GUIDELINES FOR TCS Metadata Mapping

EPOS-IP project

July 27, 2018

Table of Contents

1. Introduction

This document is a guideline for TCS DDSS providers for the metadata mapping process, in order to create a description of TCS assets in EPOS-DCAT-AP RDF/Turtle format.

At this mapping stage, the main EPOS-DCAT-AP entities that should be filled and interlinked, are: **Dataset**, **Distribution**, **WebService, Operation, Person** and **Organisation**.

A dataset can have one or more distributions. Two key pieces of information are required to access a distribution of a dataset, the address of this distribution and its type definition (which in turn determines the procedure to access the distribution). If the distribution can be accessed via web service all its information should be specified. A web service can have one or more operations which are represented by endpoint, parameters, etc.
One or more persons could be defined, and it also could assume various roles in different entities.
Example: a contactPoint role for a Dataset and WebService, or scientic contact for an Organisation.
One or more organisations could be created and can be publisher for a dataset or provider for a web service.

The mapping process is made via following steps:

1. Create EPOS-DCAT-AP RDF/Turtle file;
2. Validate EPOS-DCAT-AP RDF/Turtle file;
3. Upload EPOS-DCAT-AP RDF/Turtle file on GitHub;

The first step can be done by using the EPOS Web Metadata Editor or with the manual method, creating files from scratch (metadata practitioners).

The Web Metadata Editor allows to create EPOS-DCAT-AP RDF/Turtle file in two different ways:

- Reuse existing EPOS-DCAT-AP XML file, available on GitHub https://github.com/epos-eu/EPOS-DCAT-AP/tree/master/examples

OR

- Create EPOS-DCAT-AP RDF/Turtle description from scratch.

An example of EPOS-DCAT-AP in RDF/Turtle format is available on GitHub (https://github.com/epos-eu/EPOS-DCAT-AP/blob/EPOS-DCAT-AP-shapes/examples/EPOS-DCAT-AP_example.ttl). It contains all required entities with additional comments which help to fill the properties.

Please, pay attention to the following points:

- **Use unique identifier** (e.g., DOI, ORCID, ScopusID, PIC). In the case of you don't have for your entity a persisten identifier you can use this format to create one:
  Prefix + CommunityName + ResourceType + someID
  For instance:
  - https://www.epos-eu.org/epos-dcat-ap/Seismology/WebService/001/
  - https://www.epos-eu.org/epos-dcat-ap/Seismology/Dataset/001/
- **DDSS-ID**: this field contains the DDSS-ID. Please, note that this ID should be reported in Dataset entity and removed from WebService entity. Furthermore, you should check the correct ID from facets tree on Github (https://github.com/epos-eu/documentation/blob/master/2018.07.23%20-%20EPOS%20facets.pdf);
- **Link properly the entities between them** as shown in the next paragraphs.

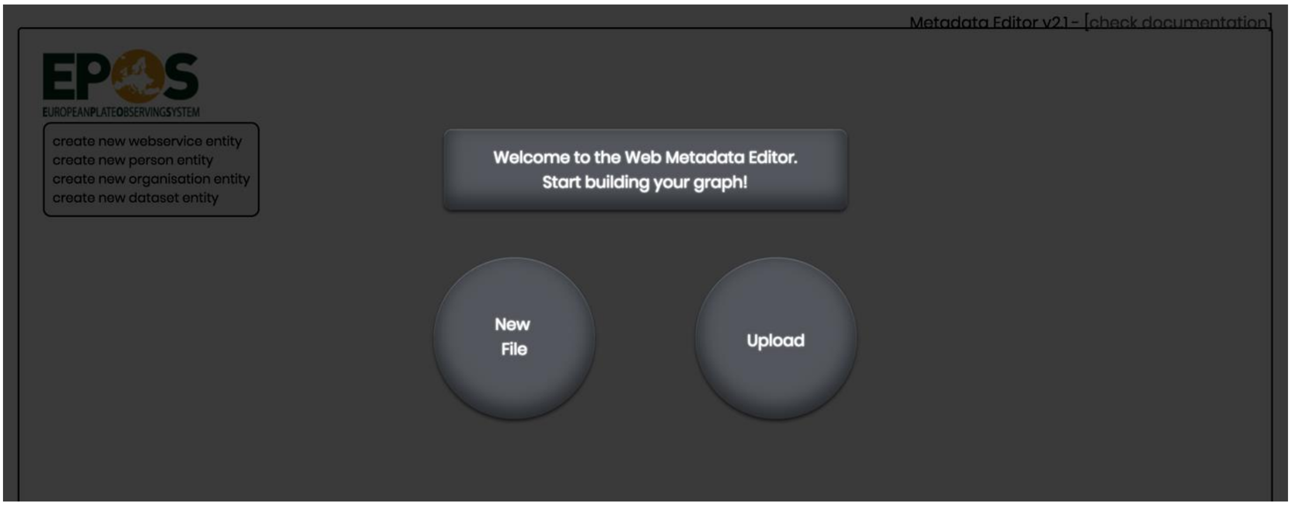
Nota bene: in case you have already uploaded the RDF/turtle files on GitHub, please review and change them if they do not contain all the entities and properties as shown here https://github.com/epos-eu/EPOS-DCAT-AP/blob/EPOS-DCAT-AP-shapes/examples/EPOS-DCAT-AP_example.ttl

2. Create the EPOS-DCAT-AP RDF/Turtle file by using EPOS Web Metadata Editor

The Metadata Editor (http://epos.cineca.it/apache/mde/public/index.php) allows the TCS to describe their metadata in EPOS-DCAT-AP RDF/Turtle format. It is a web application which permits to manage EPOS entities, providing a graph representation of them.
With this release it is possible to manage **Person**, **Organisation**, **Dataset** and **WebService** entities. **Distribution** and **Operation** entities could be created only when both Dataset and WebService are represented.

There are two options in the Metadata Editor homepage (Fig. 1):

- create new EPOS-DCAT-AP in RDF/Turtle format
- convert an existing EPOS-DCAT-AP from XML format to RDF/Turtle format


Metadata Editor v2.1 - [check documentation]
EPOS
EUROPEANPLATEOBSERVINGSYSTEM
create new webservice entity
create new person entity
create new organisation entity
create new dataset entity
New File
Upload

Welcome to the Web Metadata Editor.
Start building your graph!

Figure 1 - EPOS Metadata Editor - Homepage

2.1. Create new EPOS-DCAT-AP in RDF/Turtle format

A blank page appears by clicking on the “New File” button. At the user is given the possibility to define new entities (Fig. 2).

