TIB – Leibniz Data Manager Installation Manual

(Last update: June 2022) LDM version 2.4.1

Table of Content

Table of Content	Page
1.Description1.1 About1.2 Impact on Scientific Data Management1.3 Application of the Data Manager	3 3 3 4
2. How to install2.1 Dependencies2.2 LDM Data Manager2.3 Jupyter Notebooks visualization plugin	4 4 5 9
 3. Customizing The Data Manager 3.1 Changing configuration settings (ckan.ini) 3.2 Changing "Site Tile" and "Site Logo" 3.3 Changing The LDM Styles and Colors (CSS) 3.4 "Legal Notices" Configuration 	12 12 13 13 15
4. Useful documentation links 5. LDM Demo 6. Resource Visualization Examples	19 19 20

1. Description

1.1 About

The LDM Data Manager has been developed to support the aspect of better re-usability of research data.

The prototype supports the management and access to heterogeneous research data publications and assists researchers in the selection of relevant data sets for their respective disciplines.

The prototype currently offers the following functions for the visualization of research data:

- Supports data collections and publications with different formats
- Different views on the same data set (2D and 3D support)
- Visualization of Auto CAD files
- Jupyter Notebook(s) for demonstrating live code
- RDF Description of data collections

The file specific viewers were implemented using CKAN (Comprehensive Knowledge Archive Network) plug-ins to render existing viewers for the datasets included in the CKAN instance.

1.2 Impact on Scientific Data Management

In the research data landscape, there is a high demand for a sustainable and meaningful handling of a main product of scientific work - research data.

Since digital data production has increased rapidly in recent years and an end to this growth is not foreseeable, the availability of these growing volumes of data must be ensured not only for current research but also for future generations. The LDM Data Manager was developed to provide scientists with a **tool to improve the usability of research data**.

The LDM Data Manager provides a data management system that makes it possible to **check** the contents of research data sets for their potential application to the respective domain - without having to download them beforehand.

Therefore, the Data Manager enables the visualization of different research data formats and thus supports the **'screening' of data sets** for their potential benefits. As a visualization and management tool, TIB CKAN can be implemented on top of classical research data repositories, which often focus on the (long-term) archiving and publication of research data.

1.3 Application of the Data Manager

As an open source tool, the LDM Data Manager offers

- developers,
- scientists and
- data curators in public and academic research as well as in industry a wide range of possibilities for expanding and connecting established and developing research data management systems, such as local and discipline-specific research data repositories.

As the German National Library of Science and Technology, we advise universities, research institutions and industry on the use and implementation of the LDM Data Manager. Furthermore, we continue to develop and enhance the functionality of the system in view of the constantly growing number of scientific file formats.

2. How to install

2.1 Dependencies

The LDM Data Manager is composed by different services:

- CKAN: Open-source DMS (data management system) for powering data hubs and data portals
- PostgreSQL: Open-source object-relational database management system
- SOLR: Open-source enterprise search platform
- Postfix: Open-source mail transfer agent (MTA) that routes and delivers electronic mail
- DataPusher:

In order to avoid having to manually install these dependencies, the distribution package comes with dockerized instances of these dependencies, making it easy to get started with LDM Data Manager.

The distribution package also contains a docker-compose file, where Docker

Docker-compose 1.18.0+

2.2 LDM Data Manager

Docker is necessary so that the LDM Data Manager can be used. To be able to install it, the user must download the docker packets from Docker official website (https://docs.docker.com/install/), and afterwards follow the installation steps established in the packets.

In case the user is going to clone the LDM Data Manager code into your system is also needed to have GIT installed following the instructions in the GIT official website (https://gitforwindows.org).

Note: Before starting the steps below please check that you don't have conflicting docker containers from other projects on your machine. You can check them by firing this command in terminal:

docker container ls -aq

If you find some of them then please stop them by using the command:

docker container stop \$(docker container ls -aq)

And further after stopping them remove them by the following command:

docker container rm \$(docker container ls -aq)

After doing the steps above continue with the steps below:

To be able use CKAN with all the services it is necessary to follow these steps:

Step 1:

First, it is necessary to change the url of the site in the .env file (which is located in the "docker" folder), in the line 21 of the file the user needs to put the url of the server to get access to the CKAN through the server.

```
CKAN_SITE_URL: "<URL Server>:5000"
```

Additionally, make sure the following ports are free in the server:

- Port 5000 for data pusher.
- Port 8000 for Jupyter Notebooks.

