



Software evaluation results

Yuliia Dikova, Volodymyr Kushnarenko, Preston Rodrigues,
Nadiia Huskova, Taras Petrenko, and Thomas Bönisch

Presentation plan

- Brief overview of user requirements
- Repository software and specifications
- Software overview
- User requirements test results
- Administrative requirements test results

Brief requirements overview

- Preserve research data and metadata in a FAIR way
- Be able to share publications and research results within the framework of the central repository, differentiate access rights
- Have tools for analysing, visualizing and previewing data and metadata
- Enable operations with metadata through SPARQL
- Provide API for easy access and data management



Repository Software



Software for repositories

■ InvenioRDM

Version: 9.1

Architecture:

Load balancers: HAProxy, Nginx or others.

Web servers: Nginx, Apache or others.

Application servers: UWSGI, Gunicorn or mod_wsgi.

Distributed task queue: Celery

Database: PostgreSQL, MySQL or SQLite.

Search engine: Elasticsearch (v6 and v7).

Message queue: RabbitMQ, Redis or Amazon SQS.

Cache system: Redis or Memcache.

Storage system: Local, S3, XRootD, WebDAV and more.

Dataverse

Version: 5.12.1

Architecture:

Load balancers: HAProxy, Nginx or others.

Web servers: Nginx, Apache or others.

Application servers: Payara.

Database: PostgreSQL

Search engine: Apache Solr (v8).

Cache system: Redis or Memcache.

Storage system: Local, S3.

Each of these products were installed on virtual machines in bwCloud (<https://www.bw-cloud.org/>). This cloud-based "infrastructure as a service" environment is specifically designed for research and teaching in Baden-Württemberg.

Software for repositories

■ DSpace

Version: 7

Architecture:

Load balancers: HAProxy, Nginx or others.

Web servers: Tomcat, Jetty.

Application servers: Tomcat, nodejs.

Database: PostgreSQL, Oracle.

Search engine: Apache Solr.

Message queue: RabbitMQ, Redis or Amazon SQS.

Cache system: Redis or Memcache.

Storage system: Local, S3.

CKAN

Version: 2.9.7

Architecture:

Load balancers: HAProxy, Nginx or others.

Web servers: Nginx, Apache, Tomcat.

Application servers: UWSGI, Gunicorn or mod_wsgi.

Database: PostgreSQL, MySQL or SQLite.

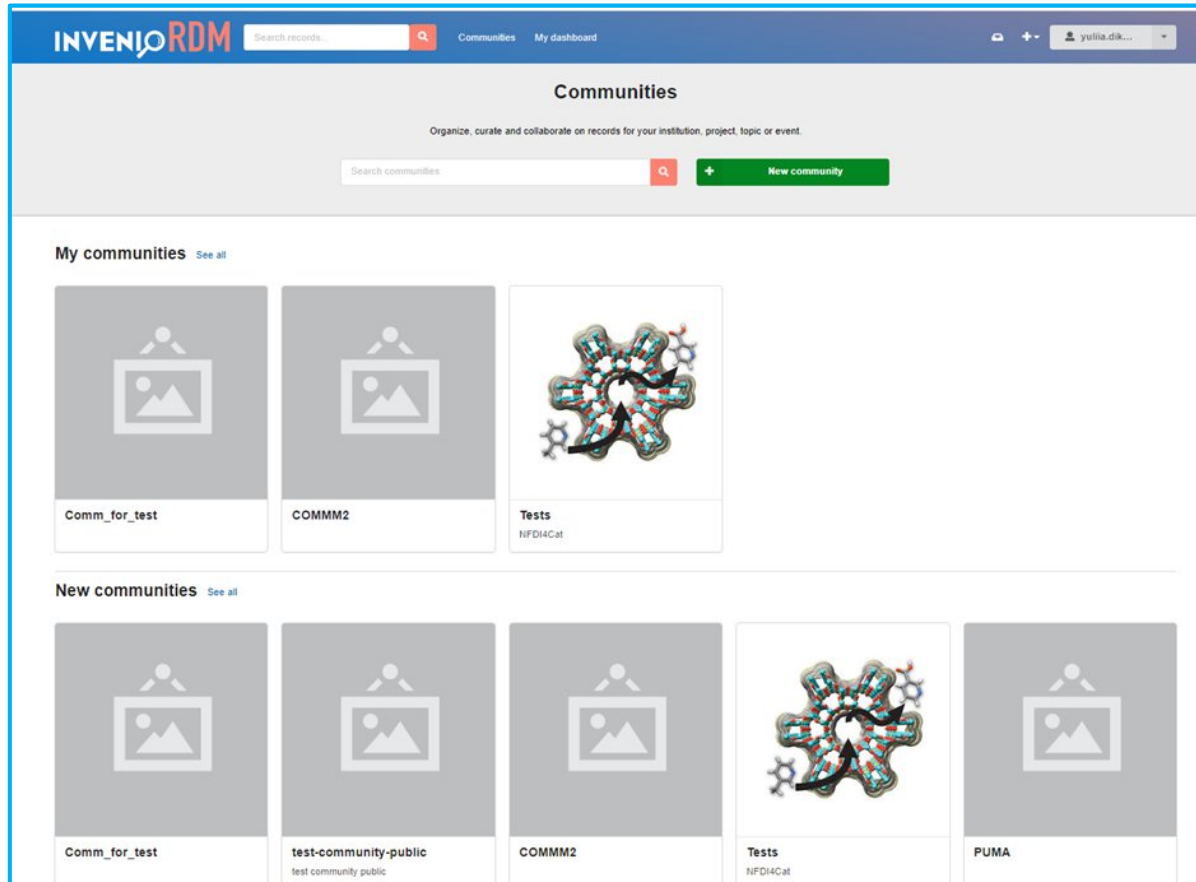
Search engine: Solr.

Message queue: RabbitMQ, Redis or Amazon SQS.