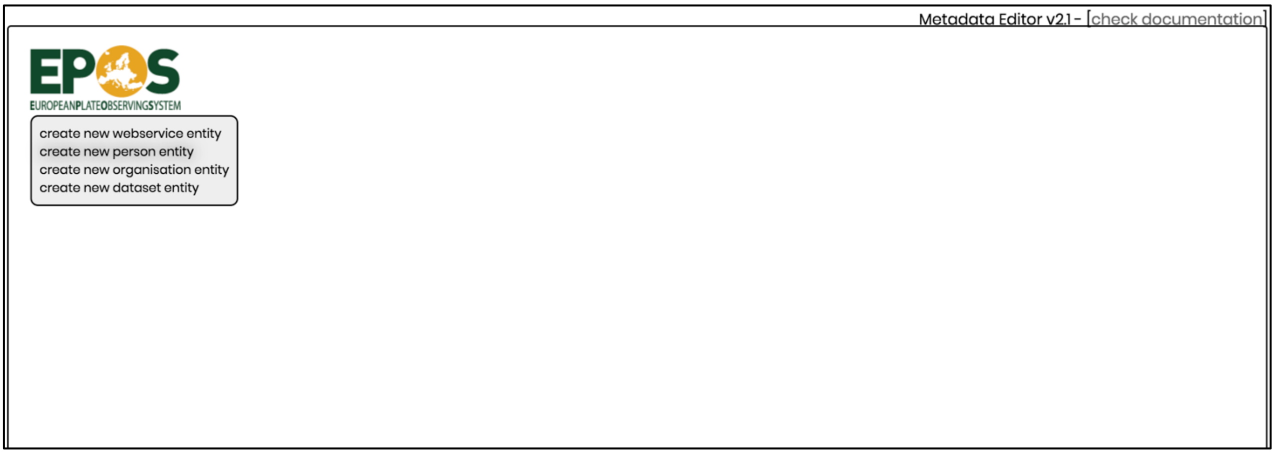
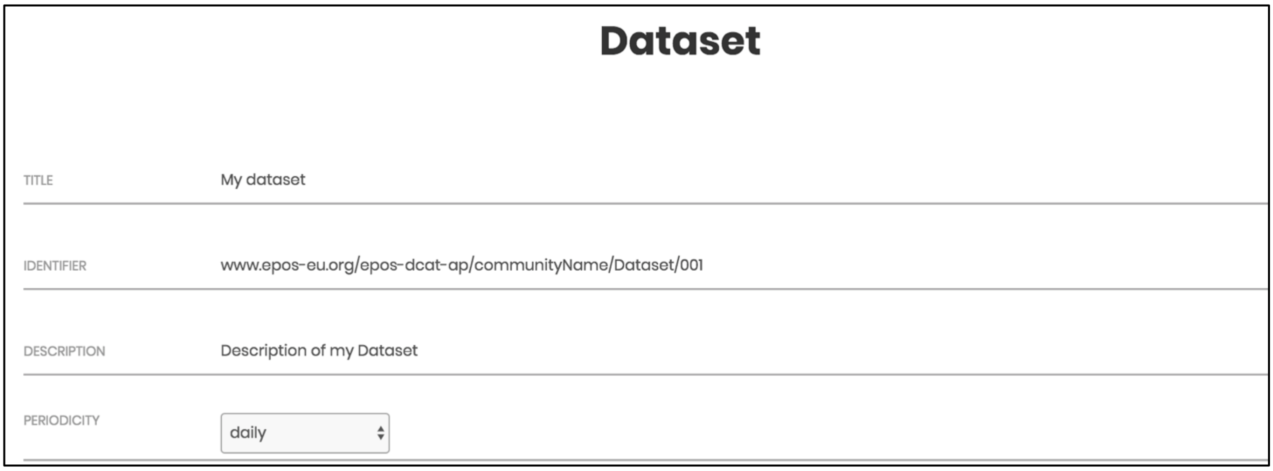

Metadata Editor v2.1 - [check documentation]
EPOS
EUROPEANPLATEOBSERVINGSYSTEM
create new webservice entity
create new person entity
create new organisation entity
create new dataset entity

Figure 2 - EPOS Metadata Editor - Screen to create new entities

Use the upper left menu to create a new Dataset entity (Fig. 3).



Dataset

TITLE	My dataset
IDENTIFIER	www.epos-eu.org/epos-dcat-ap/communityName/Dataset/001
DESCRIPTION	Description of my Dataset
PERIODICITY	daily

Figure 3 - EPOS Metadata Editor - New Dataset entity

After the completion of your Dataset information, click on *submit* and the graph starts populating with the defined node. By left-clicking on the node, its contextual menu provides additional features (Fig. 4).

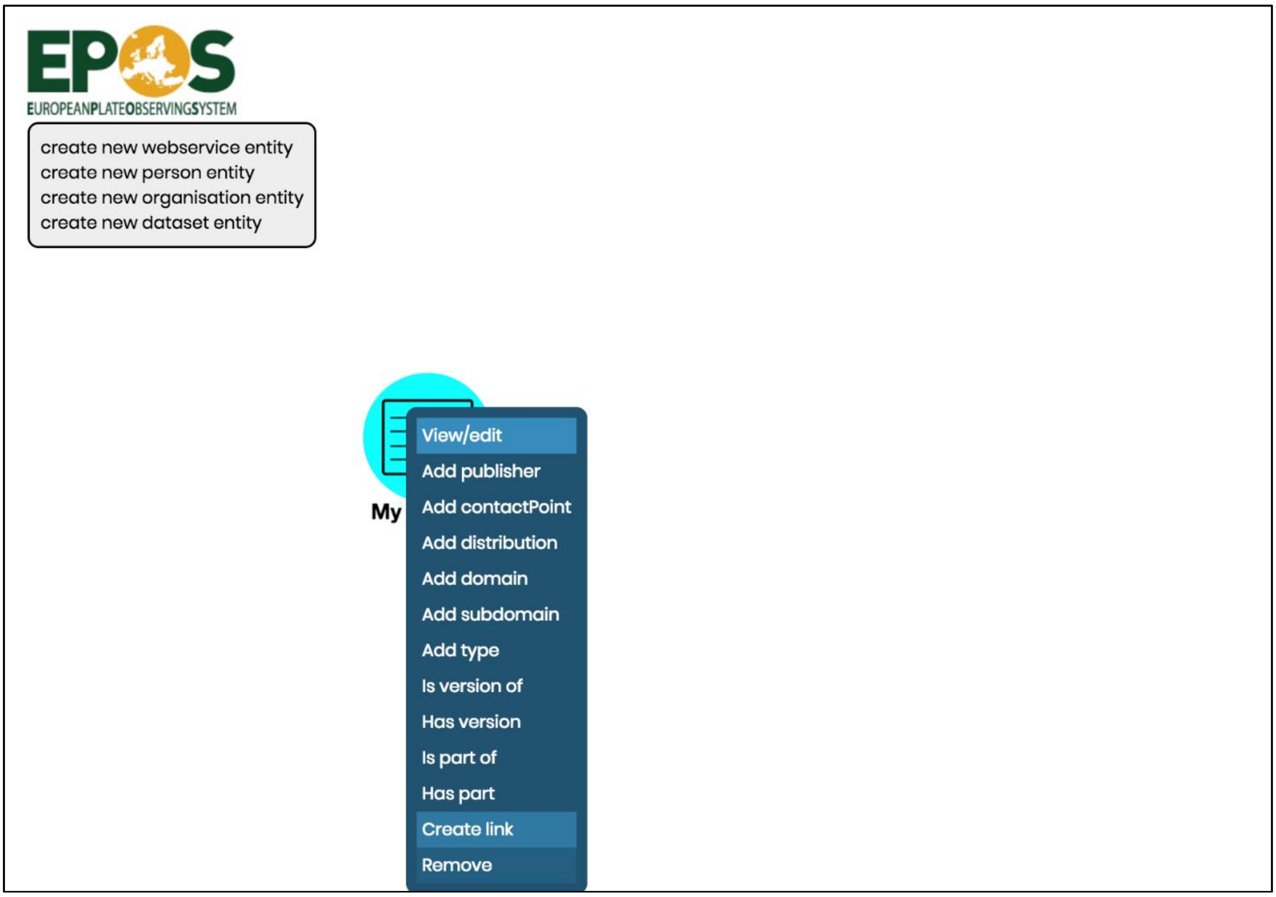
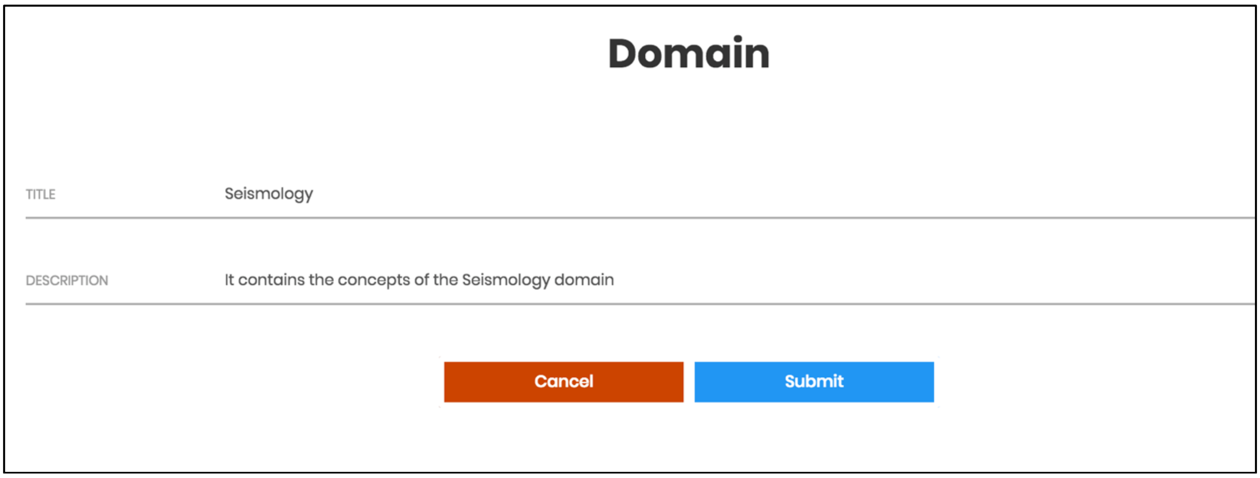


Figure 4 - EPOS Metadata Editor – Dataset node and its contextual menu

For the Dataset created, you can define the *domain* and *subdomain* by left-clicking on the node (fig. 5 - 6 – 7).

