
V1.0
March 2016
Table of Contents

1 Introduction............................................................................................................................. 4
  1.1 Purpose of the Document .................................................................................................. 4
  1.2 Reference Documents ....................................................................................................... 4
  1.3 Terminology ....................................................................................................................... 4
2 Approach .................................................................................................................................. 5
3 Main User Functions of the Portal .......................................................................................... 5
  3.1 Portal Home Page ................................................................................................................ 6
    3.1.1 How to browse through the Editorial Content of the Portal ........................................ 7
    3.1.2 How to view the “Latest News” ..................................................................................... 13
    3.1.3 How to subscribe to the EDP Newsletter ........................................................................ 14
    3.1.4 How to view “Tweets” .................................................................................................... 15
    3.1.5 How to switch to another User Language ............................................................... 16
    3.1.6 How to Search for Datasets by Category ...................................................................... 17
    3.1.7 How to Search for Datasets by Keyword ...................................................................... 19
  3.2 Datasets (CKAN) ................................................................................................................. 20
    3.2.1 Entering the Datasets-View .......................................................................................... 21
    3.2.2 How to filter datasets by using “Faceted Search” ..................................................... 21
    3.2.3 Additional information about a dataset .......................................................................... 23
    3.2.4 How to filter datasets by geographical area ............................................................... 23
    3.2.5 How to display dataset distributions .......................................................................... 26
    3.2.6 How to view licensing information .............................................................................. 27
    3.2.7 How to switch to another user language ...................................................................... 27
    3.2.8 How to use Multi-Language Search ........................................................................... 27
    3.2.9 How to browse by catalogues ....................................................................................... 28
  3.3 Visualization of Geo-Spatial Data ...................................................................................... 29
    3.3.1 How to visualize geo-spatial data from a dataset resource ........................................... 29
  3.4 Graphical Data Visualisation Tool ..................................................................................... 33
    3.4.1 How to visualize graphical data from a dataset resource ............................................. 33
  3.5 Help Desk .............................................................................................................................. 38
3.5.1 How to contact the Portal’s Help Desk .......................................................... 38
3.6 Metadata Quality Assurance (MQA) ................................................................. 39
  3.6.1 The Global Dashboard View ..................................................................... 39
  3.6.2 Catalogue details view ............................................................................. 41
3.7 SPARQL Manager ......................................................................................... 44
  3.7.1 SPARQL Search ....................................................................................... 44
  3.7.2 SPARQL Assistant ................................................................................... 45
3.8 Metadata Transformer Service ....................................................................... 47
  3.8.1 The User Dashboard ............................................................................... 47
  3.8.2 Creating Repositories ............................................................................. 48
  3.8.3 Creating Harvesters ............................................................................... 49
  3.8.4 The Harvester Overview ........................................................................ 50
  3.8.5 View or Monitor a Running Task ............................................................ 51
1 Introduction

1.1 Purpose of the Document

The main purpose of this document is to present a User Manual for the main user functionalities of the Portal Version 1, launched in production in February 2016. This document consists of an update of the user manual for the Portal Beta Version [1].

1.2 Reference Documents

<table>
<thead>
<tr>
<th>Id</th>
<th>Reference</th>
<th>Title</th>
<th>Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1]</td>
<td>EDP_S1_MAN</td>
<td>EDP_S1_MAN_Portal-Beta-Version-User-Manual_v1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

*Table 1-1: Reference Documents*

1.3 Terminology

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>API</td>
<td>Application Programmer Interface</td>
</tr>
<tr>
<td>CKAN</td>
<td>Comprehensive Knowledge Archive Network - a web-based open source data management system for the storage and distribution of meta datasets.</td>
</tr>
<tr>
<td>CSV</td>
<td>Comma separated values</td>
</tr>
<tr>
<td>DCAT-AP</td>
<td>DCAT Application Profile - Metadata specification based on the Data Catalogue vocabulary (DCAT)</td>
</tr>
<tr>
<td>EDP</td>
<td>European Data Portal</td>
</tr>
<tr>
<td>FME</td>
<td>Feature Manipulation Engine</td>
</tr>
<tr>
<td>GUI</td>
<td>Graphical User Interface</td>
</tr>
<tr>
<td>HTTP</td>
<td>Hypertext Transfer Protocol</td>
</tr>
<tr>
<td>JSON</td>
<td>JavaScript Object Notation (a lightweight data-interchange format)</td>
</tr>
<tr>
<td>maps.app</td>
<td>Geo-spatial data visualization application</td>
</tr>
<tr>
<td>MQA</td>
<td>Metadata Quality Assistant</td>
</tr>
<tr>
<td>RDF</td>
<td>Resource Description Framework</td>
</tr>
<tr>
<td>SPARQL</td>
<td>Query language for linked data (RDF)</td>
</tr>
<tr>
<td>SSL</td>
<td>Secure Socket Layer</td>
</tr>
<tr>
<td>URL</td>
<td>Uniform Resource Locator</td>
</tr>
<tr>
<td>XML</td>
<td>Extensible Markup Language</td>
</tr>
</tbody>
</table>