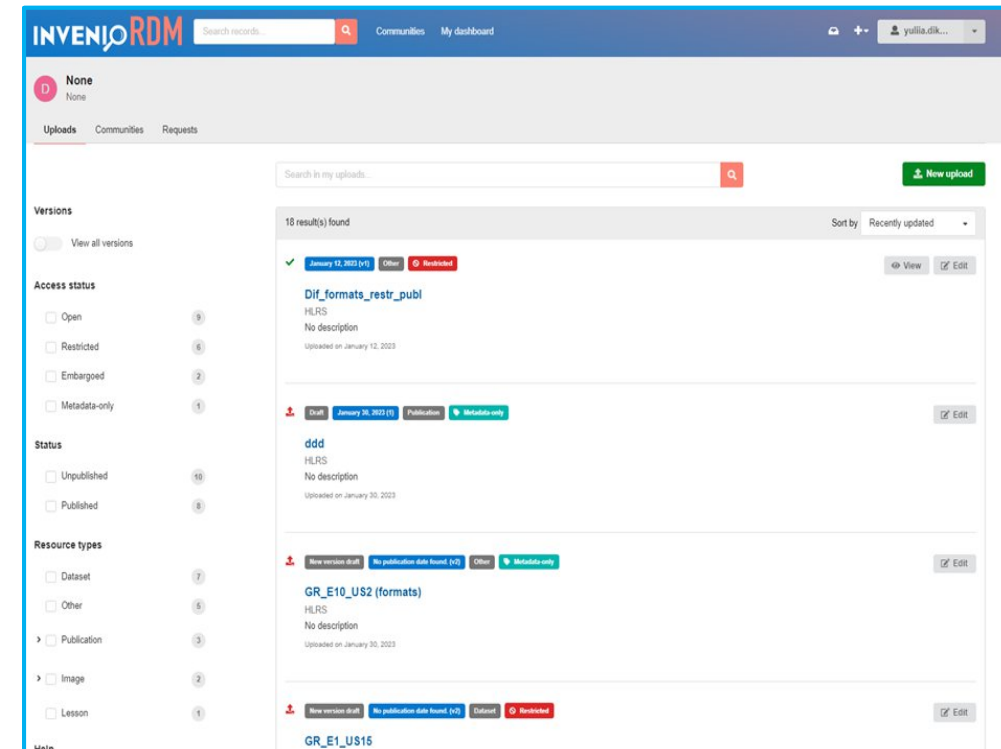
Cache system: Redis or Memcache.

Storage system: Local, S3, XRootD, WebDAV and more.

InvenioRDM overview



- Communities (similar as folder)
- Inside of each community - related uploads (as data sets)



InvenioRDM overview (continue)

The screenshot shows the InvenioRDM record creation interface. At the top, there's a navigation bar with the InvenioRDM logo, a search bar, and links to 'Communities' and 'My dashboard'. Below this, a prompt asks to 'Select the community where you want to submit your record.' with a 'Select a community' button. The main area is divided into two columns. The left column has a 'Files' section with a 'Drag and drop files' area and an 'Upload files' button. Below this is a 'Basic information' section with a 'Resource type' dropdown, a 'Title' field with an 'Add titles' button, and a 'Publication date' field with a date picker. The right column has a 'Draft' section with 'Save draft', 'Preview', and 'Publish' buttons. Below this is a 'Visibility' section with 'Full record' and 'Files only' visibility options, each with 'Public' and 'Restricted' radio buttons. The 'Public' option is selected for both. At the bottom of the right column is an 'Options' section with an 'Apply an embargo' checkbox and a note: 'Record or files protection must be restricted to apply an embargo.'

- Metadata required and non-required fields during upload

This screenshot shows the right-hand side of the InvenioRDM record creation interface. It includes a 'Creators' section with an 'Add creator' button. Below is a 'Description' section with a rich text editor and an 'Add description' button. The 'Licenses' section shows 'Creative Commons Attribution 4.0 International' as the selected license, with 'Edit' and 'Remove' buttons. Below this is a 'Recommended information' section with a 'Contributors' section (including an 'Add contributor' button), a 'Subjects' section (with a 'Suggest from' dropdown and a search bar), and a 'Languages' section (with a search bar). At the bottom is a 'Dates' section with a table for recording dates and descriptions.

Date *	Type *	Description
YYYY-MM-DD or YYYY-MM-DD/YYYY-MM-DD		

Format: DATE or DATE/DATE where DATE is YYYY or YYYY-MM or YYYY-MM-DD.

InvenioRDM overview (continue)

Version

Mostly relevant for software and dataset uploads. A semantic version string is preferred see semver.org, but any version string is accepted.

Publisher

invenio-9.1

The publisher is used to formulate the citation, so consider the prominence of the role.

Funding

Awards

+ Add award + Add custom

Alternate identifiers

Alternate identifiers

Identifier * Scheme *

+ Add identifier

Related works

Specify identifiers of related works. Supported identifiers include DOI, Handle, ARK, PURL, ISSN, ISBN, PubMed ID, PubMed Central ID, ADS Bibliographic Code, arXiv, Life Science Identifiers (LSID), EAN-13, ISTC, URNs, and URLs.

Related works

Relation * Identifier * Scheme * Resource type

Select relation... Identifier Scheme Resource type

INVENIO RDM Search records... Communities My dashboard

test-community-public Change Remove

Files

Metadata-only record Storage available 2 out of 100 files 9 bytes out of 10.00 GB

Preview	Filename	Size	Progress
<input type="checkbox"/>	test.txt md5:048f8b0d4f21d373a0e4e123217b4f8	4 bytes	100%
<input type="checkbox"/>	test2.txt md5:a0234d26205d03319b0a1107a172b	5 bytes	100%

Drag and drop files - or - Upload files

File addition, removal or modification are not allowed after you have published your upload.

Basic information

Resource type * Dataset

In review View request

Save draft Preview

Submitted for review

Visibility

Full record Public Restricted

The full record is restricted.

Restricted
The record can only be accessed by users specified in the permissions.

Options

☐ Apply an embargo
Record or files protection must be restricted to apply an embargo.

Delete

It is possible to create restricted communities. All items will be also restricted.

In public community publications can be restricted or public.

If public and restricted files are planed, separate communities are needed.

Dataverse overview

The screenshot shows the NFDI4@t repository homepage. At the top, there's a navigation bar with 'Search', 'User Guide', 'Support', and 'Yulia Dikova'. Below the header, a yellow banner says 'Welcome to NFDI4Cat repository prototype!'. The main section is titled 'NFDI4Cat collections'. There are two buttons: 'Samples' and 'test-dataverse'. Below these, a search bar and 'Advanced Search' link are present. On the left, a sidebar lists filters: 'Dataverses (5)', 'Datasets (11)', 'Files (19)', 'Dataverse Category' (Laboratory (2), Department (1), Research Project (1)), 'Metadata Source' (NFDI4Cat collections (11), SDC4Lit collections (5)), 'Publication Year' (2023 (2), 2022 (4), 2021 (5)), 'Publication Status' (Published (11), Unpublished (5), Draft (4)), 'Author Name' (Admin, Dataverse (6), Dikova, Yulia (5)), 'Subject' (Agricultural Sciences (7), Engineering (7), Computer and Information Science (3), Other (2), Mathematical Sciences (1)), and 'Deposit Date' (2023 (2), 2022 (5), 2021 (4)). The main content area shows '1 to 10 of 16 Results'. The first result is 'script_test' (Draft, Unpublished) by Dikova, Yulia, 2023, 'script_test', https://doi.org/10.82013/NFDI4Cat-proto-bwc-1035, NFDI4Cat collections, DRAFT VERSION, bash script. Other results include 'test_track_changes_published', 'Samples (Dataverse.org)', 'Picture_test', 'test-dataverse-subdataverse', 'Doc_test', and 'Pdf_test'.

