



OPEN DATA SUPPORT

Training Module 1.5

Promoting the reuse
of Open Government
Data through the
Open Data
Interoperability
Platform (ODIP)



PwC firms help organisations and individuals create the value they're looking for. We're a network of firms in 158 countries with close to 180,000 people who are committed to delivering quality in assurance, tax and advisory services. Tell us what matters to you and find out more by visiting us at www.pwc.com. PwC refers to the PwC network and/or one or more of its member firms, each of which is a separate legal entity. Please see www.pwc.com/structure for further details.

This presentation has been created by PwC

Authors:

Michiel De Keyzer, Nikolaos Loutas and Stijn Goedertier

Presentation metadata

Open Data Support is funded by the European Commission under SMART 2012/0107 'Lot 2: Provision of services for the Publication, Access and Reuse of Open Public Data across the European Union, through existing open data portals' (Contract No. 30-CE-0530965/00-17).

© 2014 European Commission

Disclaimers

1. The views expressed in this presentation are purely those of the authors and may not, in any circumstances, be interpreted as stating an official position of the European Commission. The European Commission does not guarantee the accuracy of the information included in this presentation, nor does it accept any responsibility for any use thereof. Reference herein to any specific products, specifications, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favouring by the European Commission. All care has been taken by the author to ensure that s/he has obtained, where necessary, permission to use any parts of manuscripts including illustrations, maps, and graphs, on which intellectual property rights already exist from the titular holder(s) of such rights or from her/his or their legal representative.
2. This presentation has been carefully compiled by PwC, but no representation is made or warranty given (either express or implied) as to the completeness or accuracy of the information it contains. PwC is not liable for the information in this presentation or any decision or consequence based on the use of it.. PwC will not be liable for any damages arising from the use of the information contained in this presentation. The information contained in this presentation is of a general nature and is solely for guidance on matters of general interest. This presentation is not a substitute for professional advice on any particular matter. No reader should act on the basis of any matter contained in this publication without considering appropriate professional advice.

Learning objectives

By the end of this training module you should have an understanding of:

- How you can overcome the barriers of reuse for your datasets.
- How Open Data Support can promote the reuse of datasets.
- What the DCAT Application Profile is and how it can be used.
- What Open Data Interoperability Platform (ODIP) is and how it can be used.

Content

This module contains...

- An outline of the context of Open Government Data in Europe.
- An outline of the Open Data Support project.
- Information about the DCAT Application Profile for Data Portals in Europe as a homogenised metadata model.
- Information on how to use the Open Data Interoperability Platform.

There are more than 160 portals in Europe hosting Open Government Data

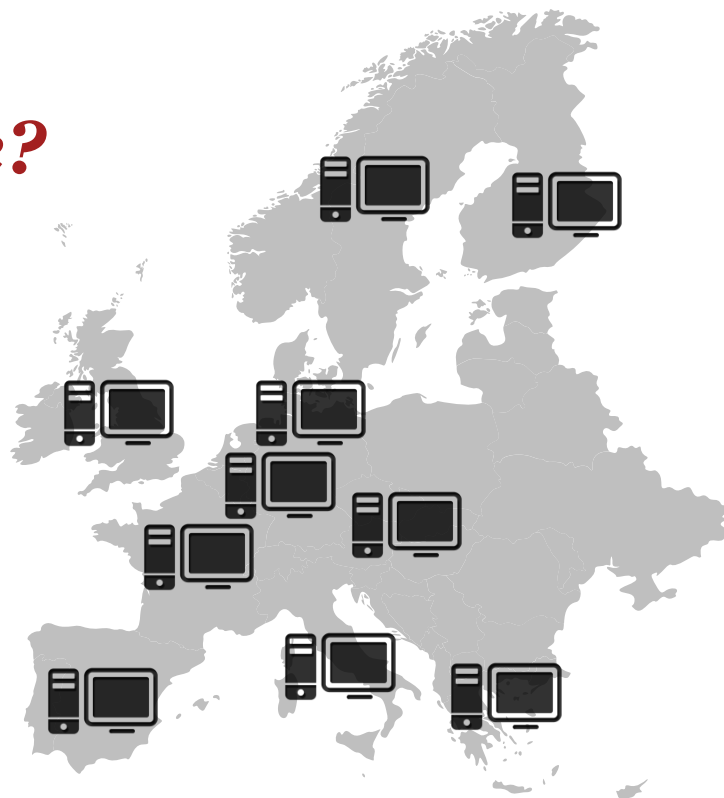
Provenance?

