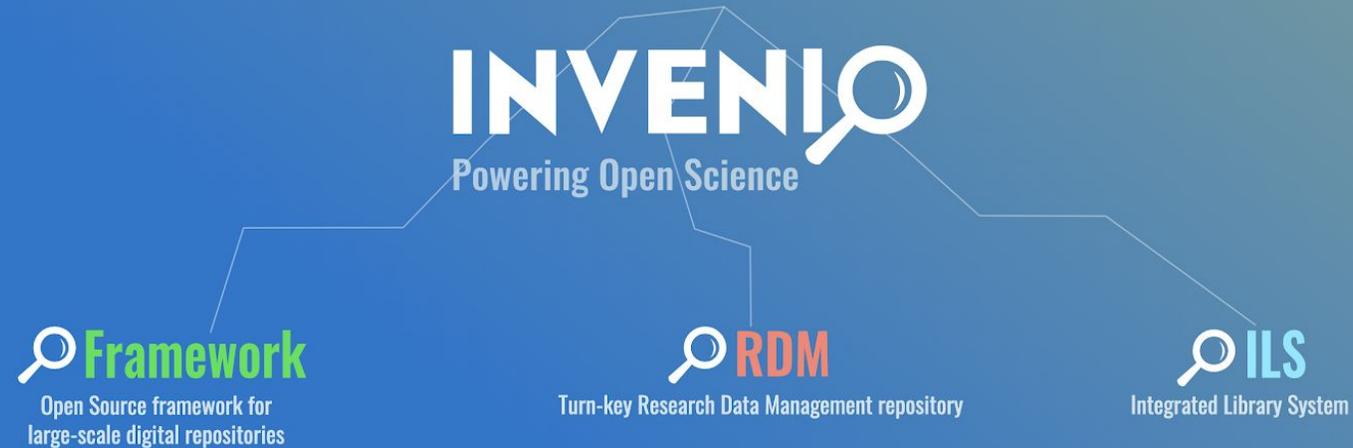
Galter Health Sciences Library & Learning Center  
Feinberg School of Medicine  
Northwestern University



@inveniosoftware

Biomedical Data Science Day: February 4, 2020

# Invenio software powers open science



## Free Open Source Software

Invenio is Free Open Source Software supported by a committed community of multidisciplinary institutions.

[Code](#) | [Docs](#) | [Examples](#)



## Friendly and Responsive Community

Although Invenio was born at CERN, its community is growing bigger every day. Talk to the team now in our chatroom or forum.

[Chatroom](#) | [Forum](#) | [Get Involved](#) | [Events](#)



## The "Safe bet"

Invenio community has been around for 20 years. **Solid services** have been built on top of it to ensure long-term confidence.

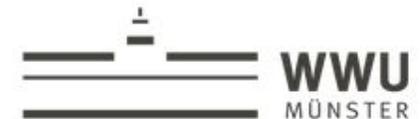
[Live services](#) | [Products](#)

# InvenioRDM features

*RDM platforms are critical to help preserve and share research, enable reproducibility, and empower reuse of datasets, protocols, engagement or study materials, & a wide range of other research products.*

- **Research, shared.** Securely share and preserve data records and research materials with collaborators. Allows easy dissemination to the community.
- **Get credit & be cited.** Get a DOI to make records easily and uniquely citable. Pre-formatted citation text makes it easy to cite your work and be cited.
- **Discoverable.** Leverages metadata standards and the powerful Elasticsearch full-text search engine retrieves, facets, sorts, and filters your searches with ease.
- **Metrics.** Industry standard usage statistics for record pages with all tracking completely anonymized.
- **FAIR.** Advanced features to make your research Findable, Accessible, Interoperable, & Reusable.
- **Communities.** Create and curate your own community (e.g., workshop, project, lab, or journal).
- **Comply.** Comply with data sharing mandates\* and acknowledge your grants.
- **Easy.** Turn-key research data management platform & index can be easily deployed in the local environment by your team or by a service provider, such as TIND. Customize the look and feel to your local environment.