The screenshot shows the NFDI4@t repository page for 'Dataverse_YL'. The page is titled 'Dataverse_YL' and has an 'Unpublished' status. Below the title, it says 'NFDI4Cat collections > test-dataverse > test-dataverse-subdataverse >'. There are buttons for 'Contact', 'Share', 'Publish', and 'Edit'. A search bar and 'Advanced Search' link are also present. On the left, a sidebar lists filters: 'Dataverses (1)', 'Datasets (3)', 'Files (7)', 'Dataverse Category' (Research Project (1)), 'Publication Status' (Unpublished (4), Draft (3)), 'Author Name' (Dikova, Yulia (3)), 'Subject' (Engineering (3)), and 'Deposit Date' (2023 (1), 2022 (2)). The main content area shows '1 to 4 of 4 Results'. The first result is 'script_test' (Draft, Unpublished) by Dikova, Yulia, 2023, 'script_test', https://doi.org/10.82013/NFDI4Cat-proto-bwc-1035, NFDI4Cat collections, DRAFT VERSION, bash script. Other results include 'Pdf_test', 'New_Test1', and 'Dataverse_YL'.

Dataverse - like folder
Dataverse can consist of several dataverses
Dataset - data

Dataverse overview (continue)

The screenshot shows the 'Dataverse_YL' form in the NFDI4@t system. The top navigation bar includes 'Search', 'User Guide', 'Support', and 'Yulia Dikova'. The breadcrumb trail reads 'NFDI4Cat collections > test-dataverse > test-dataverse-subdataverse > Dataverse_YL >'. The 'Host Dataverse' section shows 'Dataverse_YL' with a note: 'Changing the host dataverse will clear any fields you may have entered data into.' Below this, a note states '*Asterisks indicate required fields'. The 'Citation Metadata' section is expanded, showing fields for 'Title', 'Author' (Name, Affiliation, Identifier Type), 'Point of Contact' (Name, Affiliation, E-mail), 'Description' (Text, Date), 'Subject', 'Keyword' (Term, Controlled Vocabulary Name, URL), 'Related Publication' (Citation, Identifier Type, Identifier, URL), and 'Notes'.

The screenshot shows the 'Dataverse_YL' form in the NFDI4@t system, focusing on the 'Structural Metadata' sections. The top navigation bar and breadcrumb trail are the same as the previous screenshot. The 'Deposit Date' is set to '2023-03-17'. The 'Structural Metadata MODS' section is expanded, showing fields for 'Related Item' (Related Item Type, TitleInfo Authority URI, Name Authority URI, Name Type), 'TitleInfo' (TitleInfo Type, TitleInfo Value URI, TitleInfo Title, Name Value URI, Name NamePart), and 'Relationship' (Relationship Type, Relationship Sub Type, Related Object Identifier Type, Related Object Identifier Value, Related Object Sequence). The 'Structural Metadata PREMIS' section is also expanded, showing fields for 'Object Category', 'Significant Properties' (Significant Properties Type, Significant Properties Value), and 'Environment' (Environment Characteristic, Environment Purpose, Environment Note, Software Name, Software Version, Software Type, Software Other Information, Software Dependency, Hardware Name, Hardware Type, Hardware Other Information).

- More detailed description of the datasets
- Different metadata blocks (here - 3)
- It is possible to create own metadata fields and define metadata set when creating a new dataverse

Dataverse overview (continue)

Welcome to NFDI4Cat repository prototype!

Dataverse_YL
Unpublished

NFDI4Cat collections > test-dataverse > test-dataverse-subdataverse > Dataverse_YL >

Success! – This dataset has been created.

Info – This draft version needs to be published. When ready for sharing, please publish it so that others can see these changes.

test1
Draft Unpublished

Dikova, Yuliia, 2023, "test1", <https://doi.org/10.82013/NFDI4Cat-proto-bwc-1036>, NFDI4Cat collections, DRAFT VERSION

Cite Dataset [Learn about Data Citation Standards.](#)

Description [text](#)
Subject [Chemistry](#)
License/Data Use Agreement [CC0 1.0](#)

Files Metadata Terms Versions

1 File

t1.pdf
Adobe PDF - 172.9 KB
Deposited Mar 17, 2023
MD5: e03...4f1

Access Dataset [Publish Dataset](#) [Edit Dataset](#) [Contact Owner](#) [Files \(Upload\)](#) [Metadata](#) [Terms](#) [Permissions](#) [Private URL](#) [Thumbnails + Widgets](#) [Delete Dataset](#)

Copyright © 2023, NFDI4Cat

Powered by [The Dataverse Project](#) v. 5.12.1 build 1122-cf90431

Welcome to NFDI4Cat repository prototype!

Dataverse_YL
Unpublished

NFDI4Cat collections > test-dataverse > test-dataverse-subdataverse > Dataverse_YL > test1 >

Info – This draft version needs to be published. When ready for sharing, please go to the dataset page to publish it so that others can see these changes.

t1.pdf
This file is part of "test1".
Draft Unpublished

File Citation
Dikova, Yuliia, 2023, "test1", <https://doi.org/10.82013/NFDI4Cat-proto-bwc-1036>, NFDI4Cat collections, DRAFT VERSION: t1.pdf
Cite Data File [Learn about Data Citation Standards.](#)

Dataset Citation
Dikova, Yuliia, 2023, "test1", <https://doi.org/10.82013/NFDI4Cat-proto-bwc-1036>, NFDI4Cat collections, DRAFT VERSION

Preview Metadata Versions

Open in New Window

Access File [Edit File](#) [Contact Owner](#) [Share](#)

File Metrics
0 Downloads

Recent date / Accepted date

Abstract Research data management is rapidly becoming a regular concern for researchers, and institutions need to provide them with platforms to support data organisation and preparation for publication. Some institutions have adopted institutional repositories as the basis for data deposit, whereas others are experimenting with richer environments for data description, in spite of the diversity of existing workflows. This paper is a synthetic overview of current platforms that can be used for data management purposes. Adopting a pragmatic view on data management, the paper focuses on solutions that can be adopted in the long tail of science, where investments in tools and manpower are modest. First, a broad set of data management platforms is presented—some designed for institutional repositories and digital libraries—to select a short list of the more promising ones for data management. These platforms are compared considering their architecture, support for metadata, existing programming interfaces, as well as their search mechanisms and community acceptance. In this process, the stakeholders' requirements are also taken into account. The results show that there is still plenty of room for improvement, mainly regarding the specificity of data description in different domains, as well as the potential for integration of the data management platforms with existing research management tools. Nevertheless, depending on the context, some platforms can meet all or part of the stakeholders' requirements.

1 Introduction

The number of published scholarly papers is steadily increasing, and there is a growing awareness of the importance, diversity and complexity of data generated in research contexts [25]. The management of these assets is currently a concern for both researchers and institutions who have to streamline scholarly communication, while keeping record of research contributions and ensuring the correct framing of their contents [23, 18]. At the same time, academic institutions have new mandates, requiring data management activities to be carried out during the research projects, as a part of research grant contracts [14, 26]. These activities are inevitably supported by software platforms, increasing the demand for such infrastructures.

This paper presents an overview of several prominent research data management platforms that can be put in place by an institution to support part of its research data management workflow. It starts by identifying a set of well known repositories that are currently being used for other publications or data management.

Copyright © 2023, NFDI4Cat

Powered by [The Dataverse Project](#) v. 5.12.1 build 1122-cf90431

Unpublished Dataset Private URL

Use a Private URL to allow those without Dataverse accounts to access your unpublished dataset. For more information about the Private URL feature, please refer to the [User Guide](#).