Licence? Persistence?

Trust? Availability?

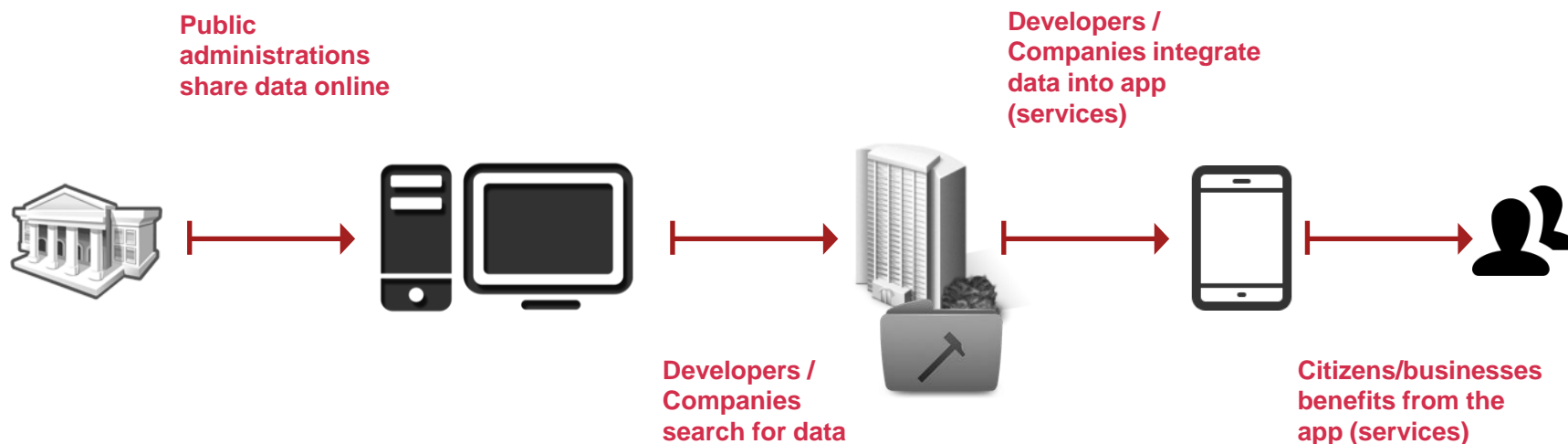
Quality?

160+



Open Data has a great potential to create social and economic value

Publishing data



Reusing data

Barriers to Open Data publishing and reuse

Data publishers

No view on which data is more likely to be reused / has a higher ROI potential.

Unclear business model for publishing Open Data.

Limited tool support.

Competing licenses for datasets.

Competing vocabularies for describing datasets.

Domain-specific metadata needs.

Effort required for keeping the metadata up-to-date.

Data reusers

Lack of overview of existing/available datasets.

Unclear business model for reusing Open Data.

Data is often of low quality, outdated, unstructured and/or not machine-readable.

Lack of licensing information or incompatible licenses.

Different vocabularies when searching for datasets.

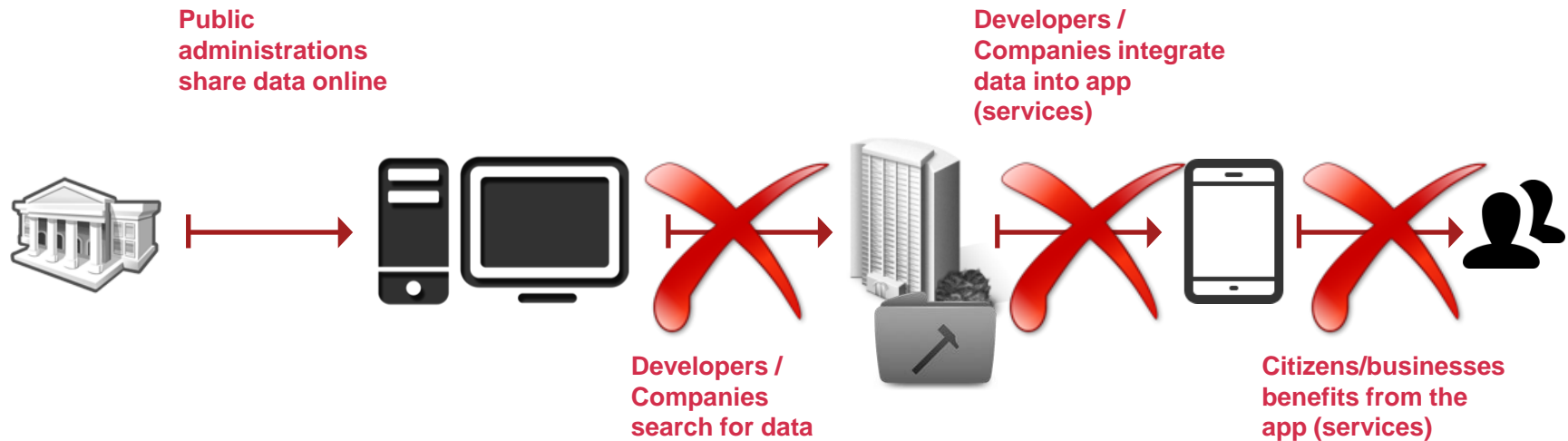
Lack of (good quality) metadata.

