
Odf Docs Documentation

Release 0.0.1

Giuseppe Ciulla

Apr 17, 2018

Contents

| | |
|-------------------------------------|----------|
| 1 Content | 3 |
| 1.1 Architecture Overview | 3 |
| 1.2 Administration Manual | 4 |
| 1.3 End User Manual | 13 |

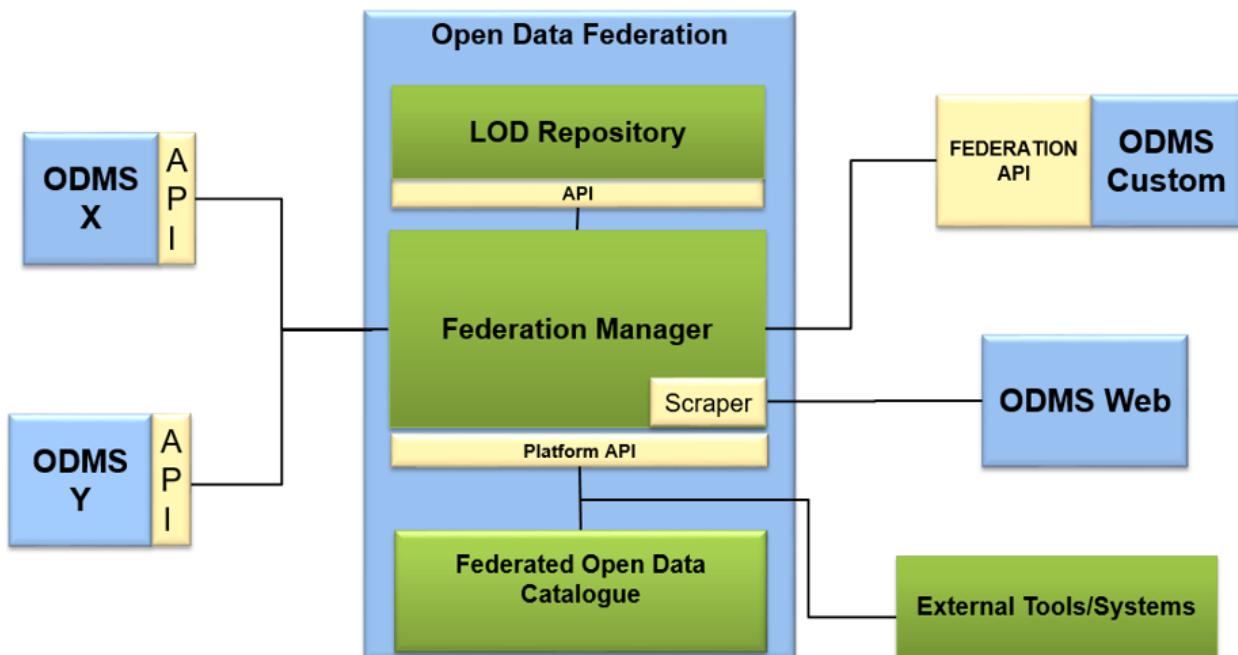
Open Data Federation (ODF) is a web application able to federate existing Open Data Management Systems (ODMS) based on different technologies; in this way ODF provides a unique access point to search and discover open data sets coming from the different federated ODMS. ODF uniforms representation of collected Open Data Set, thanks to the adoption of international standards (DCAT-AP) and provides a set of APIs to develop third party applications. ODF supports natively ODMS based on CKAN, DKAN and Socrata and provides a set of APIs to federate ODMSs not natively supported; these ODMSs have to implement and expose them. In addition, it is possible to federate a generic Web Portal, either by using the Web Scraping functionality or by uploading a dump of the datasets in DCAT-AP format. Moreover, ODF provides a SPARQL endpoint in order to perform queries on 5 stars RDF linked open data collected from federated ODMSs.

CHAPTER 1

Content

1.1 Architecture Overview

Open Data Federation provides access to resources of federated ODMSs from a single-entry point through a set of APIs and is able to retrieve, search and visualize datasets from different ODMSs. The platform is responsible for collecting metadata of Open Data from federated ODMS catalogues and then for translating them into a common and uniform format. In addition, it manages Linked Open Data (LOD), importing them into a specific repository in order to perform queries on them. The following picture illustrates the architecture of the Open Data Federation.



Its main components are:

- **Federation Manager:** is the core of the platform that interacts with federated ODMS catalogues; it is responsible for managing internal federation processes. It provides the main functionalities through Platform API in order to be accessed by external application or by the Federated Open Data Catalogue. Main functionalities provided by the FM are:
 - ODMS catalogues management: registration, removal and monitor.
 - Federated full text search: possibility to search for specific Open Data on the federated ODMS catalogues.
 - Federated queries on Linked Open Data.
 - Federation configuration management
- **LOD Repository:** is the central store in which collected Linked Open Data retrieved from federated ODMS catalogues are stored, in order to perform queries on them and to provide collected results in different formats.
- **Federated Open Data Catalogue:** is a web application that allows end users to access the FM functionalities calling the Platform API. In particular, the Federated Open Data Catalogue allows to:
 - Manage administrator authentication
 - Search for Open Data/Linked Open Data, visualise and manage results
 - Manage Federation and configuration.

The Federation Manager functionalities can be also accessed by a generic external system (e.g. client application) using the Platform API. It is important to underline that each ODMS catalogue depicted in the picture is a generic system that manages OD/LOD. Usually it consists in a web portal associated to a database. In order to be federated in the ODF, the ODMS has to provide some basic functionalities through RESTful APIs. One of the objectives of the ODF is to allow the federation of different ODMSs with minimum effort. Different type of ODMS catalogues will be natively supported by ODF: CKAN, Socrata, DKAN or portals that provides the datasets through a DCAT-AP or DCAT-AP_IT dump; ODF provides Federation API Specification to allow “custom ODMS catalogues” to join the federation; moreover, custom ODMS catalogues that does not provide APIs can join the federation through the scraping of its web portal.

1.2 Administration Manual

This section provides the description of the administration functionalities. An administrator should be able to install, deploy, perform the sanity checks on the environment and manage the platform through the Federated Open Data Portal.