*Table 1-2: Abbreviations and Acronyms*
2 Approach

The approach used for this User Manual was based on the identification of the main user functions of the Portal and the description of each function from the user’s perspective in terms of “How to...”.

Each main function documentation consists of a screen snapshot, the steps required to execute the function and optionally a screenshot with the results.

3 Main User Functions of the Portal

This section describes all of the main user functions supported by the Portal Version 1.

The table 3-1 below lists the available functions by module.

<table>
<thead>
<tr>
<th>Module Name</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Portal HomePage</td>
<td>- How to browse through the editorial content</td>
</tr>
<tr>
<td></td>
<td>- How to access the “eLearning” modules</td>
</tr>
<tr>
<td></td>
<td>- How to get to the “Training Companion”</td>
</tr>
<tr>
<td></td>
<td>- How to view the “Latest News”</td>
</tr>
<tr>
<td></td>
<td>- How to subscribe to the EDP Newsletter</td>
</tr>
<tr>
<td></td>
<td>- How to view “Tweets”</td>
</tr>
<tr>
<td></td>
<td>- How to switch to another user language</td>
</tr>
<tr>
<td></td>
<td>- How to search for Datasets by Category</td>
</tr>
<tr>
<td></td>
<td>- How to search for Datasets by Keyword</td>
</tr>
<tr>
<td>2 Datasets (CKAN)</td>
<td>Entering the Datasets-View</td>
</tr>
<tr>
<td></td>
<td>How to filter datasets by using “Faceted Search”</td>
</tr>
<tr>
<td></td>
<td>How to filter datasets by geographical area</td>
</tr>
<tr>
<td></td>
<td>How to display dataset distributions</td>
</tr>
<tr>
<td></td>
<td>How to view licensing information</td>
</tr>
<tr>
<td></td>
<td>How to switch to another user language</td>
</tr>
<tr>
<td></td>
<td>How to use Multi-Language Search</td>
</tr>
<tr>
<td>3 Visualization of Geo-Spatial Data (map.apps)</td>
<td>How to visualize geo-spatial data from a dataset resource</td>
</tr>
<tr>
<td>4 Graphical Data Visualisation Tool</td>
<td>How to visualize graphical data from a dataset resource</td>
</tr>
<tr>
<td>5 Help Desk</td>
<td>How to contact The Portal’s Help Desk</td>
</tr>
<tr>
<td>6 Metadata Quality Assurance (MQA)</td>
<td>Monitoring tool for the metadata quality:</td>
</tr>
<tr>
<td></td>
<td>- The Global Dashboard View</td>
</tr>
<tr>
<td></td>
<td>- Catalogue details view</td>
</tr>
<tr>
<td>7 SPARQL Manager</td>
<td>How to run SPARQL Queries:</td>
</tr>
<tr>
<td></td>
<td>- SPARQL Search</td>
</tr>
<tr>
<td></td>
<td>- SPARQL Assistant</td>
</tr>
<tr>
<td>8 Metadata Transformer Service</td>
<td>How to add new harvesters</td>
</tr>
</tbody>
</table>

Table 3-1: Main functions of the Portal Beta Version
3.1 Portal Home Page

![European Data Portal Home Page](image)

- **Latest News**
  - 01/03/2016: Data for Humanity: An Open Letter
  - 29/02/2016: OpenDataAP-WO Virtual Workshop
  - 28/02/2016: Hack4it Sweden