Notes:

In Windows you can use "netstat" to check whether a port is available.

Use the netstat -anp | find "port number" command to find whether a port is occupied by an another process or not. If it is occupied by an another process, it will show the process id of that process. For example, in "Command Promt" run:

•

In case the port is free the command shows no results, otherwise if is occupied will show some details like above:

netstat -ano | find ":5000"

netstat -ano | find ":5000"

TCP 0.0.0.0:5000 0.0.0.0:0 LISTENING 14284 TCP [::]:5000 [::]:0 LISTENING 14284

To check the listening ports and applications on **Linux**:

- 1. Open a terminal application i.e. shell prompt.
- 2. Run any one of the following command on Linux to see open ports:

sudo Isof -i -P -n | grep LISTEN sudo netstat -tulpn | grep LISTEN sudo Isof -i:5000 ## see a specific port such as 5000 ## sudo nmap -sTU -O IP-address-Here

3. For the latest version of Linux use the ss command. For example, ss -tulw

Step 2:

Note:

If you have knowledge about HTML and CSS and are very sure you are not going to break the source code you can now see the chapter "3. Customizing the Data Manager" before continue for avoiding rework later.

Build the container base of CKAN inside the docker folder.

Inside "Command Prompt" or "Terminal" depending of your Operating System, go to the "docker" folder using "cd" command. For example:

cd docker

and then run the command:

docker-compose up -d --build

and then you need to restart the containers to take ckan configuration:

docker restart solr db jupyternotebook

Important!:

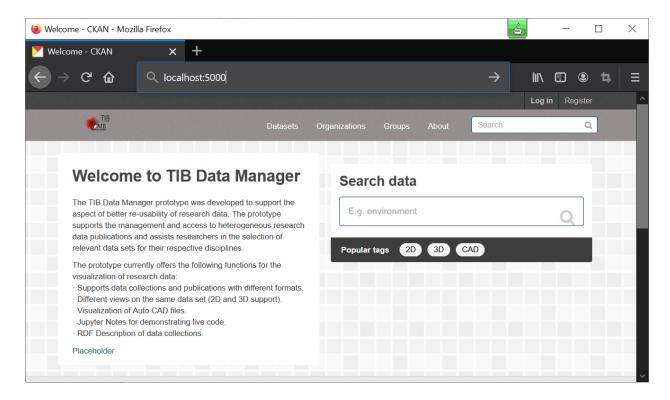
After launching the command above CKAN will start running configuration and settings procedures. Please wait until all the process is complete. You can check in docker logs if ckan container is already running (INFO [ckan.cli.server] Running CKAN on http://0.0.0.0:5000) or just accessing to localhost:5000 and getting the LDM home page as result.

For running the Data Manager in further occasions you can run:

docker-compose up ckan

Note: Please be patient the dependencies take some time to download.

To open the data manager you must open a browser and enter localhost:5000 or the url that was used in the port 5000.



Step 3:

In case the examples datasets are not showing in the system you can fix the problem running the following command:

docker exec -it ckan /reload_database.sh

Otherwise, if you need to clean the database and erase all the datasets and TIB examples you can run:

docker exec -it ckan /clean_database.sh

Svsadmin user:

In case you are using LDM with installed examples you can access the system as system administrator user with credentials: "admin" (without quotes) as user and "admin" (without quotes) as password.

Otherwise, in case you cleaned the database you can create a new user through the registration option of the tool and promote the user as sysadmin running the command:

docker exec -it ckan ckan -c /etc/ckan/default/ckan.ini sysadmin add your-user-name replacing your-user-name with the username you created before.

Troubleshooting:

- "No such file or directory" or other unexpected errors:

Docker outputs all build steps when creating an image based on a Dockerfile. On **Windows systems** is possible to the "exec user process caused "no such file or directory"" issue occurred when executing a shell script or many others unexpected errors.

The error message is misleading in terms of a wrong file path or path reference. In our case, the issue occurred due to a Windows-style file ending.

We created the Dockerfile on a Windows machine. Saving the Dockerfile used the default Windows file format. This caused the Docker build to fail on a Linux machine.

We fix this converting the file format to UNIX style using dos2unix:

dos2unix your-file.sh

You can run the dos2unix command on any Linux system. If you don't have access to a Linux system, you may use the Git Bash for Windows which comes with a dos2unix.exe.