1.2.1 Installation

This section covers the steps needed to properly install the Open Data Federation

Requirements

ODF has the following requirements that must be correctly installed and configured

| Framework | Version | License |
|-------------------------|-----------------|--|
| Java SE Development Kit | 8.0 | Oracle Binary Code License |
| Apache Tomcat | 8.0 | Apache License v.2.0 |
| MySQL | 5.7.5 Community | GNU General Public License Version 2.0 |
| RDF4J Server | 2.2.1 | EDL 1.0 (Eclipse Distribution License) |
| RDF4J Workbench | 2.2.1 | EDL 1.0 (Eclipse Distribution License) |

Libraries

ODF is based on the following software libraries and frameworks.

| Framework | Version | Licence |
|-------------------------------------|--------------|--|
| Apache SOLR-Lucene (SOLR Core) | 6.6.0 | Apache License |
| Apache Http Client | 4.5.2 | Apache License |
| Apache Http Core | 4.5.2 | Apache License |
| Mysql connector (Community Release) | 5.1.39 | GPL 2.0 (GNU General Public License Version) |
| Hibernate | 5.2.10.Final | LGPL 2.1 (GNU Lesser General Public License) |
| Hikari | 2.6.1 | Apache License 2.0 |
| Log4j | 2.7 | Apache License 2.0 |
| CKANClient-J | 1.7 | AGPL 3.0 (GNU Affero General Public License) |
| RDF4J-Runtime | 2.2.1 | EDL 1.0 (Eclipse Distribution License) |
| AngularJS | 1.5.9 | MIT |
| Angular-UI - bootstrap-ui | 0.13.3 | MIT |
| Bootstrap | 3.3.2 | MIT |
| Bootstrap-Material | 3 | MIT |
| Smart-table | 2.1.3 | MIT |
| ngImageCrop | 0.3.2 | MIT |
| spin.js | 2.3.2 | MIT |
| angular-zeroclipboard | 0.8.0 | MIT |
| angular-xeditable | 0.1.8 | MIT |
| angular-pagination | 0.11.0 | MIT |
| Ace Editor | 1.2.0 | BSD |
| Angular-UI - ace-ui | 0.2.3 | MIT |

Prerequisites

The following tools should be properly installed on your computer:

- Git
- NodeJs (with NPM)
- Bower
- Maven

Proxy configurations

In order to use the different tools behind a proxy please execute the following commands (*username* and *password* are your credential, *proxyhost* is the host name or the IP address of the proxy and *proxypot* is the TCP port of the proxy):

- **Git: open a command prompt and execute:**

```
$ git config --global http.proxy http://username:password@proxyhost:proxypot
$ git config --global https.proxy http://username:password@proxyhost:proxypot
```

- **Npm: open a command prompt and execute:**

```
$ npm config set proxy http://username:password@proxyhost:proxypot
$ npm config set https-proxy http://username:password@proxyhost:proxypot
```

- **Bower:** change the current directory to the one that contains the “*bower.json*” file and create/edit the “*.bowerrc*” file and add the proxy configuration:

```
{  
    "proxy" : "http://username:password@proxyhost:proxypor",  
    "https-proxy" : "http://username:password@proxyhost:proxypor"  
}
```

- **Maven:** edit the file “*Path_Of_Maven/conf/settings.xml*” and add to the “*<proxies>*” section the proper configuration following the example provided in the same file (please refer to maven guide <https://maven.apache.org/guides/mini/guide-proxies.html>)

Create WAR packages

Open a command prompt and Execute the following command to clone the repository:

```
$ git clone https://production.eng.it/gitlab/OPSI/OpenDataFederation.git  
$ cd OpenDataFederation
```

In this folder you will find two subfolders:

- **FederationManager:** this folder contains the server side application of the Open Data Federation
- **ODFCatalogue:** this folder contains the client side application of the Open Data Federation

FederationManager.war

Move in FederationManager folder:

```
$ cd FederationManager  
$ mvn package
```

Note. Execute this command in a network without proxy because of jitpack dependency.

ODFCatalogue.war

Move in ODFCatalogue folder:

```
$ cd ODFCatalogue  
$ cd /src/main/webapp  
$ bower install  
$ cd ../../..  
$ mvn package
```

1.2.2 Deployment

This page shows the deployment procedure of the Open Data Federation.

Artefacts

These are the artefacts that must be installed in order to run ODF:

- FederationManager.war

- ODFCatalogue.war
- rdf4j-workbench.war & rdf4j-sesame.war (you can get both [here](#), into "war" folder)
- opendata_federation.sql

Database creation

ODF relies on a MySQL database to store all the application data and collected Open Datasets.

So before deploying the application, it is necessary to create a new database, by importing in the MySQL server the provided SQL dump file:

- **opendata_federation.sql**

This dump already contains the statement that creates the “**opendata_federation**” DB automatically. In addition it creates an administration user with the following credentials:

username: admin

password: admin

Note. To change the administrator password login in the Open Data Catalogue with the previous credentials then go to the **Administration -> Manage Configurations -> Update Password** section.

WARs deployment

Move all the WAR artifacts to the “webapps” folder of Tomcat installation, start it up and wait until they are deployed.

RDF repository creation

Once the Tomcat server started, go with browser to the URL “**localhost:8080/rdf4j-workbench**”

Note. Change the port number according to the configuration of server.xml file of Tomcat “conf” folder (default 8080)

Through the RDF4J GUI, select “new repository” on the left menu, then create a new repository of type “**Native Java Store**” called “**ODF**”.

Configuration

Once all the WAR files are deployed and the server has started, modify the following configuration files, located in the deployed folders of Tomcat “webapps” folder.