- **Browse Categories**
  - Agriculture, Forestry, Fisheries & Food
  - Energy
  - Space & Satellites
  - Transport
  - Industry & Finance
  - International Issues
  - Government & Public Sector
  - Justice, Legal System & Public Safety
  - Environment
  - Education, Culture & Sport
  - Health
  - Population & Society
  - Science & Technology

- **Featured Highlights**
  - European Data Portal open calls on Open Data use cases
  - European Data Portal open calls on Open Data use cases
  - European Data Portal open calls on Open Data use cases
  - European Data Portal open calls on Open Data use cases

- **Tweets**
  - European Data Portal
  - European Data Portal
  - European Data Portal

- **FAQ**
  - Search
  - Contact
  - Cookies
  - Legal notice

Version 1.0 | Last update: 03/03/2016

Your feedback will help us to improve the overall user experience. Any suggestions?
### 3.1.1 How to browse through the Editorial Content of the Portal

The editorial content of the Portal is organized into 4 main menu items:

1. What we do
2. Providing Data
3. Using Data
4. Library

1. **Click on “What we do”**

   The system displays a separate page with information on what is done in the Portal.

---

The European Data Portal harvests the metadata of Public Sector Information available on public data portals across European countries. Information regarding the provision of data and the benefits of re-using data is also included.

**What is Open Data?**

Open (Government) Data refers to the information collected, produced or paid for by the public bodies (also referred to as Public Sector Information) and made freely available for re-use for any purpose. The licence will specify the terms of use. These principles for Open Data are described in detail in the Open Definition.

Public sector information is information held by the public sector. The Directive on the re-use of public sector information provides a common legal framework for a European market for government-held data. It is built around the key pillars of the internal market: free flow of data, transparency and fair competition. It is important to note that not all of the public sector information is Open Data.

Find out more about the PSI Directive and other non-legislative activities of DG CONNECT in this area.

**About the European Data Portal**

Going beyond the harvesting of metadata, the strategic objective of the European Data Portal is to improve accessibility and increase the value of Open Data:

- **Accessibility:** How to access this information? Where to find it? How to make it available in the first place? In domains, across domains, across countries? In what language?
- **Value:** For what purpose and what economic gain? Societal gain? Democratic gain? In what format? What is the critical mass?

The European Data Portal addresses the whole data value chain: from data publishing to data re-use.

Within the Portal, sections are dedicated to:

**Searching datasets:** Categories have been established to structure the metadata harvested from the various countries. These categories follow the revision of the DCAT Application Profile and have been mapped against the EuroVoc Thesaurus.
2. Click on “Providing Data”, then on sub-menu “Overview”

System displays a separate page with information on how to provide data to the Portal. This page mainly addresses the suppliers (harvested portals) of the data.
3. Click on “Using Data”, then on sub-menu “Overview”

The system displays a separate page with information on how the Portal data/metadata can be used. This page mainly addresses the users of the data.
4a. Click on “Library”, then on sub-menu “Overview”  
System displays a separate page with information on available eLearning modules and Library materials (downloadable documents).
4b. eLearning

By clicking on the “eLearning” header and then on the button on the subsequent page, the system switches to the training platform from which 13 training lessons can be directly taken online.
4c. Training Companion

By clicking on the “Training Companion” sub-menu item, the system provides detailed information on how to deliver training on the basics of Open Data as well as the corresponding supporting materials.
3.1.2 How to view the “Latest News”

The Home Page displays the latest 4 news items in the “Latest News” panel on the left hand side.

- Click on any of the 4 news items to display the complete news article (here: item#1).
- Click on “News Archive” in order to find previous news articles that have been published in the past.
3.1.3 How to subscribe to the EDP Newsletter

On the Home Page in the left side panel:

- **Click on the button “Subscribe to newsletter”**
  
The system opens the “Newsletter subscriptions” page.

On the “Newsletter subscriptions” page:

- **Enter your E-Mail address**
- **Click on the button “Subscribe”**

  The system will display a notification message after successful subscription.
3.1.4 How to view “Tweets”

The Home Page displays the latest tweets on the European Data Portal in the “Tweets” panel on the left hand side.

- Click on any of the tweets to display the complete tweet on twitter.
- Scroll vertically to see previous tweets.
3.1.5 How to switch to another User Language

Select another language from the language selection box located on the upper right corner of the home page.

The User Interface as well as the main editorial content is displayed in the selected language.