Lack of provenance information.

Metadata

Metadata

No reuse = No social and economic value



Open Data Support

...funded by the European Commission, DG CONNECT, aims at lowering accessibility and awareness barriers.

Open Data Support mission...

To Improve the **visibility** and facilitate the **access** to datasets published on local and national Open Data portals in order to increase their **reuse** within and across borders.

See also:

<http://www.slideshare.net/OpenDataSupport>

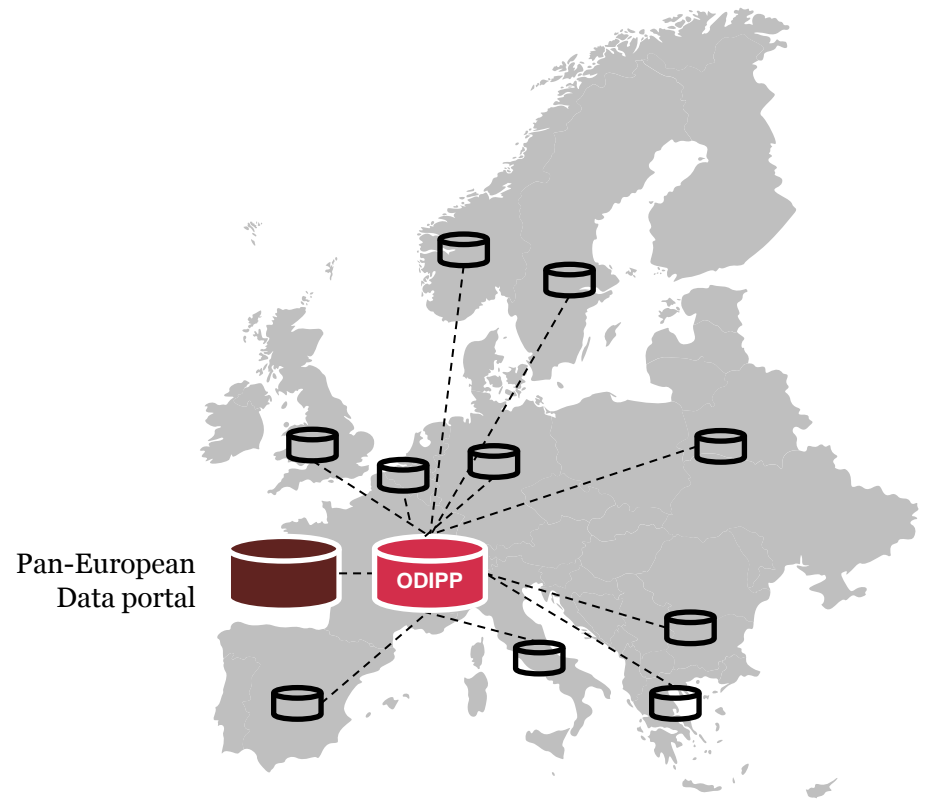
By ...

Providing

homogenised access

to metadata
descriptions of open
datasets via a

single point of access



DCAT Application Profile

A common vocabulary for describing datasets hosted in data portals in Europe, based on the Data Catalogue vocabulary (DCAT).

A shared initiative of...



Funded by the ISA Programme under Action 1.1.
“Improving semantic interoperability in European eGovernment systems” (a.k.a the [SEMIC](#) project).