This Private URL provides full read access to the dataset

<https://c8204ba4-c5d5-42c4-af03-272aa550e1e1r-bw-cloud-instance.org/privateurl.xhtml?token=e5d5e78-a4dc-485e-b583-36868c907e3a>

[Copy to Clipboard](#) [Disable Private URL](#) [Close](#)

Private URL can be used by person, who has no account in Dataverse Link is valid until owner deletes it.

Preview available only for the most common formats

Dataverse overview (continue)

The screenshot shows the NFDI4@t repository prototype interface. At the top, there's a navigation bar with the logo, search, user guide, support, and the user's name 'Yulia Dikova'. Below the navigation bar, a welcome message is displayed. The main content area shows the breadcrumb path: 'NFDI4Cat collections > test-dataverse > test-dataverse-subdataverse > Dataverse_YL > Pdf_test >'. The 'Users/Groups' section is active, showing a table with 2 users/groups. The table has columns for 'User/Group Name (Affiliation)', 'ID', 'Role', and 'Action'. The first user is 'Yulia Dikova' with ID '@Yulia.dikova' and role 'Contributor'. The second user is 'Yulia Dikova' with ID '@Yulia.dikova' and role 'Admin'. Below the table, there's a 'Roles' section listing various roles and their permissions.

Users/Groups

User/Group Name (Affiliation)	ID	Role	Action
Yulia Dikova	@Yulia.dikova	Contributor	Remove Assigned Role
Yulia Dikova	@Yulia.dikova	Admin	Role assigned at Dataverse_YL

Roles

- Admin** - A person who has all permissions for dataverses, datasets, and files, including approving requests for restricted data.
Permissions: AddDataverse, AddDataset, ViewUnpublishedDataverse, ViewUnpublishedDataset, DownloadFile, EditDataverse, EditDataset, ManageDataversePermissions, ManageDatasetPermissions, ManageFilePermissions, PublishDataverse, PublishDataset, DeleteDataverse, DeleteDatasetDraft.
- Contributor** - For datasets, a person who can edit License + Terms, and then submit them for review.
Permissions: ViewUnpublishedDataset, DownloadFile, EditDataset, DeleteDatasetDraft.
- Curator** - For datasets, a person who can edit License + Terms, edit Permissions, and publish datasets.
Permissions: AddDataverse, AddDataset, ViewUnpublishedDataverse, ViewUnpublishedDataset, DownloadFile, EditDataset, ManageDatasetPermissions, ManageFilePermissions, PublishDataset, DeleteDatasetDraft.
- File Downloader** - A person who can download a published file.
Permissions: DownloadFile.
- Member** - A person who can view both unpublished dataverses and datasets.
Permissions: ViewUnpublishedDataverse, ViewUnpublishedDataset, DownloadFile.

The screenshot shows the 'Assign Role' modal window. It has a title bar with a close button. The main content area contains a form with two fields: 'Users/Groups' (a text input with a placeholder 'Enter User/Group Name') and 'Role' (a radio button selection). The roles listed are Admin, Contributor, Curator, File Downloader, and Member. To the right of the role selection, there's a text box that says 'These are the permissions associated with the selected role.' At the bottom of the modal, there are 'Save Changes' and 'Cancel' buttons.

Assign Role

Grant permissions to users and groups by assigning them a role.

Users/Groups

Enter User/Group Name

Role

- ☐ Admin
- ☐ Contributor
- ☐ Curator
- ☐ File Downloader
- ☐ Member

These are the permissions associated with the selected role.

Save Changes Cancel

It is possible to assign different roles to work with different data or datasets

Dataverse overview (continue)

Welcome to NFDI4Cat repository prototype!

NFDI4Cat collections >

Account - NFDI4Cat collections

My Data Notifications Account Information API Token

Notification settings

- test1 was created in Dataverse_YL. To learn more about what you can do with a dataset, check out the User Guide. Mar 17, 2023, 1:52:53 PM CET
- script_test was created in Dataverse_YL. To learn more about what you can do with a dataset, check out the User Guide. Jan 30, 2023, 1:02:12 PM CET
- Picture_test was published in test-dataverse-subdataverse. Dec 12, 2022, 1:41:18 PM CET
- You have been granted the Admin role for test-dataverse-subdataverse. Dec 12, 2022, 1:41:18 PM CET
- Picture_test was created in test-dataverse-subdataverse. To learn more about what you can do with a dataset, check out the User Guide. Dec 12, 2022, 1:35:02 PM CET
- Doc_test was created in test-dataverse-subdataverse. To learn more about what you can do with a dataset, check out the User Guide. Dec 12, 2022, 1:19:44 PM CET
- Pdf_test was created in Dataverse_YL. To learn more about what you can do with a dataset, check out the User Guide. Dec 12, 2022, 1:08:59 PM CET
- Dataset New_Test1 has one or more tabular files that completed the tabular ingest process and are available in archival formats. Dec 7, 2022, 11:59:19 AM CET
- New_Test1 was created in Dataverse_YL. To learn more about what you can do with a dataset, check out the User Guide. Dec 7, 2022, 11:59:16 AM CET
- Dataverse_YL was created in test-dataverse-subdataverse. To learn more about what you can do with your dataverse, check out the User Guide. Dec 7, 2022, 11:58:42 AM CET
- You have been granted the Admin role for test-dataverse-subdataverse. Dec 2, 2022, 3:09:45 PM CET
- Welcome to NFDI4Cat collections! Get started by adding or finding data. Have questions? Check out the User Guide. Want to test out Dataverse features? Use our Demo Site. Also, check for your welcome email to verify your address. Dec 1, 2022, 3:22:20 PM CET

Copyright © 2023, NFDI4Cat. Powered by The Dataverse Project v. 5.12.1 build 1122-c90431

Notification for users activity

NFDI4Cat collections > test-dataverse > test-dataverse-subdataverse > Dataverse_YL > New_Test1 >

Info - This draft version needs to be published. When ready for sharing, please go to the dataset page to publish it so that others can see these changes.

test4-more-lines.tab

This file is part of "New_Test1".

Draft Unpublished

File Citation

Dikova, Yulia, 2022, "New_Test1", <https://doi.org/10.82013/NFDI4Cat-proto-bwc-1026>, NFDI4Cat collections, DRAFT VERSION, test4-more-lines.tab [fileName], UNF:6-yrLAVb5MCDRQdWmRVw== [fileUNF]

Cite Data File - Learn about Data Citation Standards.

Dataset Citation

Dikova, Yulia, 2022, "New_Test1", <https://doi.org/10.82013/NFDI4Cat-proto-bwc-1026>, NFDI4Cat collections, DRAFT VERSION, UNF:6-4OpRvJNR3/d+JMa2wXDUdQ== [fileUNF]

Cite Dataset - Learn about Data Citation Standards.