There are currently 6 languages supported by the Portal Version 1:

   English (en), French (fr), German (de), Italian (it), Polish (pl) and Spanish (es).

**Note:**

Detailed editorial content – apart from the landing pages - is only available in English / French:

- Goldbook: (en)
- eLearning modules: (en, fr)
- Training companion: (en)
- Library: (en)
3.1.6 How to Search for Datasets by Category

On the Home Page, the user may click on any of the data categories below: icon or on icon text (here: “Health”)

The system displays all datasets found under the selected category, as shown below (here: for dataset category “Health”).
The user can scroll through the datasets found by using the pagination at the bottom of the page.
3.1.7 How to Search for Datasets by Keyword

(please refer to section 3.2.1)
3.2 Datasets (CKAN)

The datasets section is the main access point for browsing, filtering and searching the datasets. It offers a faceted search, a full text search and a geographical search. The dataset view provides access and information to the distributions of the dataset.

The home page of this section appears like this:
3.2.1 Entering the Datasets-View

The user has the following possibilities to enter the datasets view:

- Browsing directly to [http://europeandataportal.eu/data](http://europeandataportal.eu/data)
- Clicking on “Datasets” in the data tools submenu
- Performing a full text search by clicking on “Search”, either with or without a keyword

3.2.2 How to filter datasets by using “Faceted Search”

The user can find suitable datasets by performing a “Faceted Search”. This means the user systematically adds properties, which the desired dataset should fulfill, e.g. a dataset should be part of a specific catalogue or category. The following properties are available:

- Catalogues,
- Categories,
- Tags,
- Formats,
- Licenses.

Those facets are presented on the left side of the main dataset page. The available options for each facet always reflect the availability of it in the current set of results. The numbers in brackets indicate how many datasets have that property e.g. there are 34,179 datasets in the current result with a distribution in CSV format.
By clicking on a facet the result set if filtered by the property. Multiple selections are linked together. The following example would find datasets which are from catalogue “data.gov.uk” AND are in the categories “International issues” AND “Health” AND have a distribution in the format “CSV”.

Facets can be selected and deselected. In addition a key word can be entered to filter the results even more.
3.2.3 Additional information about a dataset

Datasets in the result list may have additional icons. The balloon indicates that the metadata of the dataset is not available in the user’s language at this moment. The eye indicates that a preview for one or more resources of the dataset is available. This may be a map or a visualisation.

3.2.4 How to filter datasets by geographical area

Many datasets are associated with geographical information. It is possible to filter the datasets by providing a location or area. Only datasets which are within this geographical area are displayed in the result set. The core idea is to select the desired geographical area by highlighting it within a provided map. There are two approaches to provide this area:

- The user enters a location key word in the “Filter by location” input field on the left side, which triggers a search for suitable geo names. All the results are linked to coordinates. By selecting one result the location is automatically highlighted within the map. The result set is filtered accordingly.
The second possibility is a manual selection of an area. The users clicks on the “Pencil” button and the map enlarges. An area is selected by drawing a rectangle over the desired location. The filtering is applied by clicking on “Apply”.

---

A map with a rectangle drawn over it, labeled with the location of Paris, Ile-de-France, Paris.
3.2.5 How to display dataset distributions

Besides special formats (see chapters 3.3 and 3.4) distributions are not displayed within the platform but only linked. When the user enters the detailed view of a dataset by selecting it in the search result all available distributions are listed. The link to the distribution is the “Link” button right next to the title and description of the distribution.

Furthermore the link can be entered from the distribution detail view, which is accessible by clicking on the name of a distribution.
3.2.6 How to view licensing information

Licensing information is available for all datasets associated with common licenses, which are supported by the License Assistant. When available a link to the assistant is provided on left side of a dataset. By clicking on it the License Assistant is opened in a new window, displaying relevant information for this particular license.

3.2.7 How to switch to another user language

The user language can be switched by selecting a language from the dropdown menu on the top. This changes the interface language and the language of metadata fields of the datasets, which are provided in multiple languages, mainly title and description.