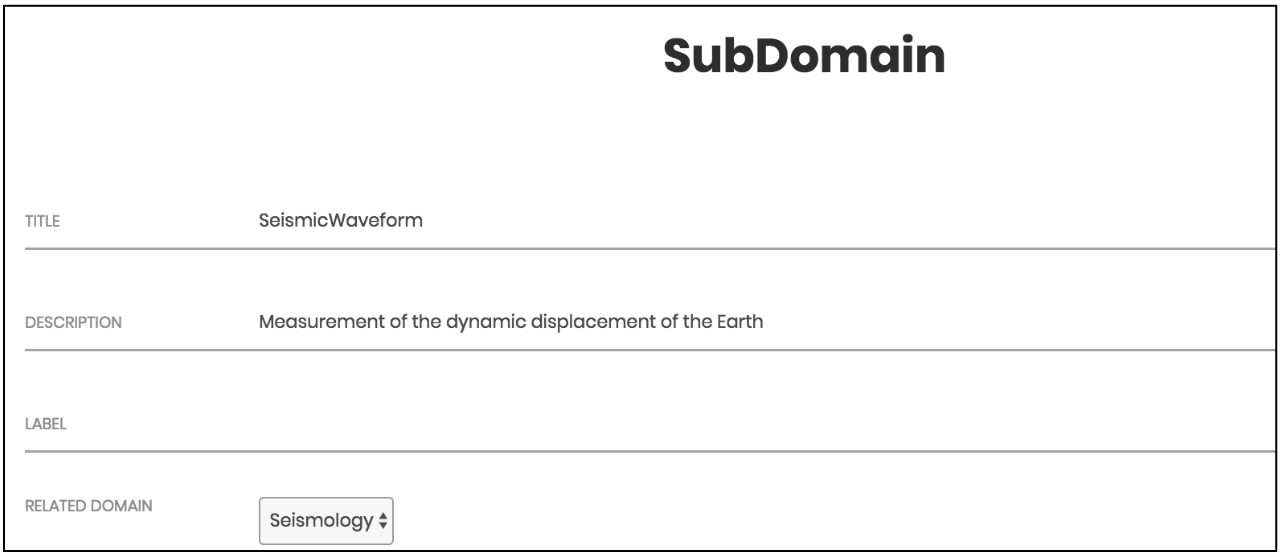


Domain

TITLE Seismology

DESCRIPTION It contains the concepts of the Seismology domain

Figure 5 - EPOS Metadata Editor – New Domain



SubDomain

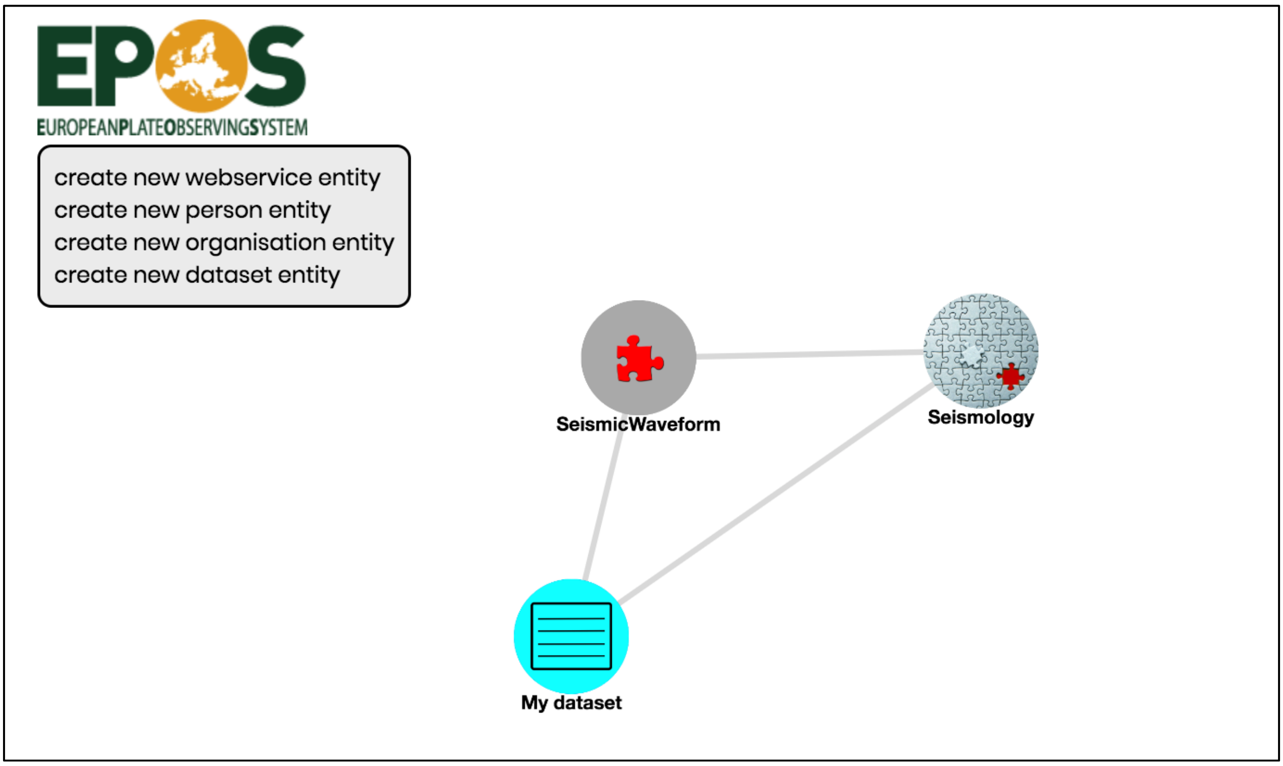
TITLE SeismicWaveform

DESCRIPTION Measurement of the dynamic displacement of the Earth

LABEL

RELATED DOMAIN Seismology

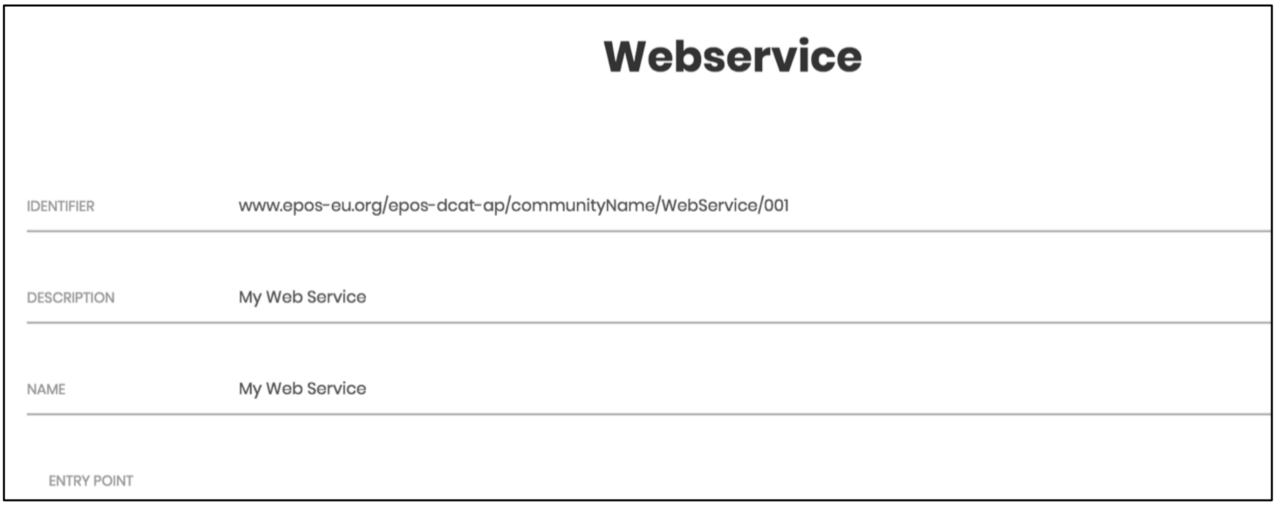
Figure 6 - EPOS Metadata Editor – New SubDomain



EPOS
EUROPEANPLATEOBSERVINGSYSTEM
create new webservice entity
create new person entity
create new organisation entity
create new dataset entity
SeismicWaveform
Seismology
My dataset

Figure 7 - EPOS Metadata Editor – Dataset node linked to Domain and SubDomain nodes

The next step is to create the WebService entity (Fig. 8).