Access File - Edit File - Contact Owner - Share

File Metrics 0 Downloads

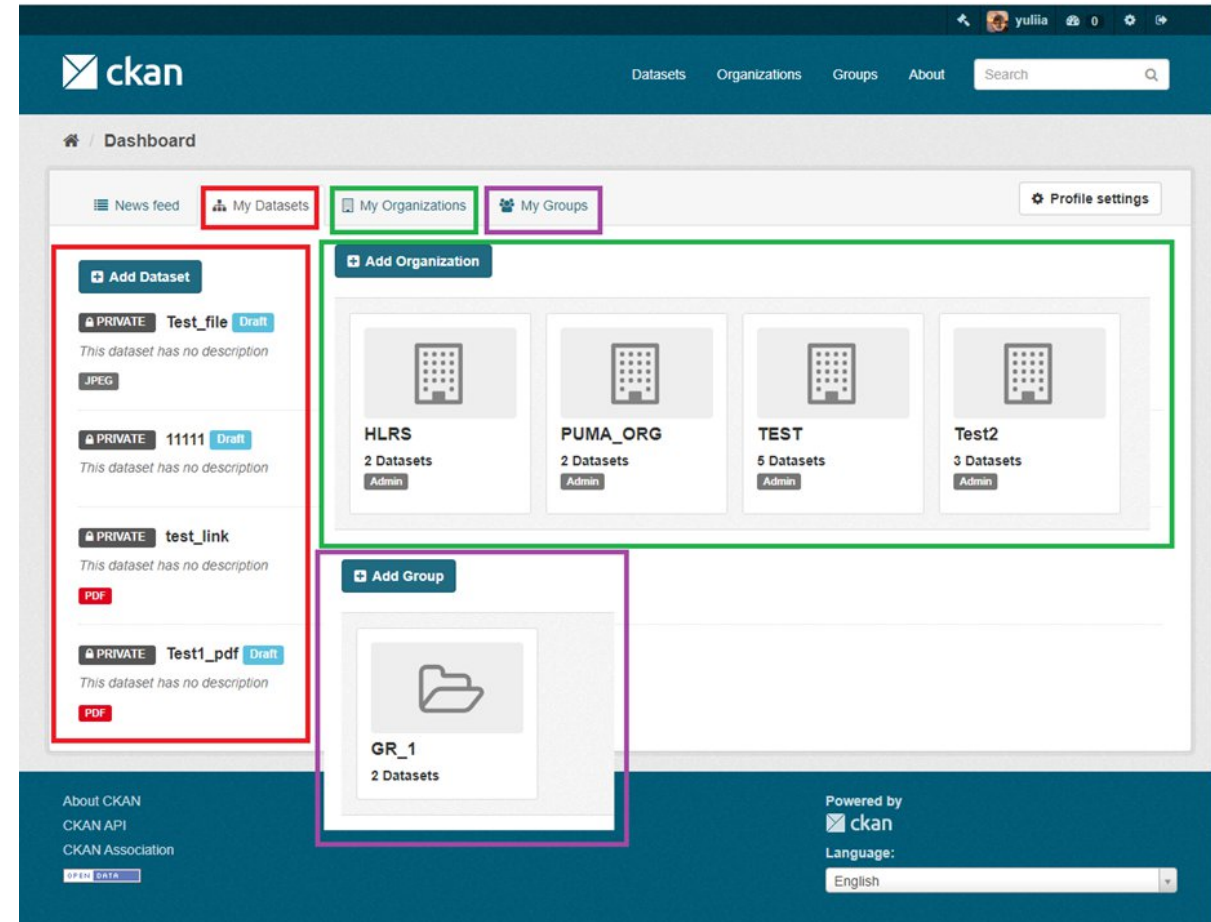
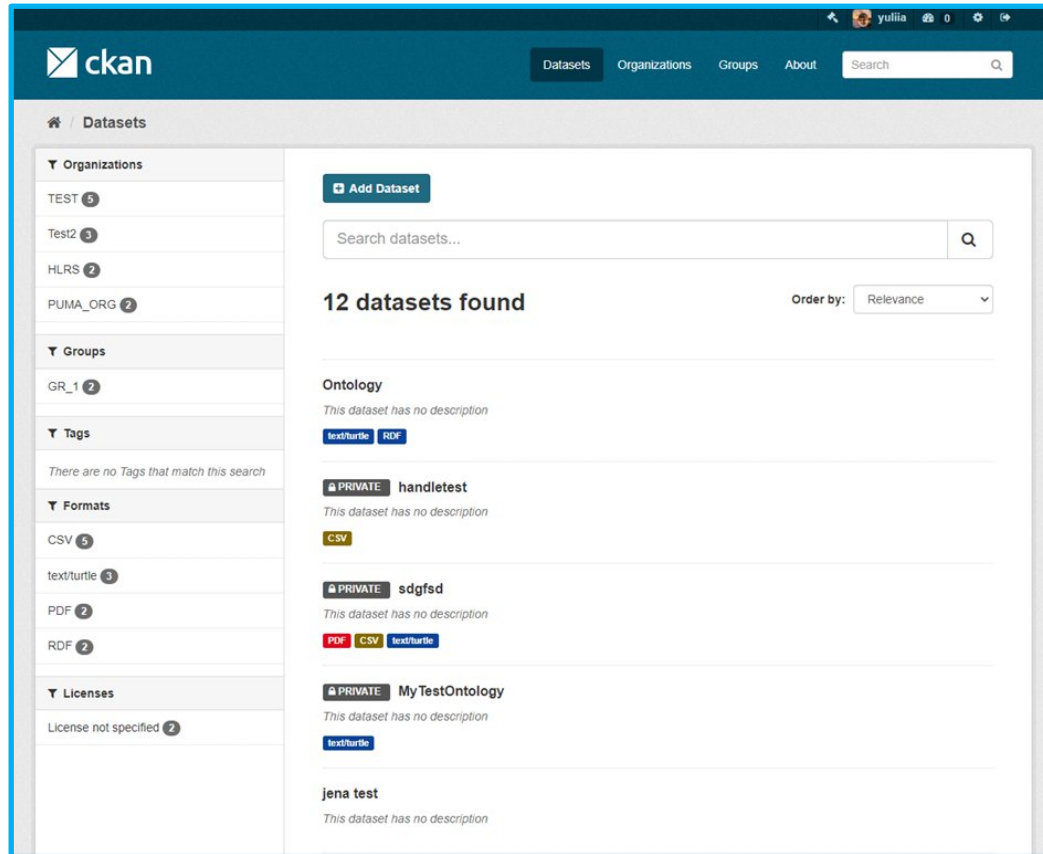
Preview Metadata Versions