3.2.8 How to use Multi-Language Search

By default the search is performed only in the user language. By clicking on “Language” on the main page the different languages can be selected and deselected. The key word will be searched in all selected languages.
3.2.9 How to browse by catalogues

By browsing to Data > Catalogues all source catalogues are listed and can be searched. In addition they can be filtered by their country of origin. This can be done by selecting one or more country in left country facet.
3.3 Visualization of Geo-Spatial Data

The visualization of geo-spatial data within the European Data Portal provides previewing functionality for spatial open data. The aim is to allow the user to assess if a dataset meets specific requirements in terms of spatial and thematic coverage. The functionality that is provided in the header (links to disclaimers and language switching) is consistent in the entire portal.

3.3.1 How to visualize geo-spatial data from a dataset resource

Accessing the geo-spatial visualization is achieved via the CKAN interface. A user searches for specific data, enters the dataset view of reasonable results and displays the available distributions (see Section 3.2.5). If a dataset distribution is supported by the geo-spatial visualization, a globe button is displayed (see Figure 1). This is the entry point into the map viewer application. Supported formats are OGC Web Map Service (WMS) and GeoJSON. If the user visits the geo-spatial visualization for the first time, an interactive user tutorial is provided to guide the use through specific functions of the user interface, similar to this written user manual.

Once within the map viewer application, the user can decide which layers to be displayed. Most Web Map Service (WMS) instances provide more than one layer. The geo-spatial visualization provides a dialog for choosing the desired layers for display (see Figure 2).
In order to examine the resource in more detail, the user can click on a geographic feature while the “Feature Info” tool is enabled (see Figure 3). This function is only enabled if the service supports this kind of detailed query.

The different displayed layers can be examined using the “Legend” tool. If the external service provides legend graphics, the user can interpret the given symbology and temporarily disable the display of layers (see Figure 4).
The two buttons on the bottom right of the map viewer (see Figure 5) can be used to display an informational disclaimer and to start the user interface tutorial.
In some cases the user might experience the display of an error message. The geo-spatial visualization tries to support all flavors of external services but cannot guarantee to work with broken services. In these situations an error message dialog is presented and the user can decide if a support ticket shall be opened (see Figure 6).

Figure 6 – Error message dialog.
3.4 Graphical Data Visualisation Tool

This section describes the features of the graphical visualisation tool for numeric data. The features are currently available for XLS (Excel) and CSV files, except for the selection of the sheet name which is applicable only for Excel files.

Most GUI elements from the “Graph” tab (records selection, search box, filters and fields buttons) are also available on the “Grid” tab and work in the same way.

3.4.1 How to visualize graphical data from a dataset resource

As a result of a dataset search, the system displays on the “Dataset” tab all distributions (resource/data files) that are part of the selected dataset. Each XLS or CSV distribution of the dataset can be further explored by using the visualization button – if available.
After clicking on the visualization button, the user should execute the following steps:

**Step 1: Select a Sheet Name**

The Portal is parsing the XLS file and lists all available worksheets. Select one of the sheets contained in the XLS file.

**Step 2: The Grid View**

The grid appears once the user has clicked on a file name (and a sheet name in case of an Excel file).
Step 3: The Graph Tab
Click on the graph tab in order to display the corresponding graph.

1. Selection of the sheet name
2. Button to go back to the “Grid” view
3. Selection of a range of data records
4. Search box
5. Filters button to open the filters form
6. Fields button to open the fields box
7. Select box to select the graph type
8. Select box to select the group column (Axis 1)
9. Select box to select the series A (Axis 2)
10. Button to add series
Step 4: Filters

Once the user clicks on the filters button the filters form appears.

After clicking on “Add filter”, the available fields appear and the user can now filter the data with a specific value or range.
Step 5: Fields

Click on the “Fields” button in order to display the fields box that contains the list of available fields.

Step 6: Adding another data series

Click on the “Add Series” button to add another data series to the current graph as shown below.
3.5 Help Desk

3.5.1 How to contact the Portal’s Help Desk

The European Data Portal Help Desk can be contacted:

- **By email**: help@europeandataportal.eu

- **By phone**: the Portal’s Help Desk is staffed by a multilingual team of experts, who can be contacted from Monday to Friday from 09:30 to 17:30 (CET).
  - EN: (+352) 31 44 01-448
  - FR: (+352) 31 44 01-449

- **By providing comments and suggestions via the online contact/feedback form available from the Portal’s home page.**

From the header:

[European Data Portal Header Image]

From the footer:

[European Data Portal Footer Image]
3.6 Metadata Quality Assurance (MQA)

The Metadata Quality Assurance is a component to monitor the quality of the metadata that is harvested from other portals or stored manually with the EDP metadata creation form. Currently quality measurement is based on metadata schema violation against DCAT-AP 1.1 and availability of the distributions of a dataset. The MQA is running checks on a weekly basis. Each check has a duration of a couple of days. This is due to the HTTP GET requests against each single distribution in the portal, checking their availability.