Webservice

IDENTIFIER www.epos-eu.org/epos-dcat-ap/communityName/WebService/001

DESCRIPTION My Web Service

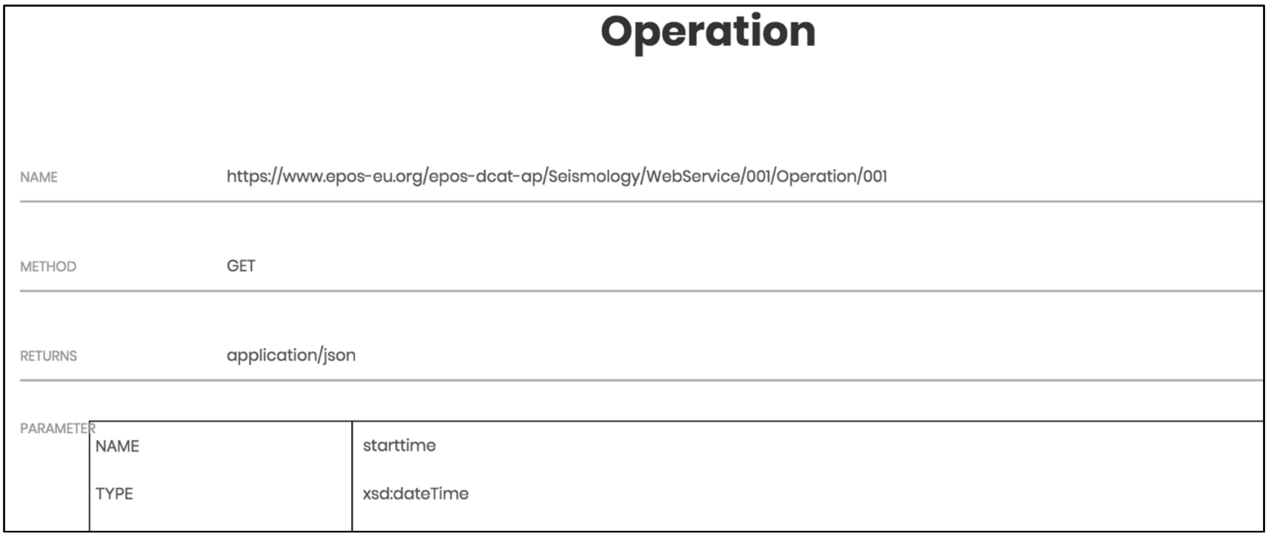
NAME My Web Service

Figure 8 - EPOS Metadata Editor – New WebService entity

For the Web Service created, you can define one or more Operations by left-clicking on the node and selecting "add operation". In this way it is possible to define enpoint and parameters to query the web service (Fig. 9 - 10).

Where possible, fill the following fields for each parameter:

Field	Description
name	The name of the parameter as required by web service specifications
type	The type of the parameter. The possible values for this field are: "xsd:string" "xsd:boolean" "xsd:date" " xsd:dateTime" "xsd:decimal" "xsd:double" "xsd:float" "xsd:int" "xsd:integer" "xsd:long".
label	The label is a short string used to describe the meaning of the parameter to the GUI user.
min Value	The minimum value of the parameter
max Value	The maximum value of the parameter
required	This field contains true if the property is required, false otherwise.
default value	This field contains the default value of the parameter. It is very important to fill this field for the proper functioning of the demonstrator.
param value	This field represents one of the possible values which should be used in the web service query; it could be repeated as many times as needed.



Operation

NAME	https://www.epos-eu.org/epos-dcat-ap/Seismology/WebService/001/Operation/001
METHOD	GET
RETURNS	application/json
PARAMETER	NAME: starttime TYPE: xsd:dateTime

Figure 9 - EPOS Metadata Editor – New Operation entity

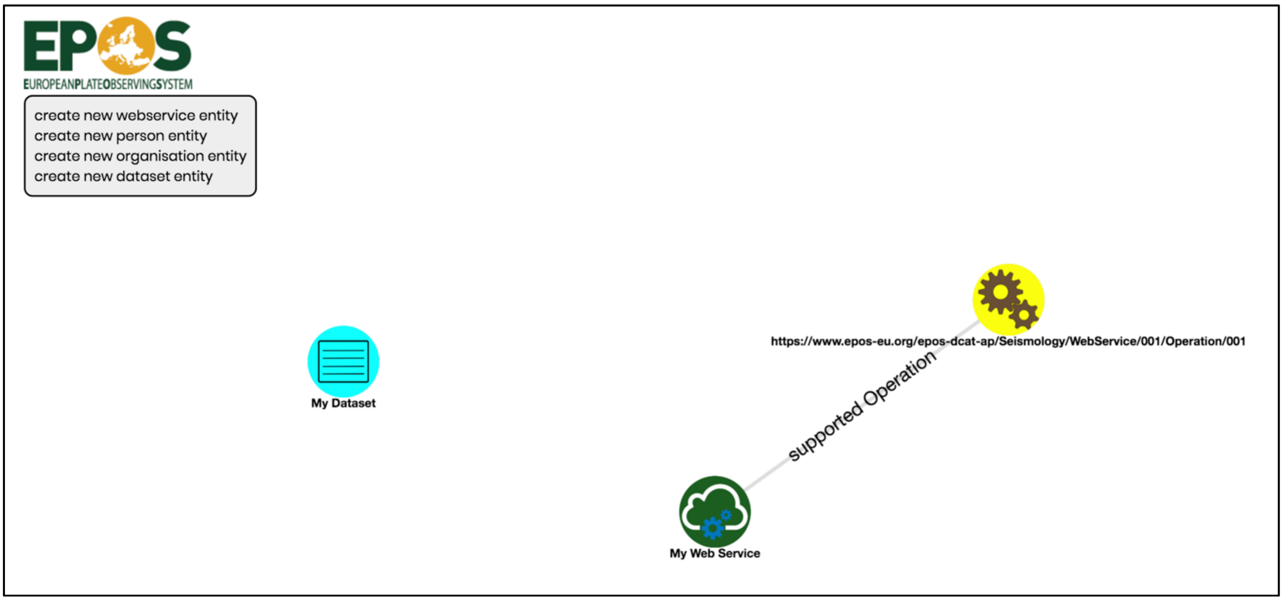

EPOS
EUROPEANPLATEOBSERVINGSYSTEM
create new webservice entity
create new person entity
create new organisation entity
create new dataset entity
My Dataset
https://www.epos-eu.org/epos-dcat-ap/Seismology/WebService/001/Operation/001
supported Operation
My Web Service

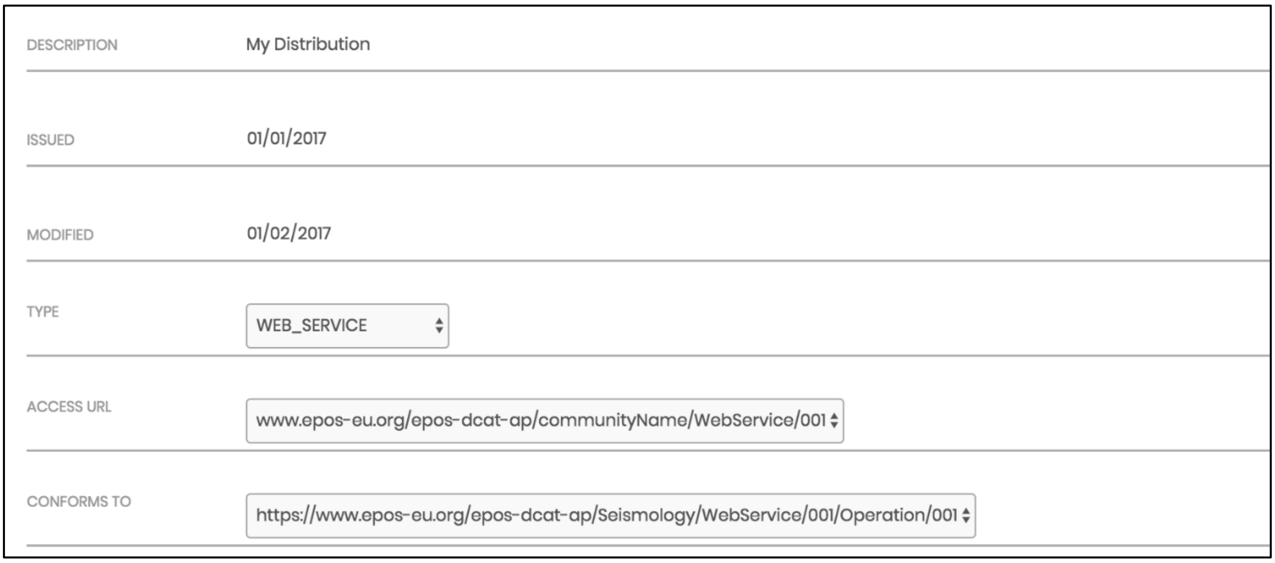
Figure 10 - EPOS Metadata Editor – Operation node linked to the WebService node

For the Web Service created, you can also define domain and subdomain as shown for Dataset (Fig. 5 - 6).