# InvenioRDM collaborators



Work presented here is supported in part by the CERN Knowledge Transfer Fund, CD2H (NCATS U24TR002306), and InvenioRDM collaborators



## The turn-key research data management repository

 Launching in the summer 2020

### Roadmap

We intend to be ready by summer 2020.

### Talk

Join our project forum and collaborate.

### Chat

Find all the partners in our official chatroom.

### Code

Have a look at InvenioRDM code evolution.

### Events

InvenioRDM project events for partners

### Features

Sneak peak at the future InvenioRDM.

# InvenioRDM as a hub for local engagement

*Northwestern University's experience on the medical campus*

Try out a 1<sup>st</sup> minimum viable product with Northwestern theming: <http://bit.ly/inveniordm-at-nu>

Search public content now (try entering “data” in the search box) or log in to create your own records:

**Test Login:** gla3975

**Password:** InvenioRDM@NU\_2019



DevSci is a critical partner to understand user experience and needs, gather requirements, and test features, given their broad range of research methodologies and populations

The screenshot shows the InvenioRDM website interface. At the top, there is a purple header with the Northwestern Medicine logo and the text "Northwestern Medicine | Northwestern University | Faculty Profiles". Below the header, the page title "InvenioRDM @ Northwestern University" is displayed. A user profile dropdown menu shows the email "kristi.holmes@northwestern.edu". The main content area features a large image of a woman in a lab coat with a "Catalog your Research" button overlaid. Below the image is a search bar with the placeholder text "What are you looking for?".

Research Data Management (RDM) platforms make it possible to preserve and share, support reproducibility, and empower reuse of collaborative research activities and outputs.

**InvenioRDM** enables organizations to securely house research products and **make them discoverable, shareable, and citable** -- from publications and presentations to datasets, software, lay summaries, policy documents, and guidelines. **InvenioRDM** is being developed as part of a large, multi-organization collaboration in partnership with the European Organization for Nuclear Research (CERN), birthplace of the World Wide Web and developers of the Zenodo RDM for the European community. Together we are working to build a wide range of features that can help power biomedical research and support data sharing, innovation, knowledge dissemination, and interdisciplinary collaboration.

Our work is driven by user needs and informed by best practices and standards, including those that help define **Next Generation Repositories** as a foundation for a distributed, globally-networked infrastructure for scholarly communication, discovery, and innovation.

**Contact us.** Email us to ask questions or if you would like a test username and password to create records of your own.

**Join us.** Get updates and test new releases on the [product site](#), [email us](#) to request a demonstration or chat, take a look at our [competitive landscape analysis](#), and get involved with collaboration and discussion through our [open source project](#).

InvenioRDM is powered by the Gatter Health Sciences Library & Learning Center and supported in part by the Feinberg School of Medicine and the NIH National Center for Advancing Translational Sciences (NCATS) through grants U24TR002306 (CD2H) and UL1TR001422 (NUCATS).



# INVENIO RDM

## User Perspectives

Sara Gonzales, MLIS  
Data Librarian

# InvenioRDM Use Cases from User Testing: Fall 2019

12 testers participated in half-hour sessions in November 2019. The majority were from the Institute for Innovations in Developmental Sciences



## Testers asserted they would use InvenioRDM to:

- Search for datasets
- Store data and other files
- Deposit materials from motivations to practice open access and open science, or to comply with data sharing requirements
- Make private collections to share materials with collaborators
- Create grey literature collections of conference abstracts and posters
- Embed repository-issued DOI's (digital object identifiers) for their deposits into manuscripts and posters, enabling tracking and citation

DevSci is a critical partner to understand user experience and needs, gather requirements, and test features, given their broad range of research methodologies and populations

# Search and retrieve datasets using standards-based documentation

## Robust search enhanced by:

- Standardized forms of name (LDAP, ORCID)
- Standard subject terms (MeSH, Library of Congress Subject terms)
- Standardized citation formats
- Clear levels of access
- Standard application of licenses

The screenshot displays the InvenioRDM @ Northwestern University search interface. The search bar contains the term "diabetes", and the results show two datasets. The first dataset is "Hyperglycemia and Adverse Pregnancy Study (HAPO) dataset" by Metzger, Boyd, Lowe, Lynn P., and Dyer, Alan R. The second dataset is "Methylomics of Prenatal Gestational Diabetes Mellitus (GDM) dataset" by Hou, Lifang, Hu, Gang, Bacarelli, Andrea, Liu, Lei, and Zhang, Wei.