Open in New Window

	найменування товару	одиниця виміру	ціна за 1 одиницю, грн	D	E	F	Номер чека	Постачальник
1	Туалетний папір Обухів, 60 м	рупок	7				1.0	Listex
2	Болога серветка Arieli, 80 шт.	упаковка	150				2.0	TOP Pack
3	Шампунь для волосся Dove 250 ml	штучка	80				3.0	Listex
4	Мило Dove, 125 gr	штучка	35				4.0	Listex
5	Антиперспірант Rexona, 100 ml	штучка	60				5.0	Listex
6	Ватний диск Sunny, 250 шт.	упаковка	20				6.0	Listex
7	Пральний порошок Ariel, 24 капс.	упаковка	100				7.0	Listex
8	Ополіскувач Silan 800 ml	штучка	80				8.0	KosmoSweet
9	Зубна паста Colgate Fresh 100 ml	штучка	35				9.0	Eva
10	Зубна щітка Oral-B	штучка	35				10.0	Eva
11							11.0	Eva
12	Постачальник						12.0	Eva
13	Thekangooroo						13.0	Thekangooroo
14	TOP Pack						14.0	Thekangooroo
15	KosmoSweet						15.0	Thekangooroo
16	Listex						16.0	Thekangooroo
17	Eva						17.0	Thekangooroo
18							18.0	TOP Pack
19							19.0	TOP Pack
20							20.0	TOP Pack
21	найменування товару	одиниця виміру	ціна за 1 одиницю, грн					
22	Туалетний папір Обухів, 60 м	рупок	7					
23	Болога серветка Arieli, 80 шт.	упаковка	150					
24	Шампунь для волосся Dove 250 ml	штучка	80					
25								

Preview is possible only for most common formats. For other some additional extension are needed

CKAN overview



Organization - folder
Dataset - data or dataset

each group can includes organization, but additional extensions are needed

CKAN overview (continue)

The screenshot shows the 'Create Dataset' form in the CKAN interface. The form is divided into two steps: '1 Create dataset' (active) and '2 Add data'. The 'Create dataset' step includes fields for Title, Description, Tags, License, Organization, and Visibility. The 'Add data' step is partially visible and includes fields for Source, Version, Author, Author Email, Maintainer, Maintainer Email, and Custom Fields. The form is titled 'What are datasets?' and includes a brief explanation of what a CKAN Dataset is.

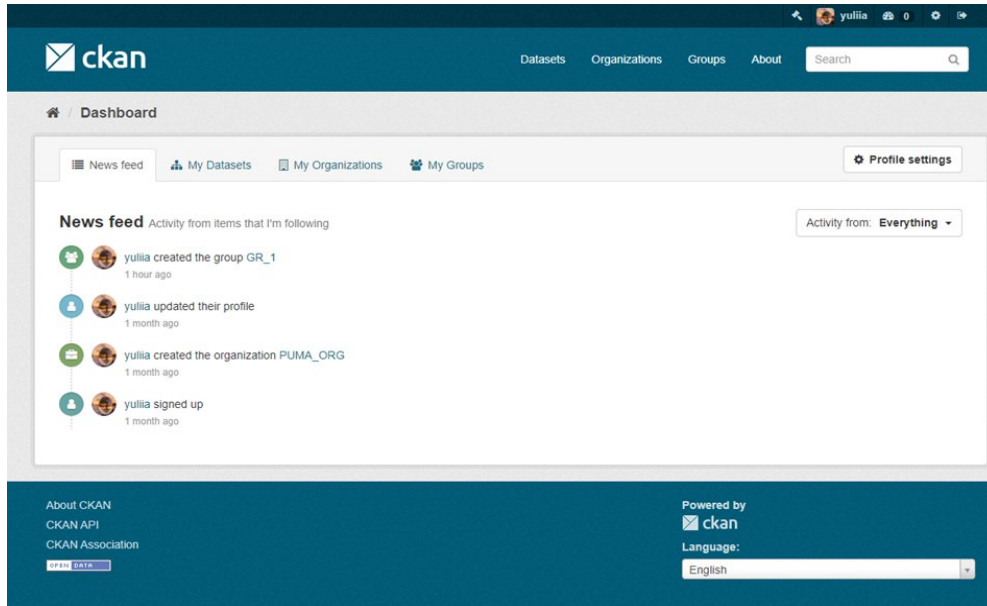
The screenshot shows the 'Add Data' form in the CKAN interface. The form is titled 'What's a resource?' and includes a brief explanation of what a resource is. The form includes fields for Name, Description, Format, and Custom Fields. The 'Add Data' step is active, and the 'Create dataset' step is partially visible. The form is titled 'What's a resource?' and includes a brief explanation of what a resource is.

The screenshot shows the 'Add New Resource' form in the CKAN interface. The form is titled 'What's a resource?' and includes a brief explanation of what a resource is. The form includes fields for Name, Description, Format, and Custom Fields. The 'Add New Resource' step is active, and the 'Add data' step is partially visible. The form is titled 'What's a resource?' and includes a brief explanation of what a resource is.

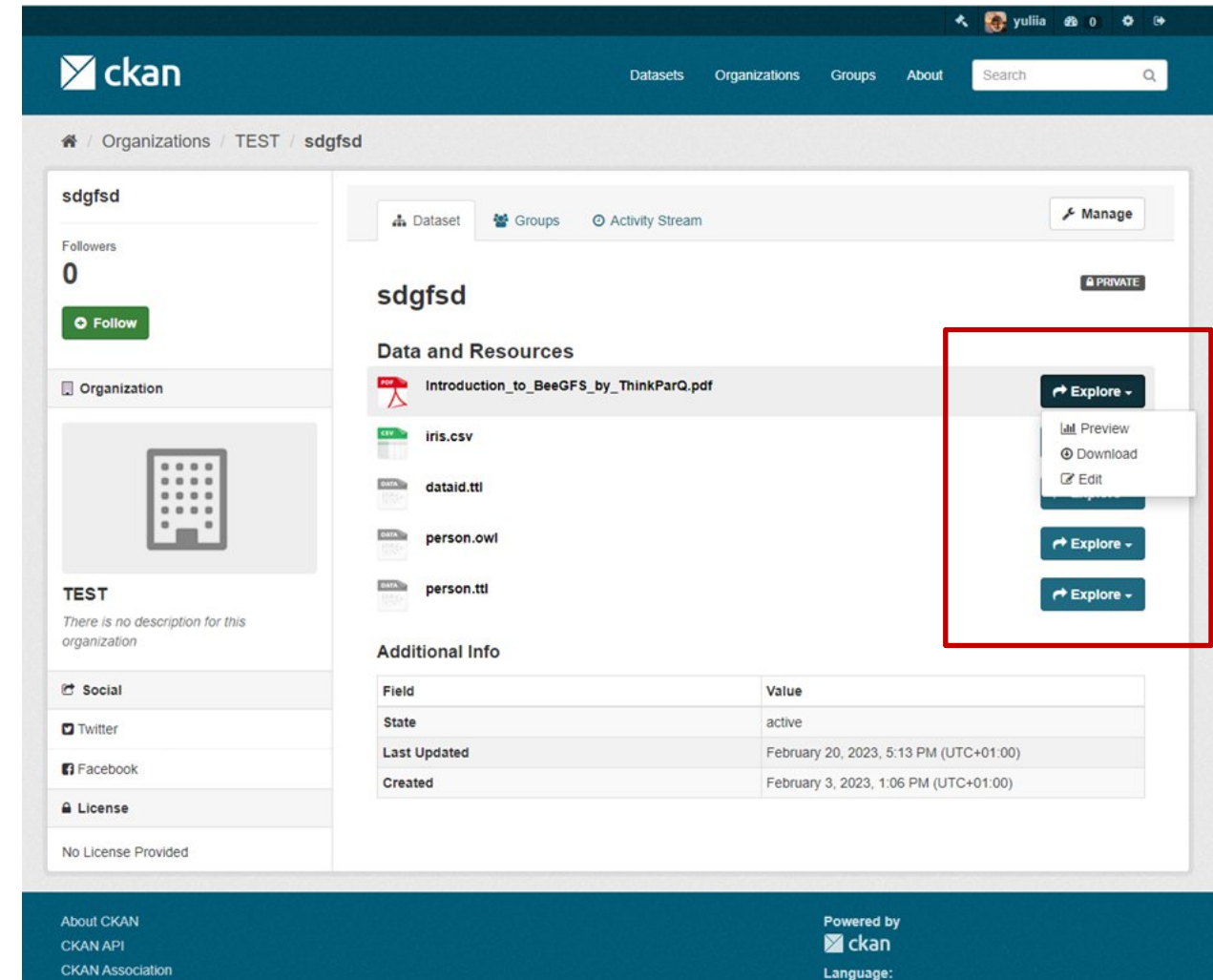
When adding a new dataset, a fairly small amount of metadata is specified. If necessary, the user can add any number of metadata fields as custom fields