Once the Dataset and Web Service (with its operation(s)) are created, link these entities by creating the Distribution entity (Fig. 11 - 12).
If Distribution is associated to the webservice you should select:

- the web service ID for "*accessURL*”;
- the ID of Operation for "*conformsTo*".

Otherwise, if the distribution type is a downloadable file you should put the URL in *downloadURL* field.



DESCRIPTION	My Distribution
ISSUED	01/01/2017
MODIFIED	01/02/2017
TYPE	WEB_SERVICE
ACCESS URL	www.epos-eu.org/epos-dcat-ap/communityName/WebService/001
CONFORMS TO	https://www.epos-eu.org/epos-dcat-ap/Seismology/WebService/001/Operation/001

Figure 11 - EPOS Metadata Editor – New Distribution entity

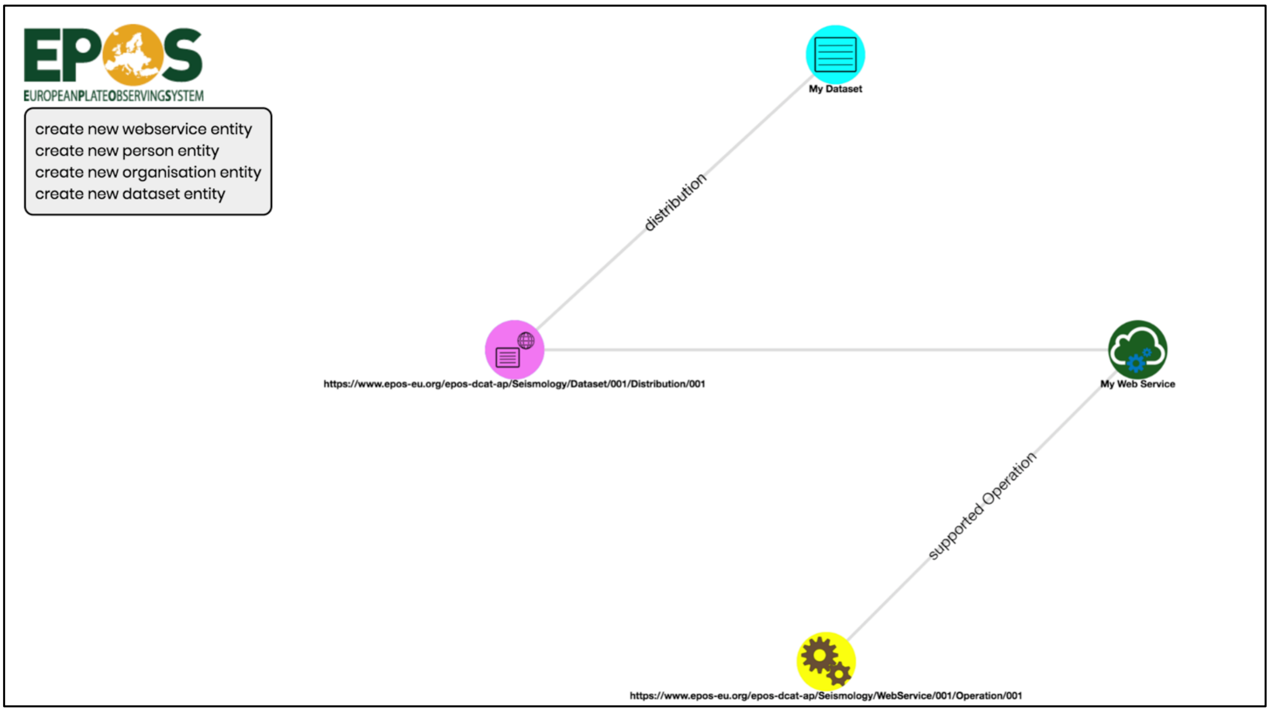
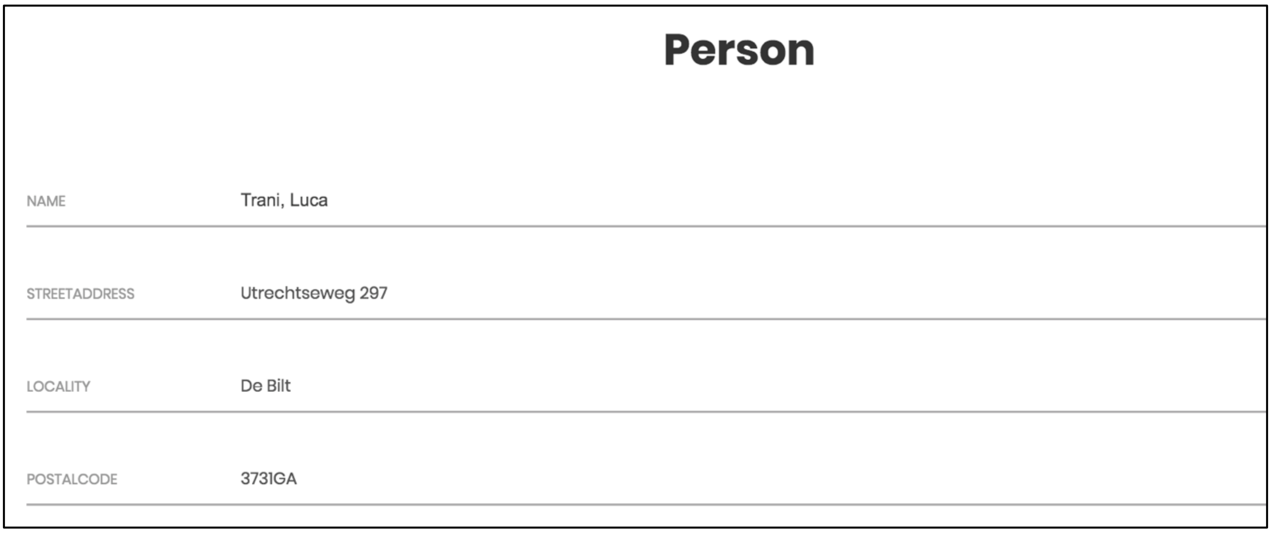

EPOS
EUROPEANPLATEOBSERVINGSYSTEM
create new webservice entity
create new person entity
create new organisation entity
create new dataset entity
My Dataset
distribution
https://www.epos-eu.org/epos-dcat-ap/Seismology/Dataset/001/Distribution/001
My Web Service
supported Operation
https://www.epos-eu.org/epos-dcat-ap/Seismology/WebService/001/Operation/001

Figure 12 – EPOS Metadata Editor – Distribution node linked to the WebService and Dataset nodes

Person entity can be created by left-clicking on the node (Dataset or WebService) choosing "add contactPoint" from the contextual menu (Fig. 13 – 14).

Person contains contact information that can be used to send comments to the Dataset and the WebServices associated.