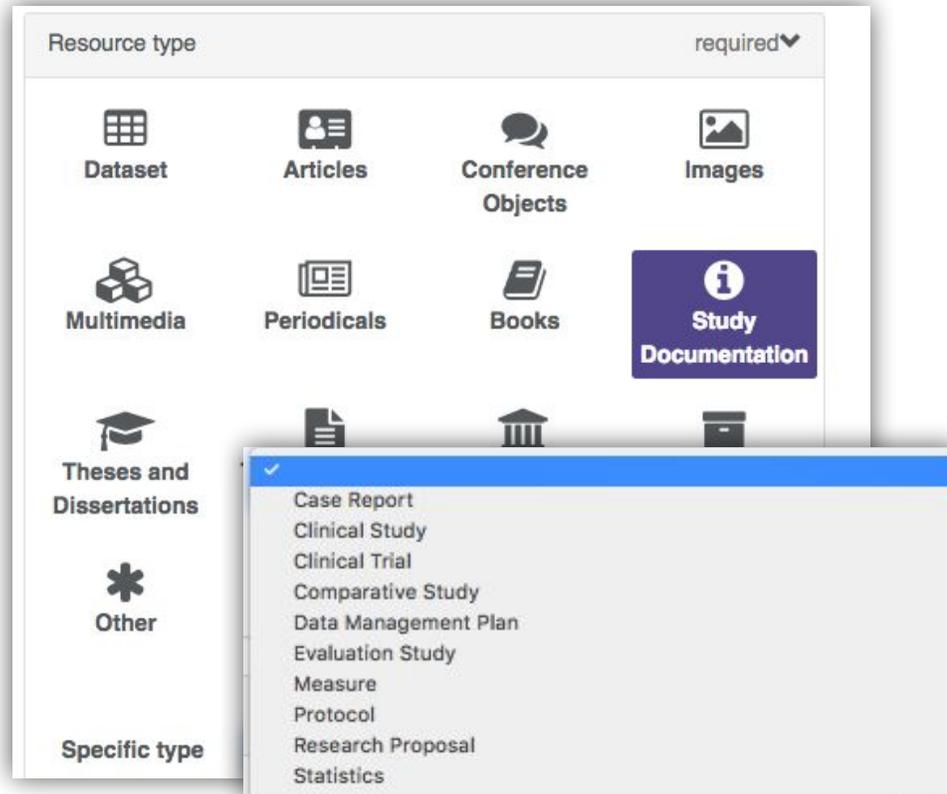
Search results for "diabetes" (found 2 results):

- Hyperglycemia and Adverse Pregnancy Study (HAPO) dataset** (Open Access)  
Metzger, Boyd; Lowe, Lynn P.; Dyer, Alan R.  
HAPO aim to assess the association between maternal glucose levels and the adverse outcome of mothers diagnosed with gestational diabetes mellitus (GDM). 25,000 pregnant women from 15 clinics across 9 countries were enrolled with 21,064 able to complete and participated in the study. Women's course of care throughout pregnancy, glucose tolerance, A1C levels and their outcomes were analyzed for any associations or independent associations to create updated guidelines for diagnosing Gestational Diabetes Mellitus (GDM) based off the data collected and adopted by WHO.
- Methylomics of Prenatal Gestational Diabetes Mellitus (GDM) dataset** (Open Access)  
Hou, Lifang; Hu, Gang; Bacarelli, Andrea; Liu, Lei; Zhang, Wei  
Building off of The Tianjin GDM Observational Study in China, this project brings the addition of genetics and lifestyle intervention to make a longitudinal study. Blood samples from both the mother and children are taken to be analyzed for DNA methylation and phenotypes. This analysis investigates the effects GDM and the risks for obesity and cardiometabolic issues that may arise in children as a result of in utero exposure. This project aims to explore the influence of lifestyle choices and intervention with children affected.