"Git for Windows" (https://gitforwindows.org/) provides a BASH emulation used to run Git from the command line.
*NIX users should feel right at home, as the BASH emulation behaves just like the "git" command in LINUX and UNIX environments.

If you already installed GIT following this manual you are able to open the Start menu by clicking on the Windows icon and typing "Git Bash" into the search bar. The icon for Git Bash and the words "Git Bash Desktop App" will appear. Click on the icon or the words "Git Bash Desktop App" to open Git Bash. Be sure to navigate to the project folder and run the following command:

find . -type f -print0 | xargs -0 dos2unix

This will recursively find all files inside current directory and call for these files dos2unix command

- SERVER ERROR in CKAN's home page:

This could happen if any problem during the installation causes the Databases not building properly. Try going into a shell inside the ckan container and rebuild the Databases executing the following commands: docker run --rm -it --entrypoint=/bin/bash ckan

cd /usr/lib/ckan/default/src/ckan

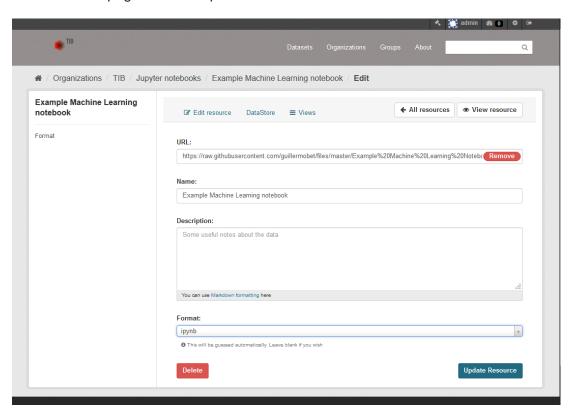
ckan -c /etc/ckan/default/ckan.ini db init

And create admin user with:

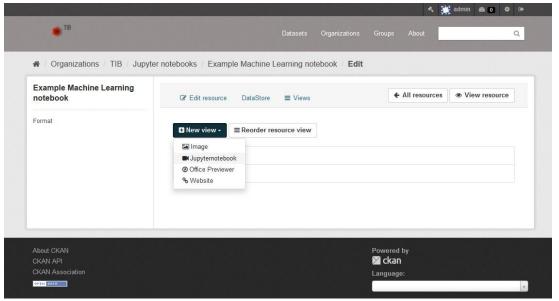
ckan-paster --plugin=ckan sysadmin -c \$CKAN_CONFIG/ckan.ini add admin email=admin@email.com password=admin

2.3 Jupyter Notebooks visualization plugin.

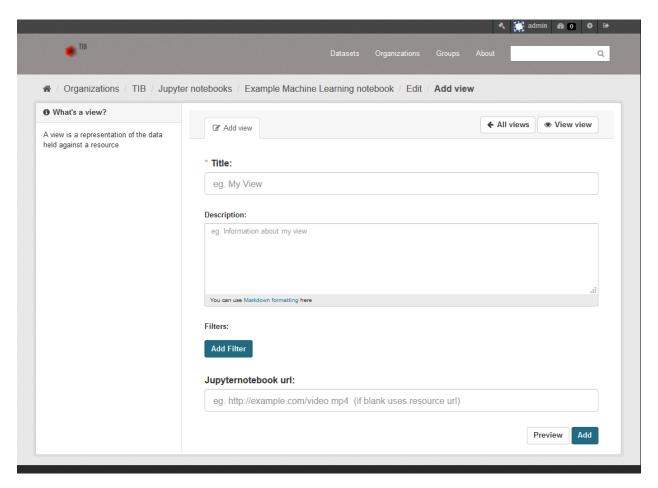
You can view Jupyter Notebooks files using the visualization plugin installed with LDM. For that make sure your notebooks files are described like "ipybn" as format in your resource's administration page. For example:



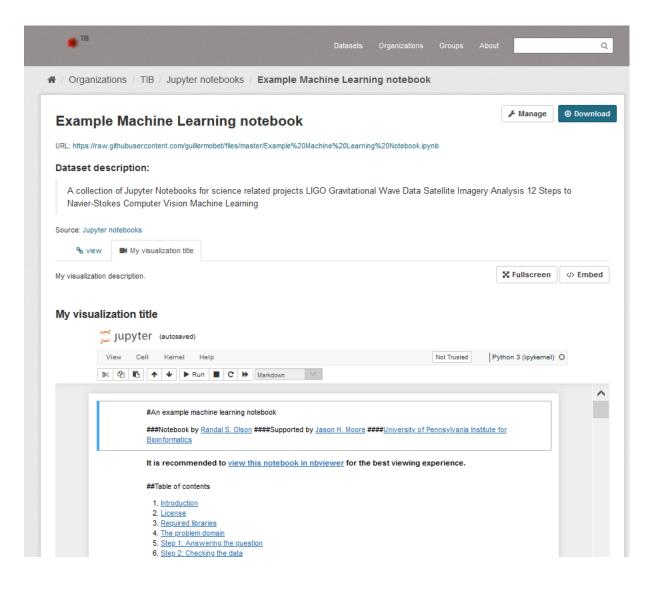
Then you can add the "JupyterNotebook" visualization in the "Views" tab clicking in the "New view" button:



Next add a title and description for your visualization. And in case the visualization should use a different notebook file than the resource itself you can set an URL in the "Jupyternotebook url" field.



Then click "Add" and your view will be ready to be shown in your resource landing page. For example:



3. Customizing The Data Manager

3.1 Changing configuration settings (ckan.ini)

Within the installation of LDM a Docker container is created with the name "ckan" containing an instance of CKAN. Inside the container a configuration file is generated:

/etc/ckan/default/ckan.ini. This configuration file stores very useful CKAN's settings that could be changed, customizing the instance behavior (See the documentation for more information: http://docs.ckan.org/en/latest/maintaining/configuration.html#ckan-configuration-file).

One simple manner for accessing and changing ckan.ini file is the following:

a. In a terminal (Linux) or command line (Windows) run the bellowing command making a copy of the file ckan.ini to your current directory.

docker cp ckan:/etc/ckan/default/ckan.ini ./ckan.ini

After the user should be able to see ckan.ini listed in the directory by running the command:

(on Windows)

dir

(on Linux)

1s

 Open and browse/edit your local copy of ckan.ini using any text editor of your preference available on your Operating System. Save the changes.

Make a copy!:

It's highly recommended making a copy of your local file **ckan.ini**.

By doing that the user should be able to restore the system in case of errors or mistakes during the edition of the settings restoring the original file into the container by running the steps "c" and "d" with the original ckan.ini as target. Without this precaution a simple spelling error could cause the system to irreversible failure in lack of the knowledge for detecting and fixing the mistake.

 Replace the configuration file inside the ckan container with your customized local version by running the command:

docker cp ckan.ini ckan:/etc/ckan/default/ckan.ini

d. Restart ckan container running the command:

docker restart ckan

3.2 Changing "Site Title" and "Site Logo"

Site Title: This is the title of this CKAN instance. It appears in several places throughout CKAN, for example in the label of the tab's browser as "Welcome – LDM" in homepage. **Site Logo:** This is the logo that appears in the header of all the CKAN instance templates.

To be able to change the site title and/or the logo of the header of the Data Manager, the user should follow the next steps:

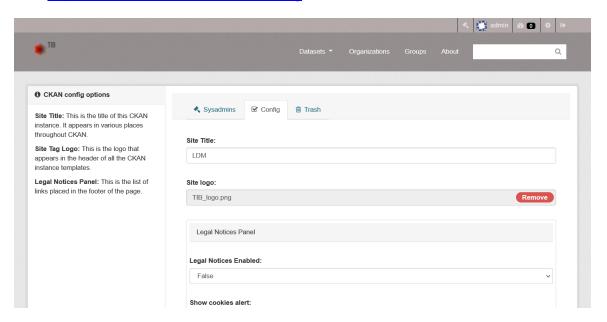
a. Login as sysadmin. Click in the "Log in" label at the upper-right corner of the LDM, and log-in introducing user and password.

Note:

If sysadmin user credentials are unknown consult the section "2.2 LDM Data Manager – step 3" of this manual for more information.

b. Click in the header's logo to access homepage and add '/ckan-admin/config'. For example: in case LDM instance is running in http://localhost:5000/ckan-admin/config.

URL http://localhost:5000/ckan-admin/config.



c. Using this form change the text in "Site Title", load a new image for "Site logo" and click on "Update Config" button at the bottom of the page.