3.6.1 The Global Dashboard View

This is the landing page of the MQA. It gives an overview of the quality situation across all catalogues, datasets and distributions in the Portal. If the user wants to investigate the mentioned issues with datasets or distributions in detail, he/she can use the navigation bar on top of the page to navigate to the catalogue details page.

An information icon in the upper right corner is indicating that the user can get explanation texts for all diagrams on the page. The introduction text shows the planned schedule for the next check to be performed by the MQA. The bottom of the page shows the date when the last check started. If the MQA is currently running a check, then a spinning wheel beside the last check date is indicating this.
Metadata Quality Dashboard

This dashboard provides an overview of the metadata quality and compliance of datasets on the European Data Portal. It includes the following sections:

1. Distribution Statistics
2. Ratio of machine-readable distributions
3. Most used distribution formats
4. Dataset Compliance Statistics

Distribution Statistics

The following charts provide a summary of the distribution statistics. The left-hand pie chart shows the percentage of accessible distributions, while the right-hand pie chart shows the percentage of error status codes. The bar chart displays the top 20 catalogs with the most accessible distributions.

Ratio of machine-readable distributions

The pie chart shows the percentage of datasets that are machine-readable, with the corresponding percentage indicated.

Most used distribution formats

The bar chart displays the most used distribution formats, with the corresponding percentage indicated.

Dataset Compliance Statistics

The following charts provide a summary of the dataset compliance. The left-hand pie chart shows the percentage of compliant datasets, with the corresponding percentage indicated. The bar chart displays the top 20 catalogs with the most compliant datasets, with the corresponding percentage indicated.
3.6.2 Catalogue details view

This page presents a detailed view of the issues per catalogue. The user can choose the catalogue she wants to investigate by selecting it in the selection box in the left menu pane. As a next step, the user can choose whether to see the catalogue dashboard, the distribution availability or dataset schema violations. The catalogue dashboard is the default view when visiting this page.
3.6.2.1 Catalog Dashboard

The catalog dashboard provides the same information as the overall dashboard gives, but only for the selected catalogue.

3.6.2.2 Distributions

This view provides a detailed report of the distributions status of the selected catalogue. The user can see the URL of the distribution that was checked (access and download URL), the corresponding HTTP response code and the date when this issue occurred the first time. Additionally, the distributions format is check for its machine readability. The machine readability check is based on the results of the Open Data Monitor project\(^1\).

\(^1\) [https://github.com/opendatamonitor/odm.restapi/blob/master/odmapi/def_formatLists.py#L44-L87](https://github.com/opendatamonitor/odm.restapi/blob/master/odmapi/def_formatLists.py#L44-L87)
### 3.6.2.3 Violations

The MQA checks for each dataset if it is DCAT-AP compliant. The validation is done on the metadata that is stored in CKAN. Therefore, a JSON schema was created and is used for validation. The user can see the location where the violation occurs and its reason for each not compliant dataset of the selected catalogue.
3.7 SPARQL Manager

The SPARQL Manager provides a graphical user interface (GUI) for sending user defined queries to the Virtuoso SPARQL query engine.

The powerful SPARQL Protocol and RDF Query Language are primarily aimed at professionals for querying metadata as Linked Data. A basic knowledge of the DCAT-AP specification is highly recommended.

In the future, users of the SPARQL Manager will be able to save their queries for scheduled execution. Additionally a notification will be send to the user when a result has changed.

Clicking the info icon in the upper right corner will display a step-by-step walkthrough of all components with a short info about their function.

This is possible in both modes of the SPARQL Manager, the search and the assistant mode, which will be described in the following sections.

3.7.1 SPARQL Search

In this mode you can load some predefined example queries from the right side into the editable text area to introduce yourself with the very basic SPARQL syntax. Limiting the number of returned results is possible by selecting a value from the Limit-dropdown or by editing the query directly. Furthermore the format for the result can be selected. After clicking the Search-Button the result is
displayed in Result data preview area below. The preview may be truncated depending on the size of
the result. The complete result could always be downloaded as a file by clicking the Download-link on
the right side.