The "Subjects" search interface allows users to refine their search using standardized terms. It features two main categories: "Medical" and "Topical".

- Medical:** Users can select Medical Subject Heading (MeSH) terms. A dropdown menu is available for "Select Medical Subject Heading (MeSH) terms". Below the dropdown, it says "Medical Subject Heading (MeSH) terms." and includes a button to "+ Add another MeSH term".
- Topical:** Users can select Library of Congress (FAST) topical terms. A dropdown menu is available for "Select Library of Congress (FAST) topical terms". Below the dropdown, it says "Library of Congress (FAST) topical terms." and includes a button to "+ Add another FAST term".

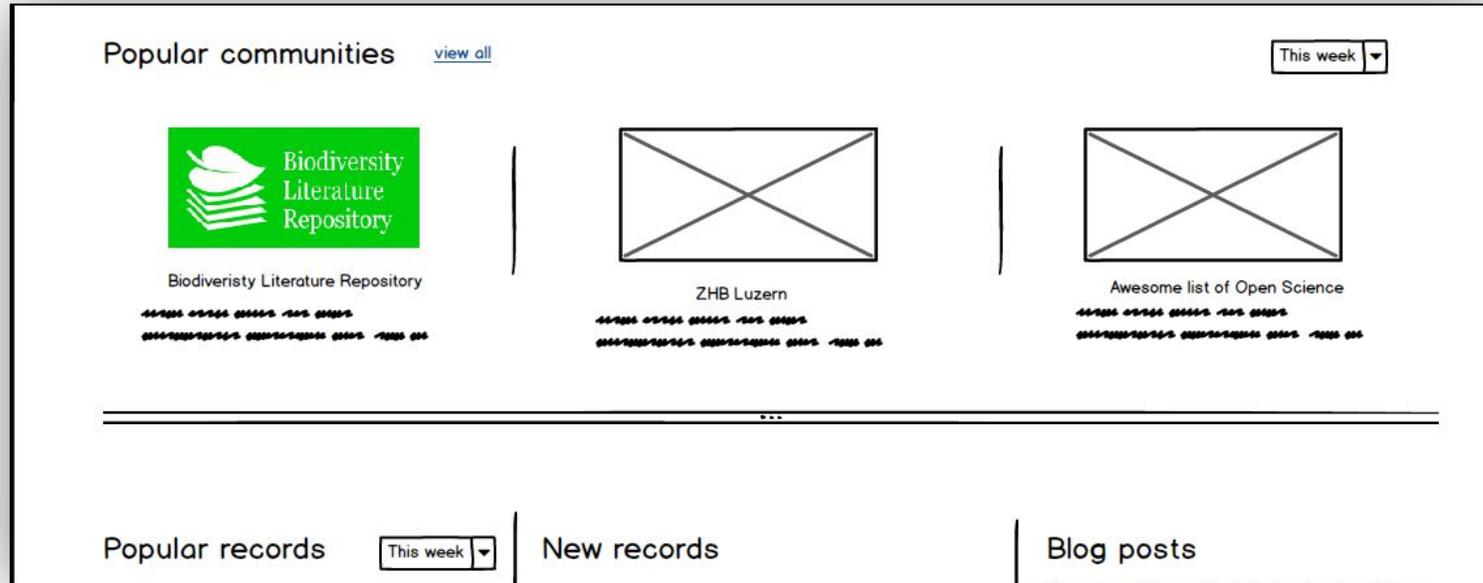
# Data management for Reproducibility and Open Access: Study-focused resource types



**InvenioRDM helps you store, manage, and if needed share your study's outputs:**

- Study-based resource types to manage a large range of assets
- **Reproducibility** is enhanced: store research proposals, datasets, code
- Be **compliant** with data sharing mandates
- **Cite** and **attribute** the work of all contributors to research
- **Reuse** deposited data or measures from other studies