- ODFCatalogue/WEB-INF/classes/
 - In **configuration.properties** file, change the following properties:
 - * Base url part of **ADMIN_SERVICES_BASE_URL** property with the **PUB-LIC** domain where is exposed the runtime environment. (Example: <https://opendatafederation.eng.it/FederationManager/api/v1/administration>)
 - * Base url part of **CLIENT_SERVICES_BASE_URL** property with the **PUB-LIC** domain where is exposed the runtime environment. (Example: <https://opendatafederation.eng.it/FederationManager/api/v1/client>)
- FederationManager/WEB-INF/classes/
 - In **configuration.properties** file, change the following properties:

- * **DB_HOST, DB_USERNAME, DB_PASSWORD** with the actual parameters of the MySQL server installation.
 - * **http.proxyHost, http.proxyPort, http.proxyUser, http.proxyPassword** with the proxy parameters, leave blank if none. Change **http.proxyEnabled** to **true** if the previous proxy parameters are provided.
 - * **odmsDumpFilePath** and **dumpFilePath** with the folder path where to save the DCAT-AP dump files.
NOTE The path **MUST** end with "\" or "/".
 - * **sesameRepositoryName** must have the same value of the newly created RDF repository.
 - * **enableRdf** to **true**, in order to enable RDF retrieval, configured with the following parameters, according to the Tomcat configuration, as described in the “**RDF repository creation**” step:
 - **sesameServerURI** with the URL where to find the "repositories" endpoint of RDF4J. Example: `http\\://localhost\\:8080/rdf4j-server/repositories/`
 - **sesameEndPoint** with the URL where to find the "query" endpoint. Example: `http\\://localhost\\:8080/rdf4j-workbench/repositories/ODF/query`
- In **hibernate.properties** file, change the following properties:
- * **hibernate.connection.url, hibernate.connection.username, hibernate.connection.password** with the actual parameters of the MySQL server installation.

1.2.3 Sanity Checks

In order to apply the previous changes, restart the Tomcat server. The Sanity Checks are the steps that the Administrator will take to verify that the installation is ready to be used and tested.

Note. Change the “BASEPATH” value with the actual host and port where is exposed the runtime environment (Tomcat).

Catalogue Access Testing

Once the server restarted, go with browser to `http://BASEPATH/ODFCatalogue`

When the home page is showed, perform the following steps:

- Check that the message "There are no federated catalogues" is showed.
- Check that you can perform the Login as Administrator, in the appropriate section in the top bar.

Platform API testing

- Open a command prompt and execute: `curl http://BASEPATH/FederationManager/api/v1/administration/info`
- Check that you get the version number as output, along with other information about API version and timestamp

1.2.4 Platform Management

This section provides the description of the Administration Functionalities. Through the Open Data Catalogue a logged administrator can:

- Manage ODMS Catalogues;

- Manage configuration parameters;
- Manage datalets;
- View platform logs.

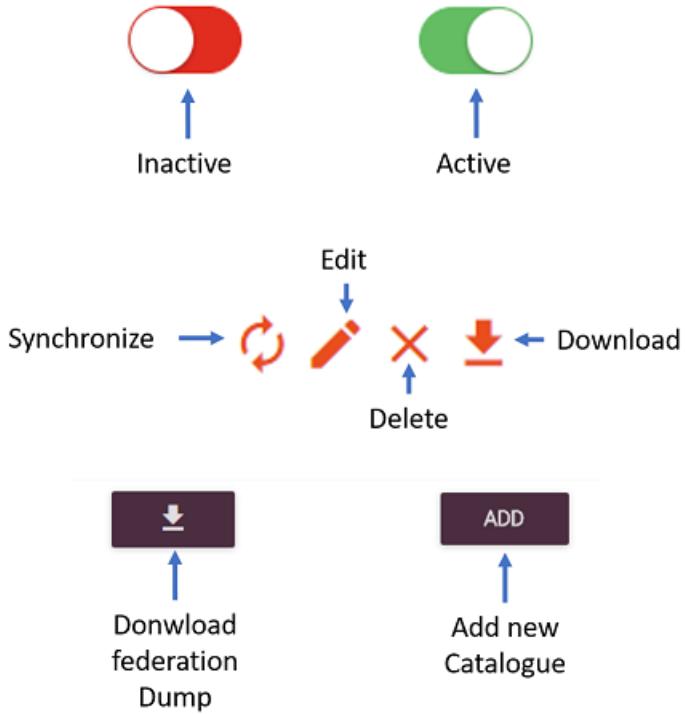
Catalogues Managements

In this page the administrator manages the Catalogues. In particular, he/she is able to:

- Add/Edit/Delete a Catalogue
- Add from a Remote Catalogues
- Activate/Deactivate a Catalogue;
- Start the synchronization of a Catalogue;
- Download a catalogue dump or the federation dump with DCAT-AP profile

| Federated Open Data Catalogue | | | | | | | | | Dataset Search | SPARQL Search | Federated Catalogues | Administration | LOGOUT | |
|-------------------------------|---------------------------------|---------------------------------|------|-------|--------|----------|---------------|---------------------|----------------|---------------|----------------------|----------------|--------|--|
| Local | | Remote | | | | | | | | | | | | |
| Active | ▲Name | Host | Type | Level | Status | Datasets | Update Period | Last Update | | | | | | |
| | Ancitel Open Data | http://okan.ancitel.it | CKAN | 3 | | 10 | 1 day | 2018-03-22 11:02:46 | | | | | | |
| | Arpaee | https://dati.arpaee.it/ | CKAN | 3 | | 21 | 1 day | 2018-03-22 18:06:11 | | | | | | |
| | Civitavecchia Open Data | http://opendata.comune.civit... | CKAN | 3 | | 43 | 1 day | 2018-03-22 18:06:45 | | | | | | |
| | Consiglio Regionale della Ca... | http://opendata-crc.di.unisa... | CKAN | 3 | | 5 | 1 hour | 2018-03-23 09:06:24 | | | | | | |
| | Consip OpenData | http://dati.consip.it | CKAN | 3 | | 16 | 1 day | 2018-03-22 19:03:50 | | | | | | |
| | Consorzio LaMMA | http://dati.lamma.toscana.it | CKAN | 3 | | 309 | 1 day | 2018-03-22 18:07:23 | | | | | | |
| | Dati Comune Genova | http://dati.comune.genova.it | DKAN | 2 | | 153 | 1 day | 2018-03-22 16:36:24 | | | | | | |
| | Dati Toscana | http://dati.toscana.it/ | CKAN | 3 | | 2844 | 1 day | 2018-03-22 17:09:26 | | | | | | |
| | Dati Trentino | https://dati.trentino.it | CKAN | 3 | | 6095 | 1 day | 2018-03-22 11:31:49 | | | | | | |
| | Fiware CKAN Catalogue | https://data.lab.fiware.org | CKAN | 3 | | 2934 | 1 week | 2018-03-21 16:39:41 | | | | | | |

The following pictures depicts the functionalities linked to every button or icons.