3.7.2 SPARQL Assistant

The SPARQL assistant extends the functionality of the simple SPARQL search described in the
previous section. More complex queries for datasets could be built by clicking several options in the
GUI.

1. Category/Theme

Select one or more Categories/Themes defined in the DCAT-AP standard to filter the results. Datasets will only be listed in the result if they belong to all selected categories/themes.

2. Distribution format

Datasets can contain distributions in different formats. You can select one or more formats to exclude datasets from the result which distributions do not contain the desired formats. If your
format selection by the DCAT-AP standard is to strict you may enter a custom format string at the bottom of the list, which will result in a simple text comparison in the format label field.

3. **Search term**
   Enter a search term which must be found in the dataset. You can limit this to the title or the description. By enabling both options the term has to be in the title and the description.

4. **Count**
   By clicking the count option only the number of results that matches the SPARQL query are returned.
3.8 Metadata Transformer Service

The metadata transformation service provides functions for developing and operating so-called harvesters. A harvester is an application which imports metadata from data catalogues with the aim to publish them on another metadata register. However, the metadata structures used by the source and target portal often differ from each other, so that the imported data cannot be stored directly in another registers. Therefore, while harvesting, the imported metadata are first transformed into the required structure and then exported to the target catalogue.

Since the datasets in the original metadata catalogues can be updated over time, the harvesters have to be executed regularly to fetch the current version. While creating or updating the harvester, the metadata transformation service allows you to set the frequency and the date for the next harvester run. If the source structure of the data is affected by the changes, it is also required to adapt the mapping rules used by the harvester. It is up to the user to customize the transformation rules to his harvester in the detail pages as needed.

**General Note:** The importing and exporting of data can affect the operation of the metadata catalogue services (performance issues).

### 3.8.1 The User Dashboard

The user dashboard shows all repositories and harvesters owned by the logged in user. On the left hand side all repositories, ordered by name, are listed. Each repository is represented by a panel, given a quick overview over some details, like name, owner, description and the homepage. On the right side, the same applies for the harvesters. Instead of the homepage, the harvesting frequency is indicated.
Repository functions:

- Create new repositories (+ New on the left side)
- Edit a repository (yellow button on the right bottom of the panel)
- Go to importing harvesters (Import drop down menu). Harvesters that uses this repository as source.
- Go to exporting harvesters (Export drop down menu). Harvesters that uses this repository target.

Harvester functions:

- Show harvester (orange button with eye on the right bottom of the panel)
- Edit harvester (yellow button on the right bottom of the panel)

### 3.8.2 Creating Repositories

To create a new Repository, click on “+ New” on the dashboard or “New Repository” on the list of repositories. This function is only available when logged in.

Fill in all fields and click on “Create”.

Required fields are:

- Name
- URL
- Publisher
3.8.3 Creating Harvesters

To create a new Harvester, click on “+ New” on the dashboard or “New Harvester” on the list of harvesters. This function is only available when logged in.

Provide on the „General“ tab the name, a description and the frequency for scheduling this harvester. Then, step forward to select the repositories.

On the „Repositories“ tab you can select from the list of your repositories or any public repository, the source and the target. Depending on the repository type, you may be provide additional information, like e.g. an API key giving you write access to the target. Finally you should edit the
transformation script on the „Transformation“ tab. If you don’t, and click now on „Create“, the service will probably provide a default example script that you can modify for your needs.

3.8.4 The Harvester Overview

The overview page of a harvester shows the source and the target, and additionally the next scheduled run time, as well as a list of all runs from the past or present.

From here you can:

- Edit the source or target if you are the owner of it.
- Select a run to see more details about it.
- Examine the transformation script.

If you are the owner you can additionally:

- Start and stop a run.
- Edit the harvester.
3.8.5 View or Monitor a Running Task

The view of a run shows all log entries and the number of added datasets, updated datasets and rejected datasets. Additionally, you can see the start time, the end time and the duration of the run.
In case this run is still ongoing, the progressbar indicates the percentage of the completion. Also, a stop button is provided to interrupt the run at any time. Otherwise, it an be removed, if you are the owner of the harvester this run belongs to.