3.3 Changing The LDM Styles and Colors (CSS)

To be able to change the visual aspect of LDM the user could access and modify the file TIBtheme.css where styles and colors are defined for the main HTML elements in the website.

Following the next steps:

a. In a terminal (Linux) or command line (Windows) run the bellowing command making a copy of the file ckan.ini to your current directory.

```
docker cp ckan:/usr/lib/ckan/default/src/ckanext-
TIBtheme/ckanext/TIBtheme/public/assets/TIBtheme.css ./TIBtheme.css
```

After the user should be able to see TIBtheme.css listed in the directory by running the command:

(on Windows)

dir

(on Linux)

1s

b. Open and browse/edit your local copy of **TIBtheme.css** using any text editor of your preference available on your Operating System. Save the changes.

Make a copy!:

It's highly recommended making a copy of your local file ckan.ini.

By doing that the user should be able to restore the original CSS file in case of errors or mistakes during the edition of the settings restoring the original file into the container by running the steps "c" and "d" with the original TIBtheme.css as target. Without this precaution a simple spelling error could cause the system to irreversible failure in lack of the knowledge for detecting and fixing the mistake.

c. Replace the CSS file inside the ckan container with your customized local version by running the command:

```
docker cp TIBtheme.css ckan:/usr/lib/ckan/default/src/ckanext-
TIBtheme/ckanext/TIBtheme/public/assets/TIBtheme.css
```

d. Restart **ckan** container running the command:

docker restart ckan

Example: How to change header and footer colors:

- a. Edit TIBtheme.css (step b) and replace it in the container (step c) following the steps above and making this changes:
 - 1. Find this line:

```
.masthead {
 background-color: #989292 !important;
}
```

2. Change it to:

```
.masthead {
   background-color: blue !important;
}
```

3. Find this line:

.site-footer, body {

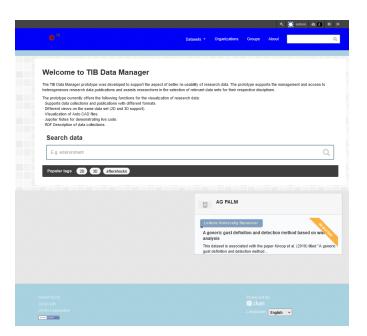
background: lightblue;

```
.site-footer,
body {
background: #b9afaf url("../../base/images/bg.png");
}

4. Change it to:
```

The result is:

}



About HTML colors:

In this example we changed the hexadecimal representation of the color '#989292' with the color name 'blue' and the hexadecimal representation of the color '#b9afaf' and an image url '("../.../base/images/bg.png")' with the color name 'lightblue'. This situation is regular en CSS definitions. For more information about the color options available visit https://www.w3.org/wiki/CSS/Properties/color/keywords.

3.4 "Legal Notices" Configuration

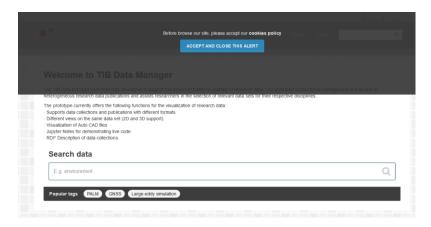
The LDM has implemented Legal Documents required for any website to be published. This feature is disabled by default, but the user could use and adapt these documents to their needs following this instructions.

When this feature is enabled LDM shows in the footer section the "Legal Notices" links:



The user is able to enable/disable the whole feature or, having it enabled, show or hide each one of the links.

As part of "Legal Notices" feature LDM also offers a "Cookies Alert" pop-up window that also can be enabled or disabled by configuration.



Finally, with some HTML knowledge the content of each Legal Document could be changed and adapted to the user's needs.

"Legal Notices" configuration:

The user should follow the next steps:

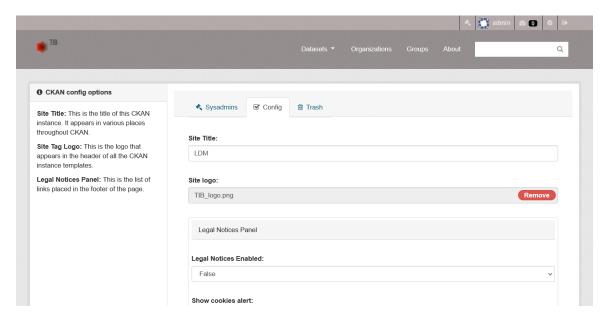
a. Login as sysadmin. Click in the "Log in" label at the upper-right corner of the LDM, and log-in introducing user and password.