Add/Edit/Delete a Catalogue

By clicking on the **ADD** button the following the Catalogue form is presented to the administrator.

The screenshot shows the 'Add Catalogue' form in the 'Federated Open Data Catalogue' application. The top navigation bar includes links for Dataset Search, SPARQL Search, Federated Catalogues, Administration, and Logout.

The form itself has the following fields:

| | | |
|--|---|---|
| | Name: <input type="text"/> | Publisher Name: <input type="text"/> |
| | Host: <input type="text"/> | Country: <input type="text"/> |
| | Active: <input type="text" value="Yes"/> | Type: <input type="text" value="OKAN"/> |
| | Refresh Period: <input type="text" value="1 hour"/> | Category: <input type="text"/> |
| Description: <hr/> | | |
| <input type="button" value="BACK"/> <input type="button" value="RESET"/> <input type="button" value="CREATE"/> | | |

Here the administrator has to insert all of the information related to the catalogue and then click on the **CREATE** button.

By clicking on the *edit* icon on the Catalogue table, the user can edit most of the Catalogue's information. He/she cannot modify the *host* and *type* attributes.

By clicking on the *delete* icon on the Catalogue table, the user deletes the Catalogue and its datasets from the federation. This operation cannot be reverted.

Remote Catalogues

New Catalogues can be added to the federation using the **remote catalogues** list. This remote list is a catalogue repository maintained by Engineering. In the remote catalogue list an ODF administrator can find certified catalogues and by clicking on the *plus* icon he can insert the selected catalogue in his/her ODF instance.

| Name | Host | Type | Level | Update Period |
|------------------------------------|--|------|-------|---------------|
| Ancitel Open Data | http://okan.ancitel.it | CKAN | 3 | 1 day |
| Arpae | https://dati.arpae.it/ | CKAN | 3 | 1 day |
| Civitavecchia Open Data | http://opendata.comune.civitavecchia.rm.it | CKAN | 3 | 1 day |
| Consiglio Regionale della Campania | http://opendata-crc.di.unisa.it | CKAN | 3 | 1 hour |
| Consip OpenData | http://dati.consip.it | CKAN | 3 | 1 day |
| Consorzio LaMMA | http://dati.lamma.toscana.it | CKAN | 3 | 1 day |
| Dati Comune Genova | http://dati.comune.genova.it | DKAN | 2 | 1 day |
| Dati Toscana | http://dati.toscana.it/ | CKAN | 3 | 1 day |
| Dati Trentino | https://dati.trentino.it | CKAN | 3 | 1 day |
| Fiware CKAN Catalogue | https://data.lab.fiware.org | CKAN | 3 | 1 week |

1 2 3 4 5

Activate/Deactivate a Catalogue

This functionality allows the administrator to manage on which catalogues the user can perform searches. Indeed, if a catalogue is **active** users will find its datasets during a search; if a catalogue is **inactive** user will not find any of its datasets during a search.

Catalogue Synchronization

By default, Catalogues are automatically synchronized from the platform taking advantage of the *refresh period* attribute. If an administrator will force the synchronization of a catalogue he/she would have to click on its *synchronize* button.

Download Dump

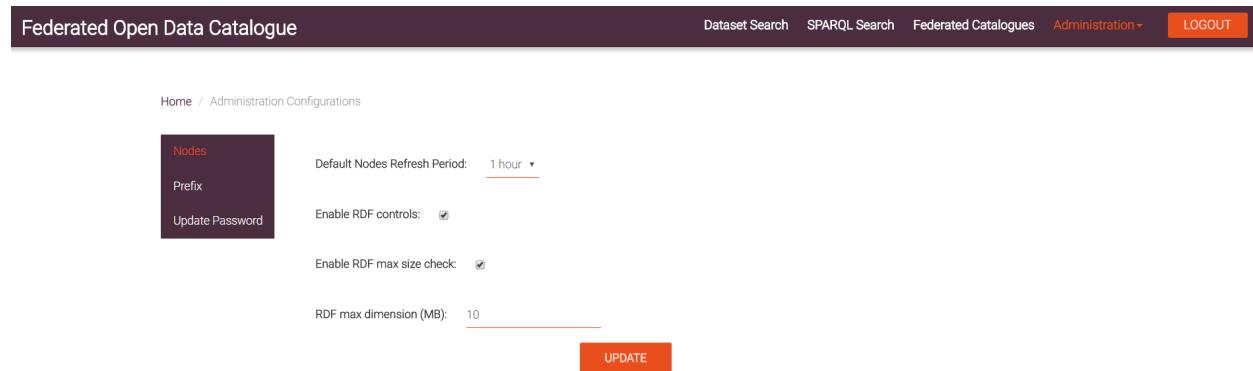
The administrator can download a DCAT-AP dump the Federated Open Data Catalogue. He/she can choose to download a single catalogue dump or the complete federation dump by clicking respectively on the download button in the catalogue's row of the table or on the global download button located at the bottom of the table.

Configuration Parameters Management

An administrator can modify some of the configuration parameters that control the loading of the RDF files into the LOD repository. In particular, he/she can:

- Enable RDF controls: if **false** all RDFs will be loaded into the LOD repository, if **true** only the RDFs which pass the controls will be loaded, the others will be discarded;
- Enable RDF max size check: this configuration parameter if **true** will enable the controls on RDFs size;
- RDF max dimension: if the previous configuration parameter is **true**, this parameter will represent the size limit of an RDF in order to be

loaded into the repository. RDFs whose dimension exceeded will be discarded. Moreover, the administrator will define the default catalogue's refresh period.