user can upload data or put link

CKAN overview (continue)



It is possible to see users activity



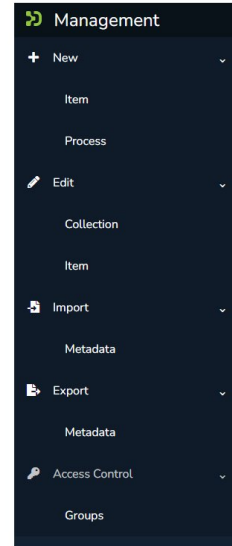
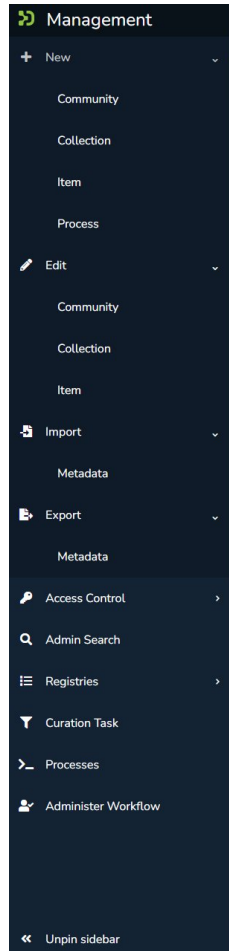
For most common formats there is preview of data

DSpace overview

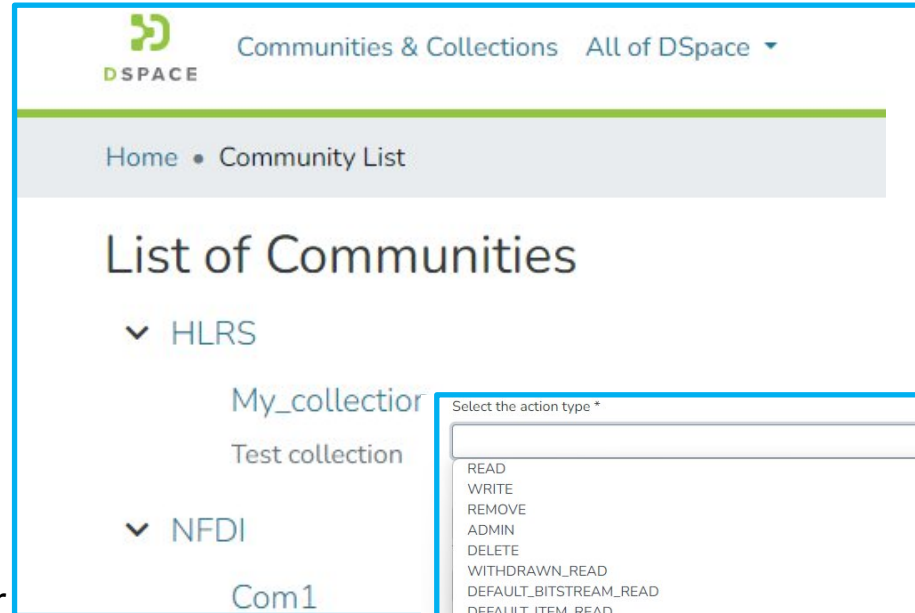
The image displays two screenshots of the DSpace 7 interface. The left screenshot shows the main DSpace 7 landing page, which includes a navigation sidebar on the left with options like 'New', 'Community', 'Collection', 'Item', 'Process', 'Edit', 'Import', 'Export', 'Access Control', 'Admin Search', 'Registries', 'Curation Task', 'Processes', and 'Administer Workflow'. The main content area features the DSpace 7 logo and a list of communities, including 'HLRS' and 'NFDI'. The right screenshot shows the 'Edit Submission' form for a collection, which includes fields for 'Author', 'Title', 'Date of Issue', 'Publisher', and 'Citation'. The form also has a 'Drop files to attach them to the item, or browse' section and a 'Discard' button at the bottom.

- Interface, upload
- Different possibilities for users
- Only main admin can create collections
- “Workflow” - admins of the collection must approve every publication - similar working process as in a library -> usage as Institutional Repository

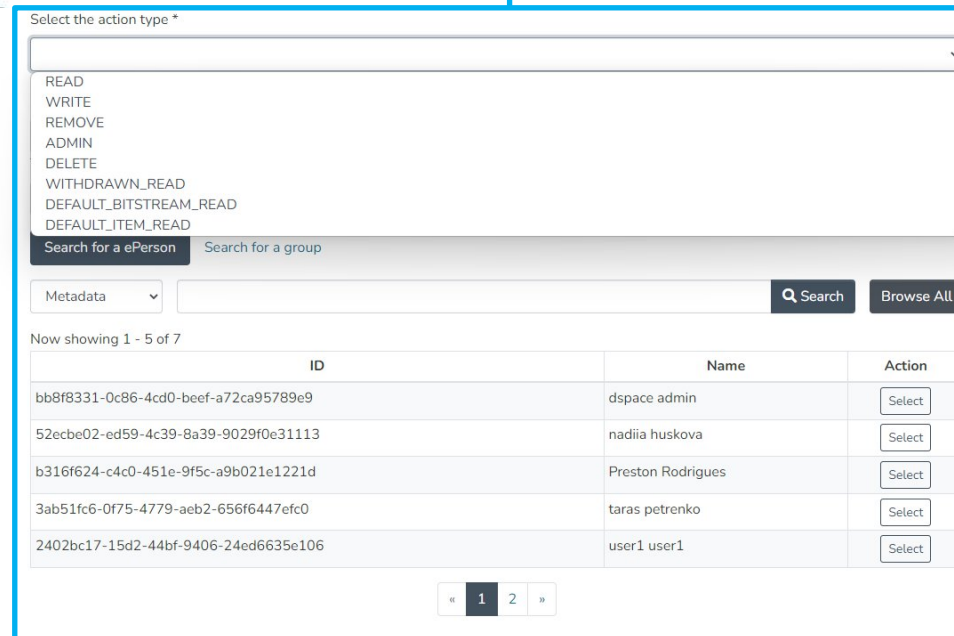
DSpace overview (continue)



Sidebar for user



Only main admin can create community and collections. Then according access rights, user can work with them



Admin can manage role of user

Sidebar for main admin

Analysis of user stories

Total amount of US - 60 (521 points)

- High - 49
- Middle - 10
- Low - 1

Priority (points)

- High 10
- Middle 3
- Low 1

Requirement can be...