Note:

If sysadmin user credentials are unknown consult the section "2.2 LDM Data Manager – step 3" of this manual for more information.

b. Click in the header's logo to access homepage and add '/ckan-admin/config'. For example: in case LDM instance is running in http://localhost:5000/ckan-admin/config.

URL http://localhost:5000/ckan-admin/config.



c. Using the options in this form the user can enable/disable the "Legal Notices" features, enable/disable the "Cookie alert" feature and/or show or hide each of the links in the footer. Each field is self-explanatory by its label. The changes are saved by clicking on "Update Config" button at the bottom of the page.

Changing the content of "Legal Notices" documents:

LDM offers the possibility of adapt each of the following "Legal Notices" documents:

- Special conditions TIB LDM
- Data Privacy Statement
- Imprint
- Accessibility Statement
- Cookies alert (pop-up window)

In case of "Special conditions TIB LDM" also the label could be changed by configuration using the form descripted in the previous section of this manual.

For changing the content of each document the user should access and edit the following files:

- For "Special conditions TIB LDM":
 - special_conditions_LDM_en.html (content in English)
 - special_conditions_LDM_de.html (content in German)
- For "Data Privacy Statement":
 - data_privacy_en.html (content in English)
 - data_privacy_de.html (content in German)
- For "Imprint":
 - imprint en.html (content in English)
 - imprint_de.html (content in German)
- For "Accessibility Statement":
 - accessibility_statement_en.html (content in English)
 - accessibility_statement_de.html (content in German)
- For "Cookies alert":
 - cookies_alert.html (content in English and German)

Example changing "Special conditions TIB LDM" content in English:

The following steps are an example. The user is able to change all mentioned documents following the same steps but changing the name of the file "special_conditions_LDM_en.html" with the correspondent file name in the list above.

a. In a terminal (Linux) or command line (Windows) run the bellowing command making a copy of the file ckan.ini to your current directory.

docker cp ckan:/usr/lib/ckan/default/src/ckanextTIBtheme/ckanext/TIBtheme/templates/home/snippets/special_conditions_LDM_en.ht
ml ./special_conditions_LDM_en.html

After the user should be able to see special_conditions_LDM_en.html listed in the directory by running the command: (on Windows)

dir

(on Linux)

ls

b. Open and browse/edit your local copy of special_conditions_LDM_en.html using any text editor of your preference available on your Operating System. Save the changes.

Make a copy!:

It's highly recommended making a copy of your local file **special_conditions_LDM_en.html**. By doing that the user should be able to restore the original HTML file in case of errors or mistakes during the edition of the settings restoring the original file into the container by running the steps "c" and "d" with the original file as target.

c. Replace the special_conditions_LDM_en.html file inside the ckan container with your customized local version by running the command:

docker cp special_conditions_LDM_en.html ckan:/usr/lib/ckan/default/src/ckanext-TIBtheme/ckanext/TIBtheme/templates/home/snippets/special_conditions_LDM_en.html

d. Restart **ckan** container running the command:

docker restart ckan

The content of the document is changed in the website.

4. <u>Useful documentation links:</u>

- CKAN User guide: http://docs.ckan.org/en/2.9/user-guide.html
- CKAN Sysadmin guide: http://docs.ckan.org/en/2.9/sysadmin-guide.html
- CKAN Mantainer's guide: http://docs.ckan.org/en/2.9/maintaining/index.html
- CKAN API guide: http://docs.ckan.org/en/2.9/api/index.html

5. <u>LDM Demo</u>

The LDM's demo presents the Leibniz Data Manager (LDM) and illustrates how Semantic Web technologies and FAIR principles empower research data management. The demonstration shows how various digital objects are created and puts in perspective the crucial role of metadata in efficient and effective management and analysis of research data management.

The <u>demonstration</u> comprises:

 A short paper is published in the proceeding of the Extended Semantic Web Conference (ESWC 2022) in the track of poster and demos. It describes the main features of the Leibniz Data Manager,



A poster summarizing the project,



· A short video describing the motivation of this project, and



A video showing the LDM features.



6. Resource Visualization Examples

Both in the installed version (<u>Link</u>) and in the online version (<u>Link</u>) are Datasets examples available showing visualizations of resources (csv, json, xlsl, pdf, videos, images, etc.) in different formats and representations.