# Communities and Collections: Bringing resources together

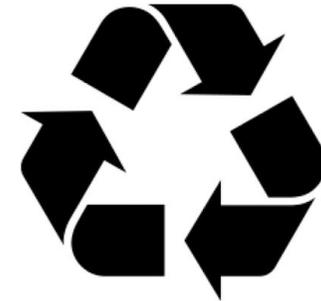
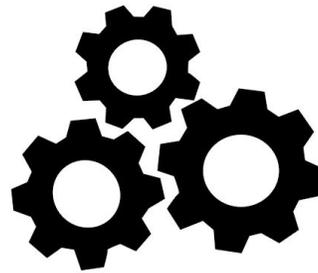


## Create a Community to:

- Create gray literature collections for your conference abstracts, posters, etc.
- Spotlight the work and resources of a department, core, project, or collaboration
- Organize resources related to a clinical study. The Community record describes the study and the records within represent the array of digital resources created through the study
- Facilitate open access, open science, and reproducibility by gathering related resources for sharing
- Optimize your file storage by using Communities and Collections to organize your deposits

# Digital Object Identifiers: Preservation and Impact

F<sub>indable</sub> A<sub>ccessible</sub> I<sub>nteroperable</sub> R<sub>eusable</sub>



InvenioRDM adheres to the FAIR data principles, which includes the repository's commitment to robust descriptive metadata from accepted standards for interoperability, and provision of licenses to make data reusable.

Data is findable and accessible if it is assigned a unique identifier on the web and can be easily be retrieved by that identifier. DOI's (digital object identifiers) assigned by InvenioRDM, allow a deposited work to be cited and retrieved online at any time. DOI's for datasets or other elements of research studies can be included in manuscripts, posters, and anywhere where the researcher would like to publicize the deposited resource. Any digital object with a DOI is **citeable**.

# Collaborators: Work with them and discover new ones

*User 1's files*



*User 2's files*



InvenioRDM will allow **private record sharing**, so researchers can:

- Share files with each other, but not anyone else in the university community or the public
- Vet materials collaboratively and privately before switching records to 'public' for open access/data sharing

InvenioRDM will have a social component, allowing researchers to:

- Follow other researchers
- Receive updates when someone they follow deposits something
- Manage requests to access files represented by a metadata-only record

## Next Generation Research Repository

guillaume.viger@northwestern.edu

Catalog your Research

Start your Research

### Catalog your Research

Drop files or folders here

— OR —

Click to select

Title

Required.

Author(s)

Required.

Description

Abstract, summary or description that will appear in search results.

Required.

License

Required. The selected license applies to all of your files. If you want to upload some of your files under different licenses, please do so in separate uploads. If you cannot find the license you're looking for, include a relevant LICENSE file in your record and choose one of the 'Other' licenses available - 'Other (Open)' or 'Other (Not Open)'.

Save Draft

Publish Record

Northwestern  
University

© 2018 Northwestern University

[Contact Northwestern University](#)

[Careers](#)

[Disclaimer](#)

[Campus Emergency Information](#)

[University Policies](#)

[Northwestern University  
Feinberg School of Medicine](#)

[Giving](#)

[Galter Health Sciences Library &  
Learning Center](#)

[Galter Health Sciences Library &  
Learning Center](#)

303 E. Chicago Avenue, Chicago, IL  
60611

312-503-8126

[Contact Galter Reference](#)



Next Generation Research Repository

[Contact Us](#)



February 9, 2002 Dataset Open Access

## Little, Clark and Green's dataset

Susan Mack

aggregate 24/365 paradigms

Fusce blandit massa scelerisque tempor lectus, ante curabitur nisi massa vitae, nascetur quam pede venenatis semper parturient venenatis fringilla,. Velit sociis taciti nostra quisque, imperdiet tincidunt tellus eros sagittis maecenas, ipsum ad lobortis vulputate amet euismod, laoreet. Pretium consequat faucibus iaculis habitasse dapibus integer, nisi conubia rhoncus enim vel, mollis etiam phasellus morbi nec, mollis suscipit. Inceptos ligula lacinia risus, etiam rhoncus morbi potenti parturient nulla.