The screenshot shows the 'Administration Configurations' page with the following settings:

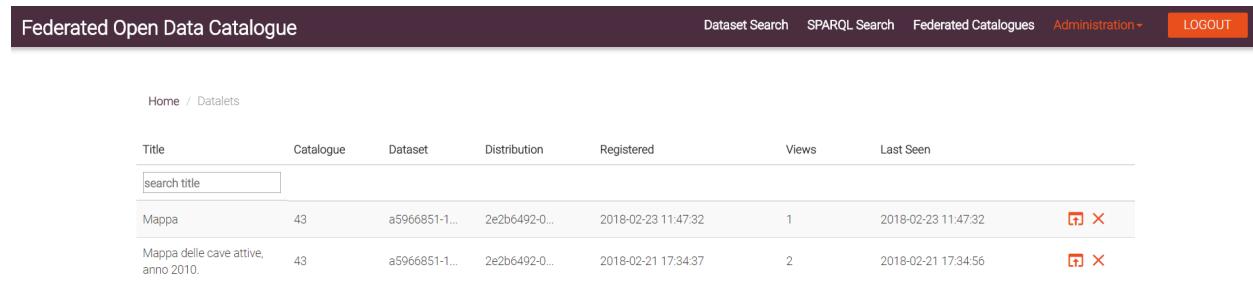
- Nodes:**
 - Default Nodes Refresh Period: 1 hour
 - Enable RDF controls:
 - Enable RDF max size check:
 - RDF max dimension (MB): 10
- Prefix:** (not visible in the screenshot)
- Update Password:** (not visible in the screenshot)

An 'UPDATE' button is located at the bottom right.

The administrator can also update his/her password and he/she can manage the RDFs' prefixes through the console.

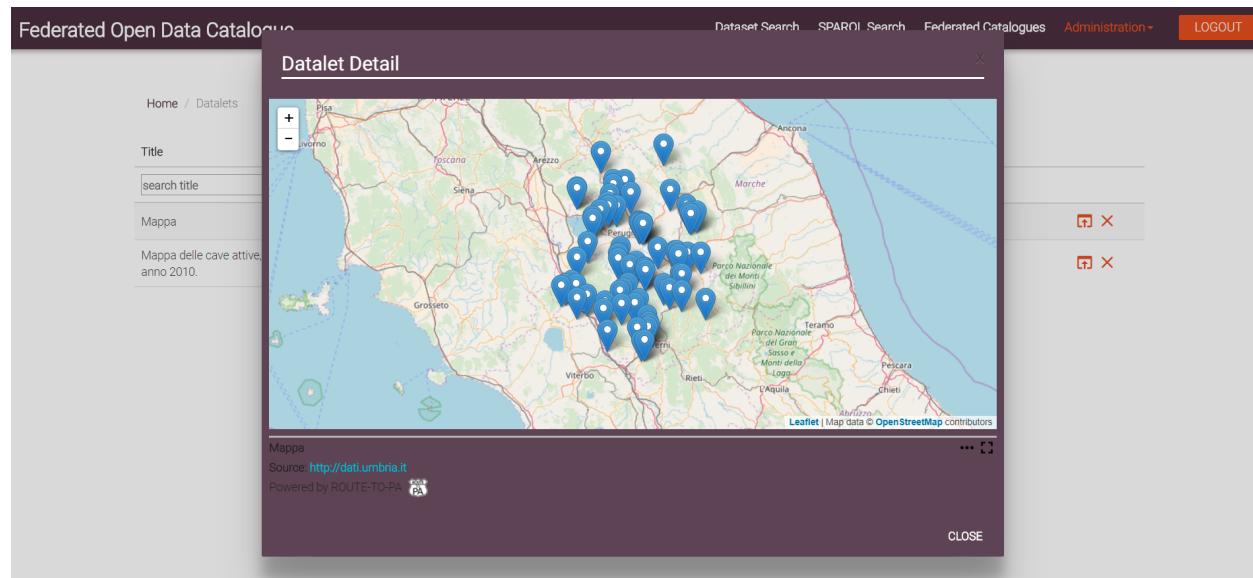
Datalets Management

Through this page, the administrator can manage all of the datalets produced by the end users.



| Title | Catalogue | Dataset | Distribution | Registered | Views | Last Seen |
|-------------------------------------|-----------|---------------|---------------|---------------------|-------|---------------------|
| search title | | | | | | |
| Mappa | 43 | a5966851-1... | 2e2b6492-0... | 2018-02-23 11:47:32 | 1 | 2018-02-23 11:47:32 |
| Mappa delle cave attive, anno 2010. | 43 | a5966851-1... | 2e2b6492-0... | 2018-02-21 17:34:37 | 2 | 2018-02-21 17:34:56 |

The administrator will check the number of views and the last time the datalets was seen by end users. The administrator will be able to delete the datalet or to see its preview.



The 'Datalet Detail' modal displays a map of Italy focusing on the central region, specifically Umbria. Numerous blue location pins are scattered across the map, primarily concentrated in the provinces of Perugia and Terni. The map includes labels for major cities like Roma, Firenze, Siena, Perugia, Ancona, and others. A legend in the top left corner of the modal indicates that the pins represent active cave locations. The bottom left of the modal shows the title 'Mappa' and the source 'Source: <http://dati.umbria.it>'. The bottom right contains a 'CLOSE' button.

Platform Logs

This page will show the *Logs* produced by the back-end server in the GUI. The administrator will be able to query the logs in order to search for a particular event. The following figure depicts this functionality.