Person

NAME Trani, Luca

STREETADDRESS Utrechtseweg 297

LOCALITY De Bilt

POSTALCODE 3731GA

Figure 13 – EPOS Metadata Editor – New Person entity

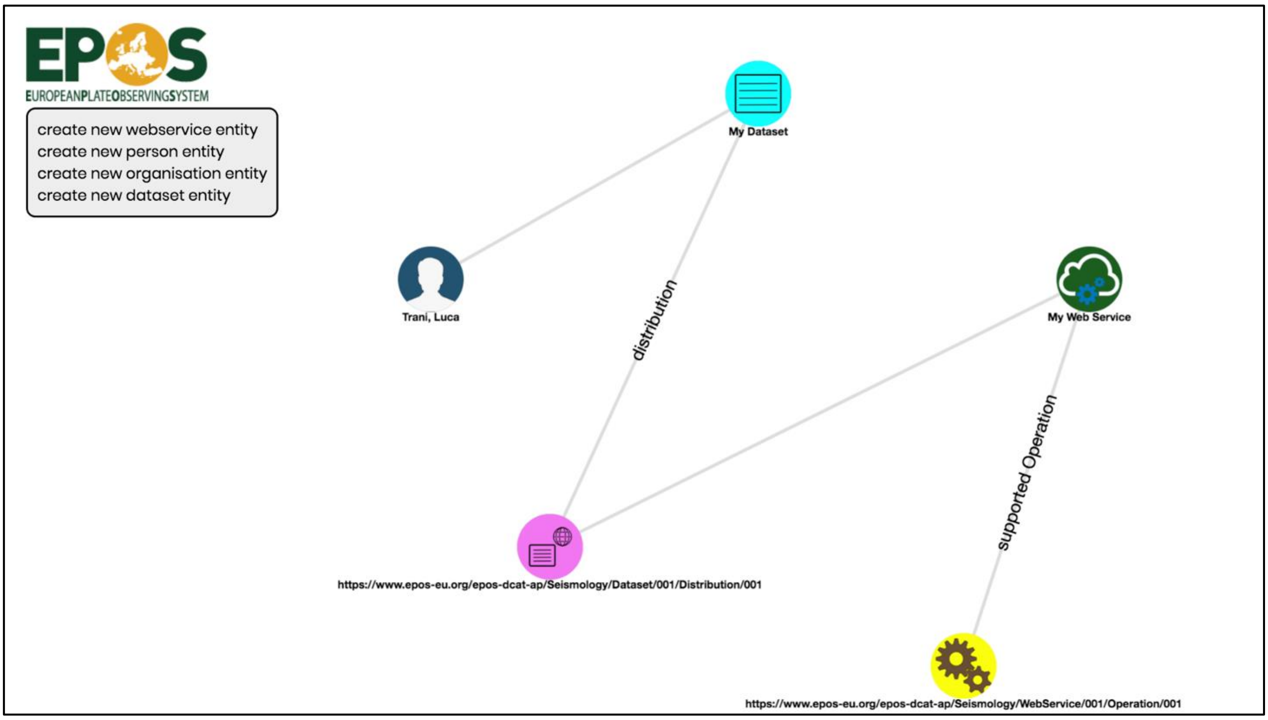
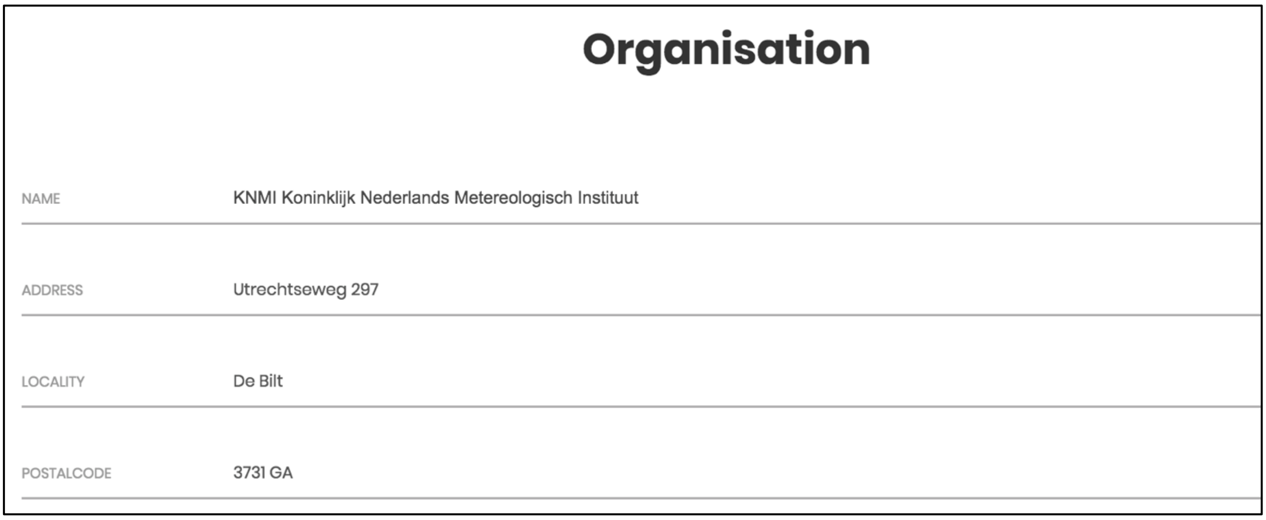

EPOS
EUROPEANPLATEOBSERVINGSYSTEM
create new webservice entity
create new person entity
create new organisation entity
create new dataset entity
My Dataset
Trani, Luca
distribution
My Web Service
supported Operation
https://www.epos-eu.org/epos-dcat-ap/Seismology/Dataset/001/Distribution/001
https://www.epos-eu.org/epos-dcat-ap/Seismology/WebService/001/Operation/001

Figure 14 - EPOS Metadata Editor – Person node linked to the Dataset as contact point

As Person, it is possible to create the Organisation entity which could be a publisher for a Dataset or a provider for a WebService (Fig. 15 – 16).



Organisation

NAME	KNMI Koninklijk Nederlands Metereologisch Instituut
ADDRESS	Utrechtseweg 297
LOCALITY	De Bilt
POSTALCODE	3731 GA

Figure 15 - EPOS Metadata Editor – New Organisation entity

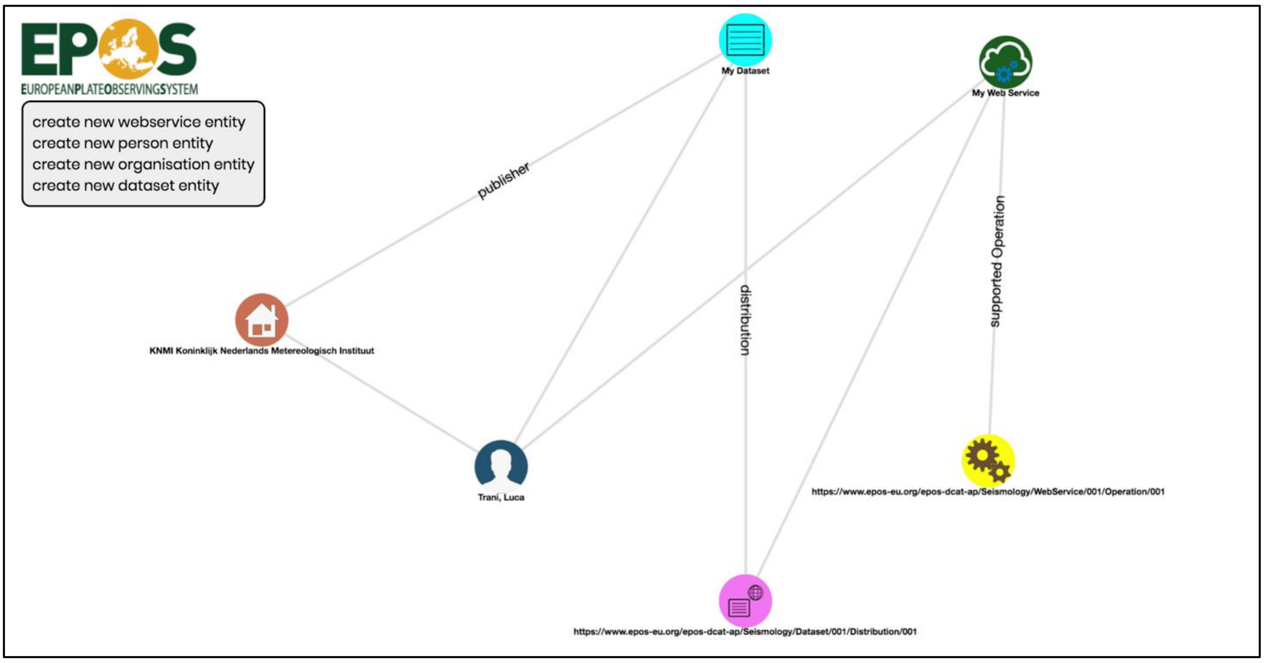

EPOS
EUROPEANPLATEOBSERVINGSYSTEM
create new webservice entity
create new person entity
create new organisation entity
create new dataset entity
My Dataset
My Web Service
publisher
distribution
supported Operation
KNMI Koninklijk Nederlands Metereologisch Instituut
Trani, Luca
https://www.epos-eu.org/epos-dcat-ap/Seismology/WebService/001/Operation/001
https://www.epos-eu.org/epos-dcat-ap/Seismology/Dataset/001/Distribution/001