- 1 = fulfilled,
- 0,5 = partially fulfilled,
- 0 = not fulfilled

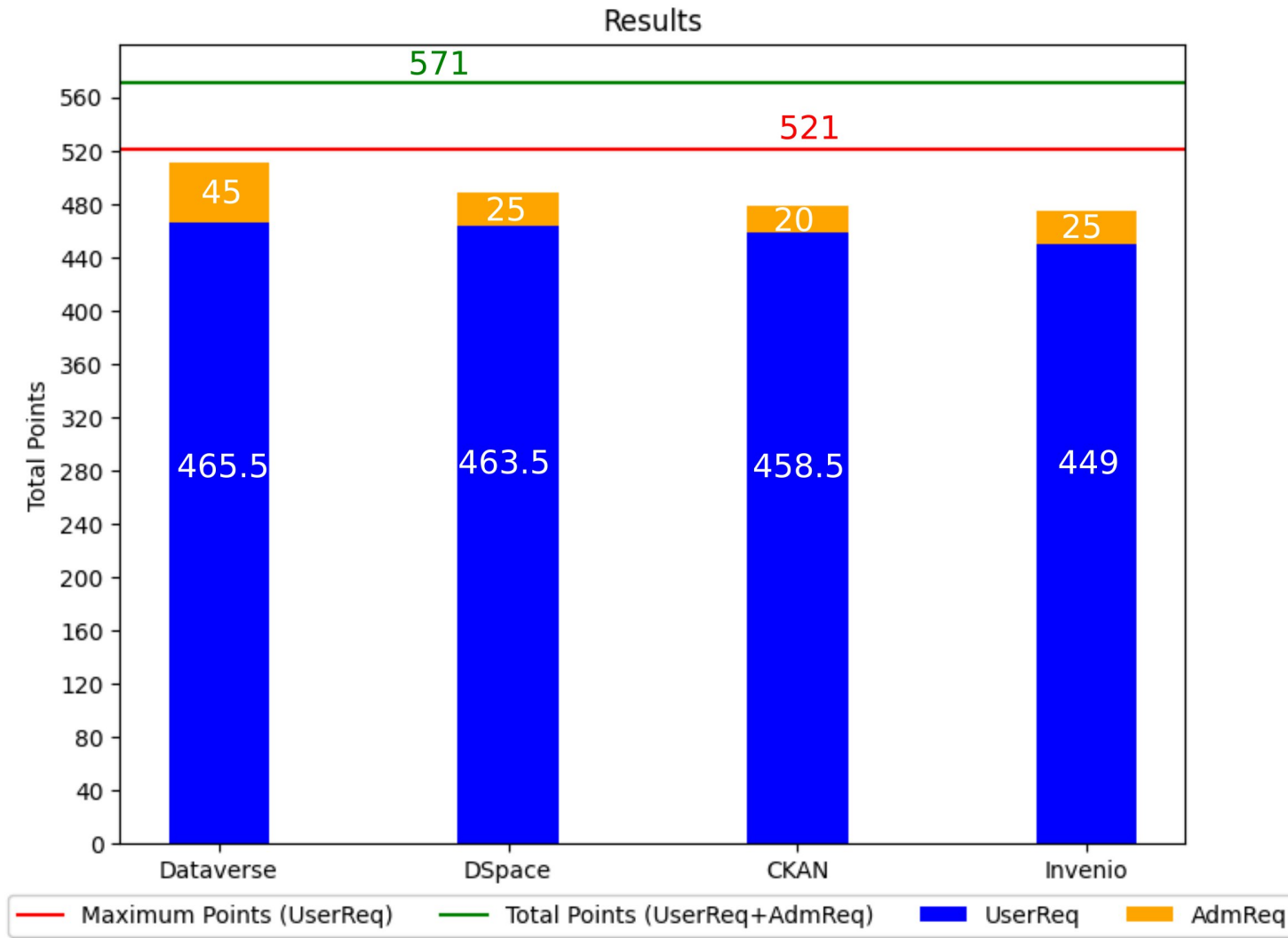
Priority	DSpace	Dataverse	CKAN	InvenioRDM
High	40	43	39	40
middle	7	6	7	5
low	1	1	1	1
High	8	3	9	6
middle	1	1	1	2
low	0	0	0	0
High	1	3	2	3
middle	2	3	2	3
low	0	0	0	0
Summ	463,5	465,5	458,5	449

	fully realizable
	partly realizable
	Non realizable

Additional requirements (from administrators)

	Priority	Dataverse	Result	Invenio	Result	CKAN	Result	DSpace	Result	Comments
Installation procedure	10	1	10	1	10	1	10	1	10	1 – simple Installation, Guides available, active Community 0.5 – additional technical support needed 0 – no guides, difficult installation process
Tools for enhancing functionality not needed	10	0.5	5	1	10	0.5	5	0.5	5	0 – need a lot of additional extensions 0.5 – extensions needed partly 1 – without any additional extensions
Role Based access	10	1	10	0	0	0	0	0.5	5	0 - strict set of roles 0.5 – different roles 1 – different role for the same user
Private link to data/data sets	10	1	10	0	0	0	0	0	0	0 – impossible share data/datasets with private link 1 - possible share data/datasets with private link
Support for metadata schemes	10	1	10	0.5	5	0.5	5	0.5	5	0 – there is no fields for primary metadata 0.5 – limited set of primary metadata 1 – flexible metadata schemes
Final evaluation administrative			45		25		20		25	

Common results



According user requirements Dataverse has the most points (465,5)

According administrative requirements Dataverse also has the most points (45)

Common result shows, that Dataverse (total points = 510,5) is the most suitable candidate for the central repository

Main problems

- It is not possible to cover 100% of requirements (user stories) - additional implementation will be needed
- Repositories do not provide an ability to export all or selected publications directly from the database. Scripts needed.
- Additional extensions required to ensure analytical, statistical and visualization capabilities
- Limited functionality with metadata (Linked Data usually is not supported)
- Repositories do not provide interfaces for working with SPARQL
- Previewers usually support only the most common data formats (image formats, pdf, csv, etc...) – for specific formats extensions are needed
- Complex implementation of communication between the central repository and existing local systems and software products – integrating the already available pilot systems of our partners with the central repository into the new created NFDI4Cat infrastructure will be the next challenge.

Thank you

Contact: Yuliia Dikova (HLRS), Email: yuliia.dikova@hlrs.de