The screenshot shows a log viewer interface with the following details:

- Header:** Federated Open Data Catalogue, Dataset Search, SPARQL Search, Federated Catalogues, Administration, LOGOUT.
- Toolbar:** Home / Platform Logs, Date range (21-March-2018 to 22-March-2018), Level (All), SEARCH button.
- Log Content:**

```

100 [ INFO ] 2018-03-21T00:00:48Z it.eng.rspa.odf.odfscheduler.job.ODMSynchJob Starting synch job for catalogue: 88
101 [ INFO ] 2018-03-21T00:00:48Z it.eng.rspa.odf.odfscheduler.job.ODMSynchJob Starting synchronization for node: OpenData Matera2019 ID: 88
102 [ INFO ] 2018-03-21T00:00:48Z org.ckan.Connection CONNECTION: OPEN
103 [ INFO ] 2018-03-21T00:00:48Z it.eng.rspa.odf.odfscheduler.job.ODMSynchJob The ODMs Node with name: OpenData Matera2019 and ID: 88 is ONLINE
104 [ INFO ] 2018-03-21T00:00:48Z it.eng.rspa.odf.connectors.CkanConnector Starting to retrieve present datasets of the node from cache
105 [ INFO ] 2018-03-21T00:00:48Z org.ckan.Connection CONNECTION: OPEN
106 [ INFO ] 2018-03-21T00:00:48Z it.eng.rspa.odf.connectors.CkanConnector Start to compare the new and old dataset IDs lists: New size: 3 - Old size: 3
107 [ INFO ] 2018-03-21T00:00:48Z it.eng.rspa.odf.connectors.CkanConnector Deleted Packages: 0
108 [ INFO ] 2018-03-21T00:00:48Z it.eng.rspa.odf.connectors.CkanConnector Last update date from which to start:2018-03-20T23:08:48Z
109 [ INFO ] 2018-03-21T00:00:48Z org.ckan.Connection CONNECTION: OPEN
110 [ INFO ] 2018-03-21T00:00:49Z it.eng.rspa.odf.connectors.CkanConnector -- CKAN Connector Response - Result TOTAL count:0
111 [ INFO ] 2018-03-21T00:00:49Z it.eng.rspa.odf.odfscheduler.job.ODMSynchJob Synchronization complete for node: OpenData Matera2019 ID: 88
112 [ INFO ] 2018-03-21T00:00:49Z it.eng.rspa.odf.odfscheduler.job.ODMSynchJob Added RDF = 0
113 [ INFO ] 2018-03-21T00:00:49Z it.eng.rspa.odf.odfscheduler.job.ODMSynchJob Updated RDF = 0
114 [ INFO ] 2018-03-21T00:00:49Z it.eng.rspa.odf.odfscheduler.job.ODMSynchJob Added Datasets : 0
115 [ INFO ] 2018-03-21T00:00:49Z it.eng.rspa.odf.odfscheduler.job.ODMSynchJob Updated Datasets : 0
116 [ INFO ] 2018-03-21T00:00:49Z it.eng.rspa.odf.odfscheduler.job.ODMSynchJob Deleted Datasets : 0
117 [ INFO ] 2018-03-21T00:00:49Z it.eng.rspa.odf.odfscheduler.job.ODMSynchJob Deleted Datasets : 0
118 [ INFO ] 2018-03-21T00:00:49Z it.eng.rspa.odf.cache.MetadataCacheManager (nodeID=88)
119 [ INFO ] 2018-03-21T00:00:49Z it.eng.rspa.odf.cache.MetadataCacheManager -- Search- Matched Datasets in cache: 3
120 [ INFO ] 2018-03-21T00:00:49Z it.eng.rspa.odf.dcat.Serializer Writing model to file: /home/ubuntu/datasetDump/globalDump_node_88
121 [ INFO ] 2018-03-21T00:00:49Z it.eng.rspa.odf.dcat.Serializer Writing model to file: PTO_0055
122 [ INFO ] 2018-03-21T00:00:49Z it.eng.rspa.odf.dataset.DCATAPSerializer Writing model to file: /home/ubuntu/datasetDump/globalDump_node_88.zip
123 [ INFO ] 2018-03-21T00:00:50Z it.eng.rspa.odf.dcat.dump.DCATAPSerializer File writing completed in: PTO_0065
124 [ INFO ] 2018-03-21T01:03:40Z it.eng.rspa.odf.odfscheduler.job.ODMSynchJob Starting synch job for catalogue: 80
125 [ INFO ] 2018-03-21T01:03:40Z it.eng.rspa.odf.odfscheduler.job.ODMSynchJob Starting synchronization for node: Consiglio Regionale della Campania ID: 80
126 [ INFO ] 2018-03-21T01:03:40Z org.ckan.Connection CONNECTION: OPEN
127 [ INFO ] 2018-03-21T01:03:41Z it.eng.rspa.odf.odfscheduler.job.ODMSynchJob The ODMs Node with name: Consiglio Regionale della Campania and ID: 80 is ONLINE
128

```

1.3 End User Manual

This section provides the description of the End User Functionalities. Through the Open Data Catalogue a user can:

- Search datasets filtering by their metadata;
- Create graphical representation of dataset resources called *Datalet*;
- Execute SPARQL queries on RDF resources;
- View the federated ODMS in the platform.

1.3.1 Metadata Search

Each user can perform a dataset. Two types of search are provided by the GUI: a simple search or an advanced search.

Simple search

To perform a simple search the user should click on the search icon to perform the search on all of the federated datasets. The user could insert one or more keyword into the search bar to perform a filtered search. Moreover, the user could:

- select a tag from the tag-cloud to filter the search using the selected tag;
- search dataset by Categories.

Federated Open Data Catalogue Dataset Search SPARQL Search Federated Catalogues LOGIN

Search over **32780** dataset from **46** catalogues!

Insert comma separated keywords 

**profili di consumo
ambiente popolazione comune-pisa
trasparenza**

consigliere edifici medio ambiente
servizi urbanistica
università statistica
trasporti 21
energia anagrafe jamma incidenti bolzano
mobilità bilancio gfs circoscrizione
dipendenti citta-metropolitana-firenze

Browse by Categories

Agriculture, fisheries, forestry and food Economy and finance Education, culture and sport Energy Environment Government and public sector Health

International issues Justice, legal system and public safety Regions and cities Population and society Science and technology Transport

Credits



Advanced Search

To perform an advanced search the user should click on the expand icon. An advanced form appears to him/her where the user can fill one or all of the fields in order to filter the results.