Figure 16 - EPOS Metadata Editor – Organisation node linked to the Dataset as the publisher

2.2. Convert an existing EPOS-DCAT-AP from XML format to RDF/Turtle format

The Metadata Editor allows to create EPOS-DCAT-AP RDF/Turtle file by choosing the *upload* option from the homepage (Fig. 1). A user can upload an existing EPOS-DCAT-AP XML file from local directory or using the URL.
Note that:

- FILE option: an EPOS-DCAT-AP XML valid local file is required. No different files/formats are allowed.
- URL option: the URL must point to an already valid EPOS-DCAT-AP XML file, not to a page containing an EPOS-DCAT-AP XML file.

For instance:
Wrong URL: *https://github.com/epos-eu/EPOS-DCAT-AP/blob/master/examples/WP08/EPOS-DCAT- AP_WP08_ODC.xml;*

Right URL**:** https://raw.githubusercontent.com/epos-eu/EPOS-DCAT-AP/master/examples/WP08/EPOS- DCAT-AP_WP08_ODC.xml

All the entities (Person, Organisation, Webservice) contained in the file/URL are represented as nodes of a graph as shown in the example of figure 17.
All the functionalities illustred in the previous paragraphs, can be used to create and manage the entity in order to produce an EPOS-DCAT-AP RDF/Turtle file.

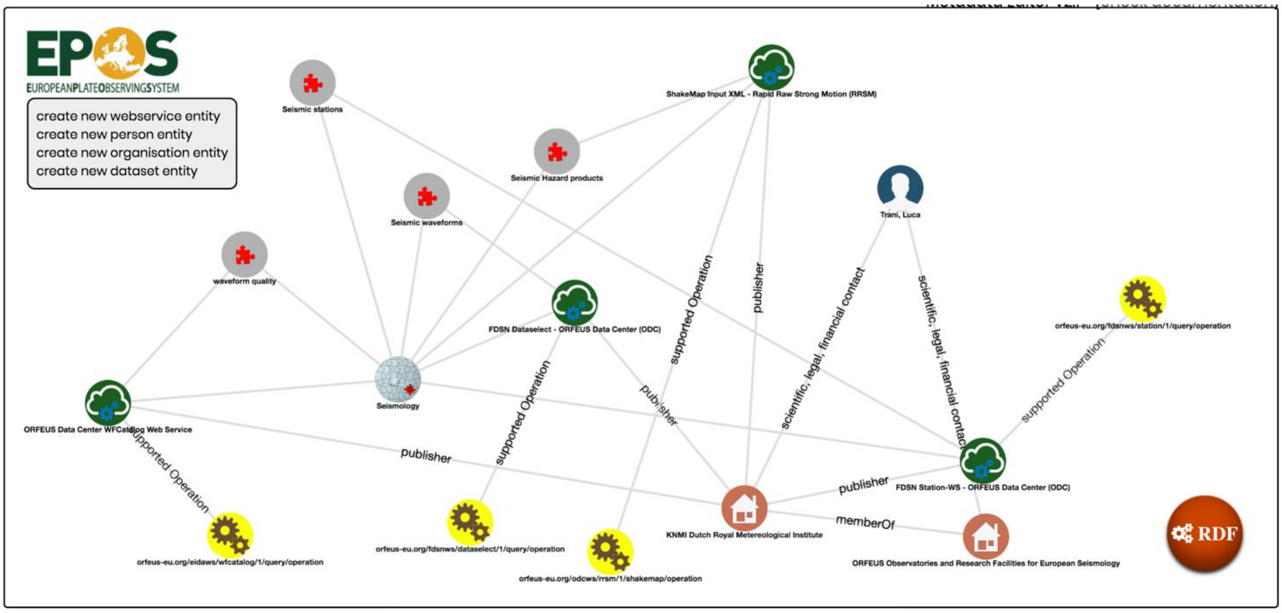
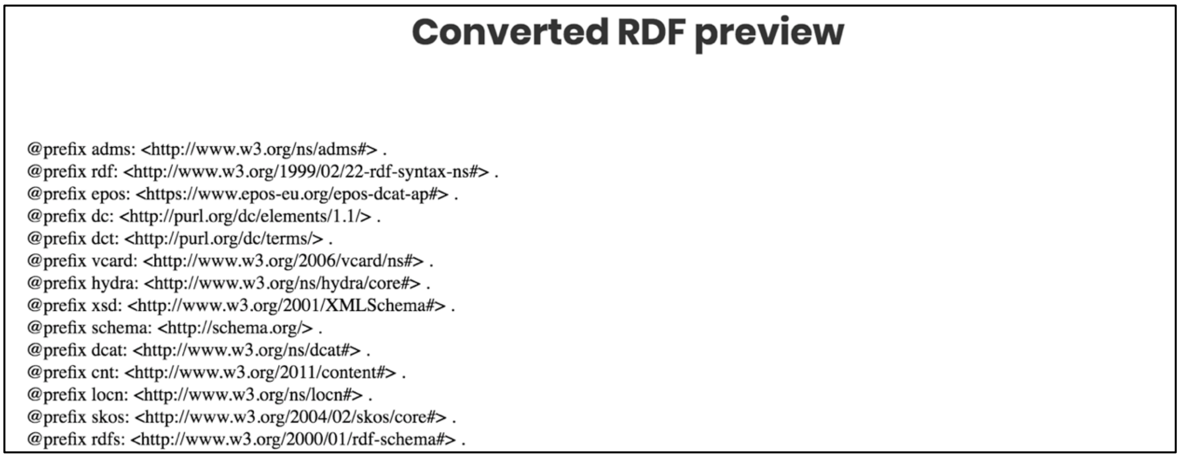

EPOS
EUROPEANPLATEOBSERVINGSYSTEM
create new webservice entity
create new person entity
create new organisation entity
create new dataset entity
Seismic stations
ShakeMap Input XML - Rapid Raw Strong Motion (RRSM)
Seismic Hazard products
Seismic waveforms
Trani, Luca
waveform quality
supported Operation
publisher
scientific, legal, financial contact
scientific, legal, financial contact
FDSN Dataselect - ORFEUS Data Center (ODC)
orfeus-eu.org/fdsnws/station/1/query/operation
Seismology
publisher
supported Operation
supported Operation
ORFEUS Data Center WFCatalog Web Service
supported Operation
publisher
publisher
FDSN Station-WS - ORFEUS Data Center (ODC)
memberOf
KNMI Dutch Royal Metereological Institute
orfeus-eu.org/eidaws/wfcatalog/1/query/operation
orfeus-eu.org/fdsnws/dataselect/1/query/operation
orfeus-eu.org/odcws/rrsm/1/shakemap/operation
ORFEUS Observatories and Research Facilities for European Seismology
RDF

Figure 17 - EPOS Metadata Editor – Upload existing EPOS-DCAT-AP XML file

2.3. Save EPOS-DCAT-AP in RDF/Turtle format

After the definition of the entities needed, by clicking on the RDF button, an overlay pops up showing a preview of the EPOS-DCAT-AP RDF/Turtle file (Fig. 18).
The created file is downloadable.