Blandit mattis potenti feugiat sodales pede, penatibus vitae suspendisse interdum mauris congue ipsum, non habitant dapibus enim quam,. Habitant hac netus vestibulum ultrices ut, sociis augue adipiscing curae ornare dictum tristique, maecenas est fusce tempus luctus conubia, platea. Pulvinar enim consectetur auctor tempor vulputate.

Morbi parturient magna luctus pretium vulputate nulla eu, sapien inceptos orci platea, montes ad mattis elementum semper posuere, consequat nostra aliquam torquent sed aptent aliquet, nec. Eu magna at porttitor lacinia congue elit, dis id convallis sem laoreet quam.

Preview - snow-doge.jpg



Name	Size	Action
<a href="#">snow-doge.jpg</a>	131.0 kB	<a href="#">Download</a>

# INVENIO RDM

## Demo

Guillaume Viger  
Senior Developer

## Next Generation Research Repository

guillaume.viger@northwestern.edu

Catalog your Research

Start your Research

### Demo-03-15

Filename	Size	
coar-logo.png	80.5 kB	<a href="#">Download</a>
<b>Total Size</b>	80.5 kB	

[Edit](#)

Guillaume

*Demo Record with long description. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vestibulum at nibh vel turpis tempor sagittis ac a tellus. Ut purus tellus, iaculis sit amet eleifend sed, vulputate vitae est. Etiam dolor metus, imperdiet quis mi nec, faucibus tincidunt augue. Etiam sit amet lectus mi. Aliquam in felis vitae turpis volutpat dignissim at vitae magna. Fusce bibendum at mauris et porttitor. Proin velit metus, lacinia non tortor non, interdum tristique justo. Mauris vel mattis ipsum.*

**Digital Object Identifier (DOI)** <https://doi.org/10.24418/7E94-9996>

**License** Creative Commons Attribution

Northwestern  
University

© 2018 Northwestern University

[Contact Northwestern University](#)

[Careers](#)

[Disclaimer](#)

[Campus Emergency Information](#)

[University Policies](#)

[Northwestern University  
Feinberg School of Medicine](#)

[Giving](#)

[Galter Health Sciences Library &  
Learning Center](#)

[Galter Health Sciences Library &  
Learning Center](#)

303 E. Chicago Avenue, Chicago, IL  
60611

312-503-8126

[Contact Galter Reference](#)



[Next Generation Research Repository](#)

[Contact Us](#)



# Search

**File Type**

- pptx (1)

**License**

- Creative Commons Attribution (2)
- MIT License (1)

**Subjects**

- Topical (1) -
  - Diabetes--Complications (1)
- Medical (1) +

Found 3 results.

Sort by most relevant match ▾

 [Test-05:09::7:56](#)

 Guillaume

*Test long file name*

 [Test-05-14::14:18](#)

 Guillaume

*Using Nested for subjects*

 [Test-05-07::14:47](#)

 Guillaume

*Access UI test*

# The InvenioRDM Team at Northwestern

## Galter Health Sciences Library & Learning Center



*Sara Gonzales*



*Lisa O'Keefe*



*Matt Carson*



*Kristi Holmes*



*Guillaume Viger*

# Links

- **Official InvenioRDM site:** <https://inveniosoftware.org/products/rdm/>
- **Roadmap:** <https://inveniosoftware.org/products/rdm/roadmap/>
- **GitHub:** <https://github.com/inveniosoftware/invenio-app-rdm>
- **Documentation:** <https://invenio-app-rdm.readthedocs.io/en/latest>
- **Project Boards:** <https://github.com/orgs/inveniosoftware/projects>
- **RFC (Request for Comments):** <https://github.com/inveniosoftware/rfcs>

**Northwestern's Proof of Concept:** <http://bit.ly/inveniordm-at-nu>

- Test Login: gla3975
- Password: InvenioRDM@NU\_2019

**Install your own instance!** <https://github.com/inveniosoftware/invenio-app-rdm/wiki>

# INVENIO RDM

Thanks!

For questions or to let us know your needs as a potential user, please contact me at:

sara.gonzales2@northwestern.edu  
(312) 503-1209  
GitHub: @saragon02