The screenshot shows the search interface of the Federated Open Data Catalogue. At the top, there are links for Dataset Search, SPARQL Search, Federated Catalogues, and a LOGIN button. Below the header, a message says "Search over 32780 dataset from 46 catalogues!". The search form includes a "Filter 1" dropdown set to "All", a keyword input field with "water" typed in, and date range inputs for Release Date and Update Date. Under "Catalogues (46/46)", the "Multi Language" checkbox is checked. The "Source" language is set to English, and the "Targets" language is set to Deutsch, Français. The "Sort By" dropdown is set to "Title" and "Ascendent". The "Max results per page" is set to 25. A "RESET" button and a "SEARCH" button are at the bottom.

The advanced search functionality allows the user to search using a multilanguage approach provided by the platform taking advantage of EuroVoc thesaurus. The user should select one source language and one or more target language to use this functionality. The following picture shows an example of multilanguage search with keyword *water*, source language *English* and target language *Italian*.

This screenshot shows the search results for the keyword "water". The search interface is identical to the one above, with "water" in the keyword field and English as the source language. A dropdown menu is open under "Targets", showing "Italiano" selected. The results page displays 69 datasets found. Three datasets are listed, each with a title in Italian and a link to the dataset. The titles are "MODELLO WRF-ARW a 3km - Maschera Terra - Acqua - 2018...", "MODELLO WRF-ARW a 3km - Maschera Terra - Acqua - 2018...", and "MODELLO GFS a 50km - Maschera Terra - Acqua - 2018-02...". Each dataset entry includes a small preview image and download links for WMS, ZIP, and UNKNOWN formats.

Search Result

The result of both the simple search or the advanced search are a list of the dataset that match with the requested filter. Next figure illustrates the result of a search operation.

The screenshot shows the search results for the query "comune-di-firenze". The results are sorted by title and show 32780 datasets. The interface includes facets for Tags (e.g., ambiente, popolazione) and Formats (e.g., csv, json). A detailed view of a dataset titled "Comune di Rovereto: Stradario WFS 1.1.0 beta" is shown, along with its description and download links for GML, CSV, XLS, and PDF.

In this page the user can navigate results, he/she can change the order and the number of the results per page; moreover, he/she can filter the data using a *facet* approach. Different facets are available, in particular:

- Tags
- File Formats
- File Licenses
- Catalogues
- Categories

Dataset Detail

By clicking on a dataset in the search result page, the detailed presentation of all its metadata is showed to the user. The following picture shows an example of dataset detail.

The screenshot shows the detailed view of a dataset titled "Cognomi di tutti i residenti di Matera aggiornati al 31/08/2013". It includes sections for Tags (cognomi, residenti, statistica), Resources (JSON and CSV files), and Additional Info (Landing Page, Publisher, License, Release Date, Update Date, Frequency, Language, Rights Holder, Creator, Theme, Identifier, Other Identifier, Contact, and Spatial Coverage).

In this page the user can download the resources associated to the dataset by clicking on the download button; moreover, the user can create a graphical representation of the resources by clicking on the *create datalet* button.

1.3.2 Datalet Creation

A **Datalet** is a view WC, which is used to create rich, reusable visualization of open data. It was developed under the [ROUTE-TO-PA project](#). The datalet creator tool called *Datalet-Ecosystem Provider (DEEP)* was integrated with ODF in order to provide to users an open data visualization tool. For any further references about datalets please check <https://github.com/routetopaspod/wiki/Datalets>.

In order to create a datalet the user should follow this steps:

- Select fields
- Select the graphical representation

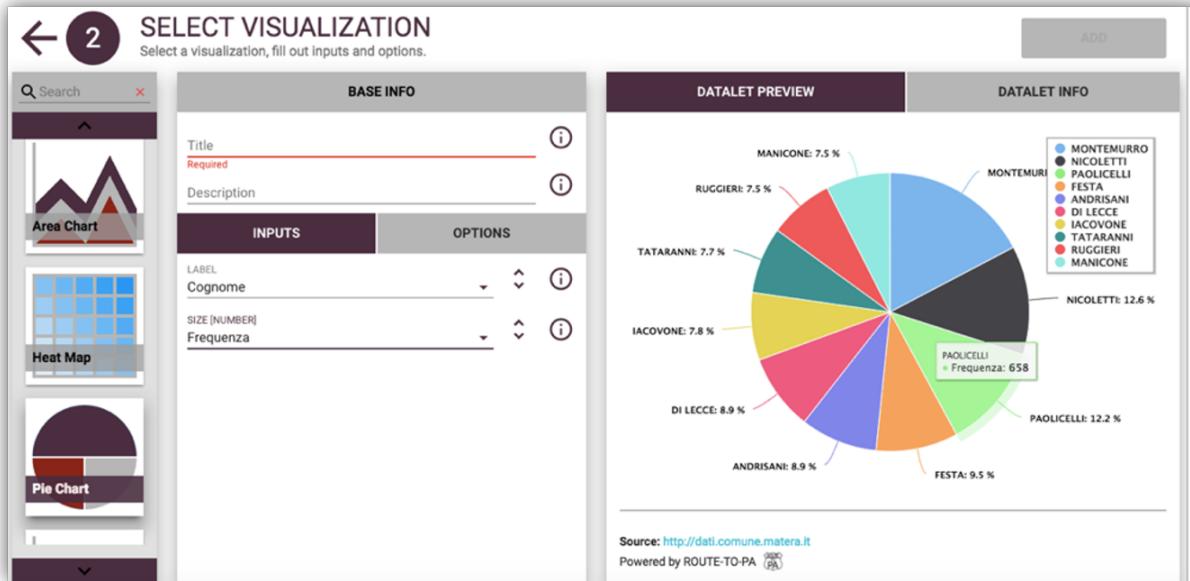
Selecting fields

The datalet creation process starts with the selection of the fields from the resource. In this page the user can add all or a subset of the original fields. Moreover, the user can also filter the data through a dedicated panel. The user should then click on the right arrow to continue the process.

| Cognome | Frequenza |
|------------|-----------|
| MONTEMURRO | 937 |
| NICOLETTI | 682 |
| PAOLICELLI | 658 |
| FESTA | 516 |
| ANDRISANI | 484 |

Select the graphical representation

The next step is to choose the graphical representation of the selected fields and the proper association among the selected fields and the chart inputs. The following picture depicts a pie chart example.