Converted RDF preview

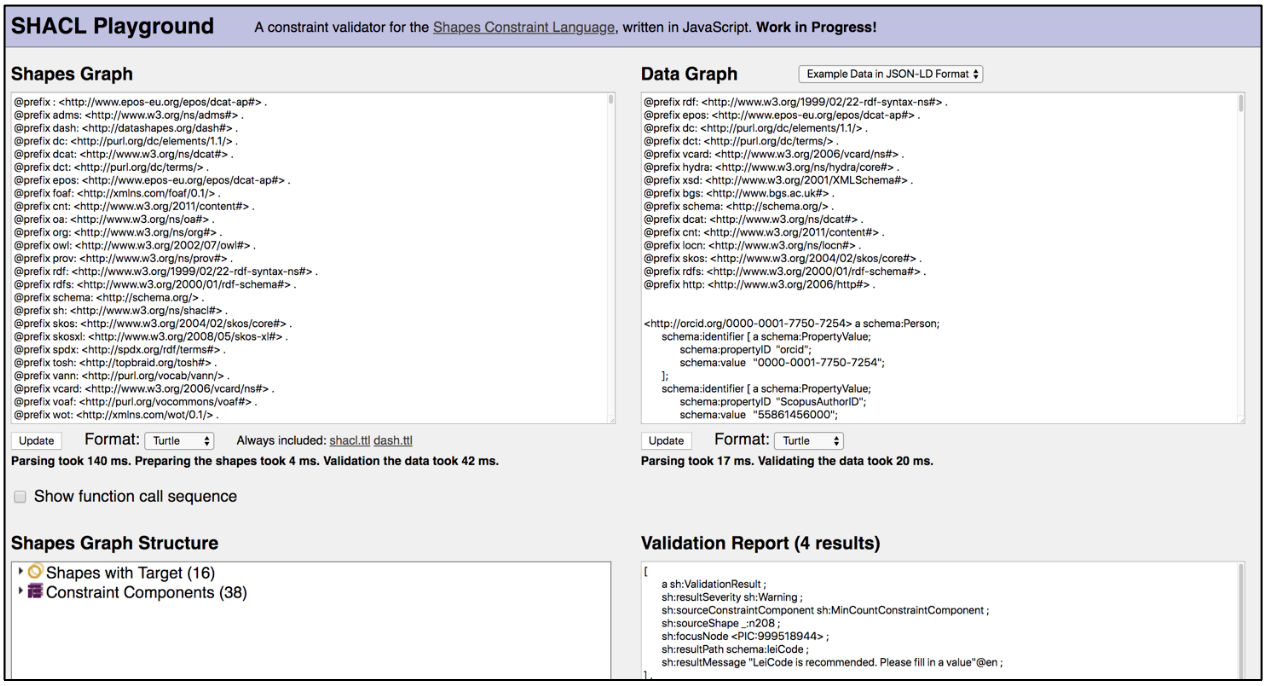
```
@prefix adms: <http://www.w3.org/ns/adms#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix epos: <https://www.epos-eu.org/epos-dcat-ap#> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .
@prefix dct: <http://purl.org/dc/terms/> .
@prefix vcard: <http://www.w3.org/2006/vcard/ns#> .
@prefix hydra: <http://www.w3.org/ns/hydra/core#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix schema: <http://schema.org/> .
@prefix dcat: <http://www.w3.org/ns/dcat#> .
@prefix cnt: <http://www.w3.org/2011/content#> .
@prefix locn: <http://www.w3.org/ns/locn#> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
```

Figure 18 - EPOS Metadata Editor – EPOS-DCAT-AP RDF/Turtle preview

3. Validate the RDF/Turtle file

In order to validate RDf/Turtle file, this online tool http://shacl.org/playground/ (Fig. 19) can be used as follows:

1. Copy and paste the content of the EPOS-DCAT-AP model file (available at https://raw.githubusercontent.com/epos-eu/EPOS-DCAT-AP/EPOS-DCAT-AP-shapes/epos-dcat-ap_shapes.ttl) into the “Shapes Graph” upper left box;
2. Click on the "UPDATE" button; a red notification will advise you when syntactic errors are reported;
3. Copy and paste the content of your RDF/Turtle file in the “Data Graph” upper right box;
4. Select Turtle format from drop-down list;
5. Click on the "UPDATE" button; a red notification will advise you when syntactic errors are reported;
6. The “Validation Report” box shows the result of validation (compliant to EPOS-DCAT-AP model). It reports violations (sh:Violation) and warnings (sh:Warning). All violations must be fixed in order to have a valid RDF/Turtle file.


SHACL Playground
A constraint validator for the Shapes Constraint Language, written in JavaScript. Work in Progress!
Shapes Graph
Data Graph
Example Data in JSON-LD Format
Update
Format: Turtle
Always included: shacl.ttl dash.ttl
Update
Format: Turtle
Show function call sequence
Shapes Graph Structure
Shapes with Target (16)
Constraint Components (38)
Validation Report (4 results)

```
@prefix : <http://www.epos-eu.org/epos/dcat-ap#> .
@prefix adms: <http://www.w3.org/ns/adms#> .
@prefix dash: <http://datashapes.org/dash#> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .
@prefix dcat: <http://www.w3.org/ns/dcat#> .
@prefix dct: <http://purl.org/dc/terms/> .
@prefix epos: <http://www.epos-eu.org/epos/dcat-ap#> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix cnt: <http://www.w3.org/2011/content#> .
@prefix oa: <http://www.w3.org/ns/oa#> .
@prefix org: <http://www.w3.org/ns/org#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix prov: <http://www.w3.org/ns/prov#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix schema: <http://schema.org/> .
@prefix sh: <http://www.w3.org/ns/shacl#> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix skosxl: <http://www.w3.org/2008/05/skos-xl#> .
@prefix spdx: <http://spdx.org/rdf/terms#> .
@prefix tosh: <http://topbraid.org/tosh#> .
@prefix vann: <http://purl.org/vocab/vann/> .
@prefix vcard: <http://www.w3.org/2006/vcard/ns#> .
@prefix voaf: <http://purl.org/vocommons/voaf#> .
@prefix wot: <http://xmlns.com/wot/0.1/> .
```

Parsing took 140 ms. Preparing the shapes took 4 ms. Validation the data took 42 ms.

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix epos: <http://www.epos-eu.org/epos/dcat-ap#> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .
@prefix dct: <http://purl.org/dc/terms/> .
@prefix vcard: <http://www.w3.org/2006/vcard/ns#> .
@prefix hydra: <http://www.w3.org/ns/hydra/core#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix bgs: <http://www.bgs.ac.uk#> .
@prefix schema: <http://schema.org/> .
@prefix dcat: <http://www.w3.org/ns/dcat#> .
@prefix cnt: <http://www.w3.org/2011/content#> .
@prefix locn: <http://www.w3.org/ns/locn#> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix http: <http://www.w3.org/2006/http#> .


<http://orcid.org/0000-0001-7750-7254> a schema:Person;
    schema:identifier [ a schema:PropertyValue;
        schema:propertyID  "orcid";
        schema:value   "0000-0001-7750-7254";
    ];
    schema:identifier [ a schema:PropertyValue;
        schema:propertyID  "ScopusAuthorID";
        schema:value   "55861456000";
```

Parsing took 17 ms. Validating the data took 20 ms.

```
[
    a sh:ValidationResult ;
    sh:resultSeverity sh:Warning ;
    sh:sourceConstraintComponent sh:MinCountConstraintComponent ;
    sh:sourceShape _:n208 ;
    sh:focusNode <PIC:999518944> ;
    sh:resultPath schema:leiCode ;
    sh:resultMessage "LeiCode is recommended. Please fill in a value"@en ;
]
```

Figure 19 – SHACL online Validator

4. Commit and pull request

- Log in GitHub https://github.com/epos-eu/EPOS-DCAT-AP;
- Create a fork on your personal account by the "fork" button;
- Browse into your forked project (at the top of the page you should read youraccount/EPOS-DCAT-AP forked from epos-eu/EPOS-DCAT-AP) and go to the directory https://github.com/epos-eu/EPOS-DCAT-AP/tree/EPOS-DCAT-AP-shapes/examples/WPXX (replace XX with WP number);
- Click on "Upload files" button;
- Drag your file into the graphic dashed square;
- Do the commit, inserting a properly label (es. add WPXX RDF/turtle file);
- Move to the (second) tab ("pull requests") and click on "New pull request" button and then "Create Pull Request". Insert a title for the request and finally confirm by clicking on "Create Pull Request".
- The file will be available as soon as the Administrator will approve the request.

5. Contacts

- Rossana Paciello, rossana.paciello@ingv.it, for Metadata model issues;
- Riccardo Rabissoni, riccardo.rabissoni@ingv.it, for Metadata Web Editor issues;
- Jan Michalek, jan.michalek@uib.no, for DDSS-ID and other issue.