In order to show the datalet in the ODF environment the user should click on the *Add* button.

1.3.3 SPARQL Queries

This functionality allows the user to search over LOD downloaded from the federated dataset and stored into RDF4J triple store. In this page the user can write his SPARQL query and select the format of the output between *XML* or *JSON*.

The screenshot shows the 'Federated Open Data Catalogue' interface. At the top, there are navigation links: 'Dataset Search', 'SPARQL Search', 'Federated Catalogues', and a 'LOGIN' button. Below this is a 'Sparql Query' section with a text input containing the following SPARQL code:

```

1 PREFIX dc:<http://purl.org/dc/elements/1.1/>
2 SELECT ?object
3 WHERE {
4 ?subject ?predicate ?object
5 }
6 LIMIT 50

```

Below the query input, there is an 'Output Format' dropdown set to 'XML'. At the bottom right are 'CLEAR' and 'SEARCH' buttons.

The result of the query is showed to the user and he/she can download.

Federated Open Data Catalogue Dataset Search SPARQL Search Federated Catalogues LOGIN

Home / Sparql Query / Sparql Result

[BACK](#)

```

1 <?xml version="1.0"?>
2 <sparql xmlns="http://www.w3.org/2005/sparql-results#"
3 <head>
4   <variable name="object"/>
5 </head>
6 <result>
7   <result>
8     <binding name="object">
9       <literal datatype="xs:date">21/05/2013</literal>
10    </binding>
11   <result>
12     <binding name="object">
13       <literal datatype="xs:date">31/10/2017</literal>
14     </binding>
15   </result>
16   <result>
17     <binding name="object">
18       <literal datatype="xs:string">62</literal>
19     </binding>
20   </result>
21   <result>
22     <binding name="object">
23       <literal datatype="xs:string">M</literal>
24     </binding>
25   </result>
26   <result>
27     <binding name="object">
28       <uri>http://dati.inail.it/opendata/downloads/datisemestralimalattieprofessionali/rdf/Gestione.rdf/I</uri>
29

```

[COPY](#) [DOWNLOAD](#)

1.3.4 Catalogues overview

In this page all of the federated Catalogues are showed to user. The user can have a brief description of the Catalogue, check its country and category. Moreover, by clicking on the search button, the user can see all of the its datasets. The user can select between two views:

- Card

Federated Open Data Catalogue Dataset Search SPARQL Search Federated Catalogues LOGIN

| | |
|---|---|
| <p> Opendata Bari 🔍</p> <p>Datasets: 93 Last update: 2018-03-22 Category: Municipality</p> <p>Open data del comune di Bari</p> | <p> Open Data Catania 🔍</p> <p>Datasets: 49 Last update: 2018-03-22 Category: Municipality</p> <p>Dati Aperti del Comune di Catania</p> |
| <p> Open Data Regione Sicilia 🔍</p> <p>Datasets: 21 Last update: 2018-03-21 Category: Region</p> <p>Open Data della Regione Siciliana relativi ai settori di attività dell'amministrazione regionale e delle pubbliche amministrazioni regionali</p> | <p> Open Data Umbria 🔍</p> <p>Datasets: 361 Last update: 2018-03-21 Category: Region</p> <p>Repertorio regionale dei dati pubblici aperti dell'Umbria ai sensi dell'art.15 della l.r. n.8/2011, realizzato nell'ambito della Community Network regionale (CN-Umbria) di cui all'art.10 della l.r. n.8/2011. Il portale ha lo scopo di valorizzare il ...</p> |
| <p> Open Data Palermo 🔍</p> <p>Datasets: 826 Last update: 2018-03-22 Category: Municipality</p> <p>No description available</p> | <p> Opendata Sardegna 🔍</p> <p>Datasets: 618 Last update: 2018-03-22 Category: Region</p> <p>Con la deliberazione 4/2 del 5/2014 la Regione Autonoma della Sardegna ha avviato un processo finalizzato all'adozione di un Piano d'Azione regionale per la valorizzazione del patrimonio informativo pubblico della Regione e degli Enti Locali della ...</p> |

- Table

| Federated Open Data Catalogue | | | | | | Dataset Search | SPARQL Search | Federated Catalogues | LOGIN |
|-------------------------------|---------|---|---------------------|----------|---------------------|---|-------------------------------|--------------------------------------|--|
| Name | Country | Host | Category | Datasets | Last Update | | | | |
| Opendata Bari | IT | http://opendata.comune.bari.it | Municipality | 93 | 2018-03-22 11:08:47 |  | | | |
| Open Data Catania | IT | http://opendata.comune.catania.gov.it/ | Municipality | 49 | 2018-03-22 11:08:48 |  | | | |
| Open Data Regione Sicilia | IT | https://dati.regione.sicilia.it/ | Region | 21 | 2018-03-21 17:18:53 |  | | | |
| Open Data Umbria | IT | http://dati.umbria.it/ | Region | 361 | 2018-03-21 16:38:58 |  | | | |
| Open Data Palermo | IT | https://www.comune.palermo.it/opendata_did.php | Municipality | 826 | 2018-03-22 09:49:16 |  | | | |
| Opendata Sardegna | IT | http://dati.regione.sardegna.it | Region | 618 | 2018-03-22 10:31:22 |  | | | |
| Fiware CKAN Catalogue | - | https://data.lab.fiware.org | Private Institution | 2934 | 2018-03-21 16:39:42 |  | | | |
| Open Data Milano | IT | http://dati.comune.milano.it/ | Municipality | 292 | 2018-03-22 10:39:41 |  | | | |
| Dati Toscana | IT | http://dati.toscana.it/ | Region | 2844 | 2018-03-21 17:07:48 |  | | | |
| Regione Marche Open Data | IT | http://goodpa.regione.marche.it | Region | 27 | 2018-03-22 10:52:25 |  | | | |
| Open Data Regione Lazio | IT | http://dati.lazio.it/catalog | Region | 348 | 2018-03-21 17:56:16 |  | | | |
| Open Data Regione Puglia | IT | http://www.dataset.puglia.it/ | Region | 111 | 2018-03-21 17:57:04 |  | | | |