An international Working Group of experts

- Chair: Antonio Carneiro (Publications Office)
- 59 Working Group members representing:
 - 15 different European Member States
(UK,IT,ES,DK,DE,SK,BE,AT,SE,FI,FR,IE,NL,GR,SI)
 - US
 - Several European Institutions and international organisations
 - 40 different Data Portals

See also:

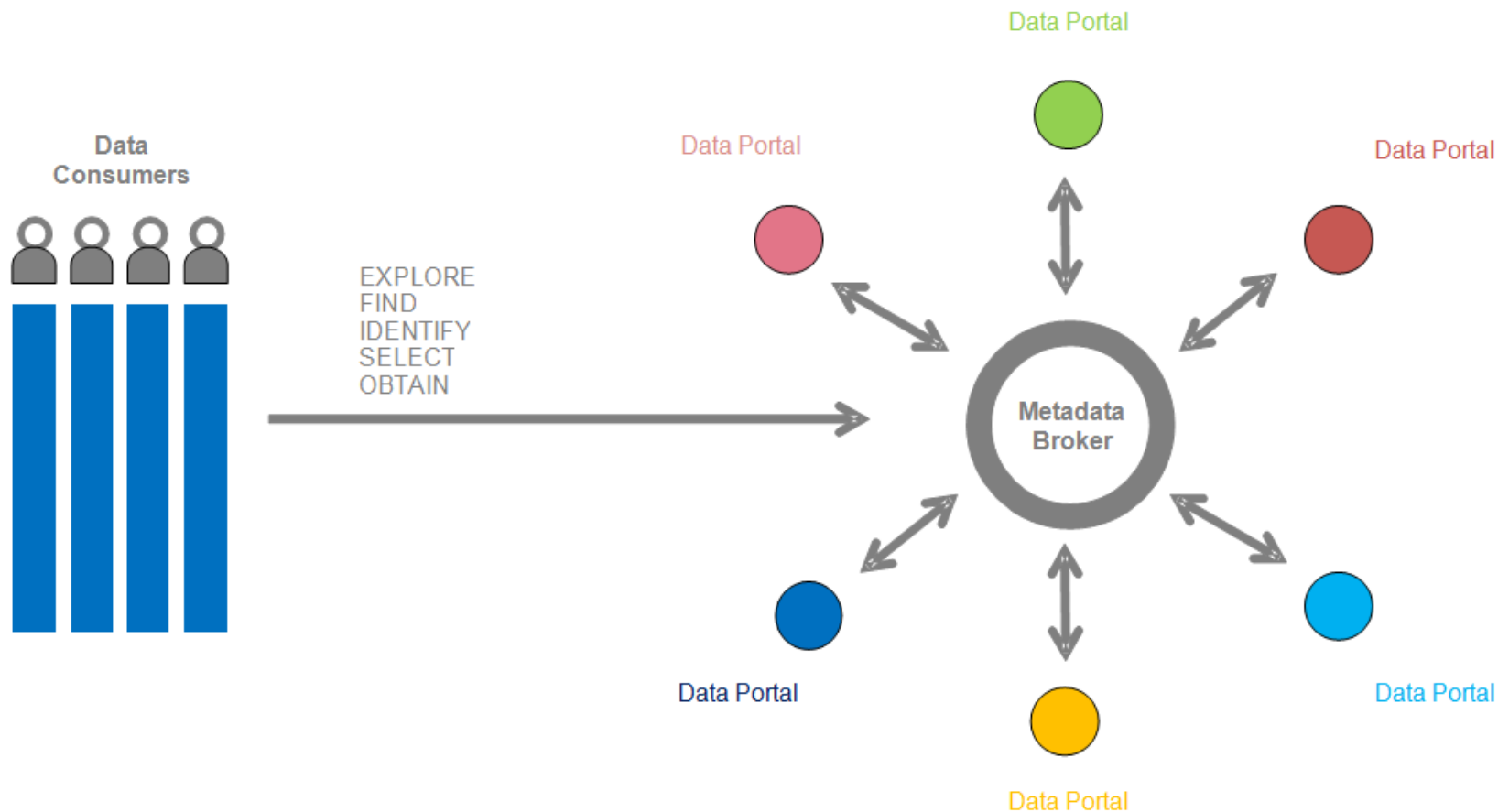
https://joinup.ec.europa.eu/asset/dcat_application_profile/description

By using a common metadata schema to describe datasets and sharing metadata...

- **Data publishers** increase discoverability and thus reuse of their data.
- **Data reusers** can uniformly search across platforms without facing difficulties caused by the use of separate models or language differences.

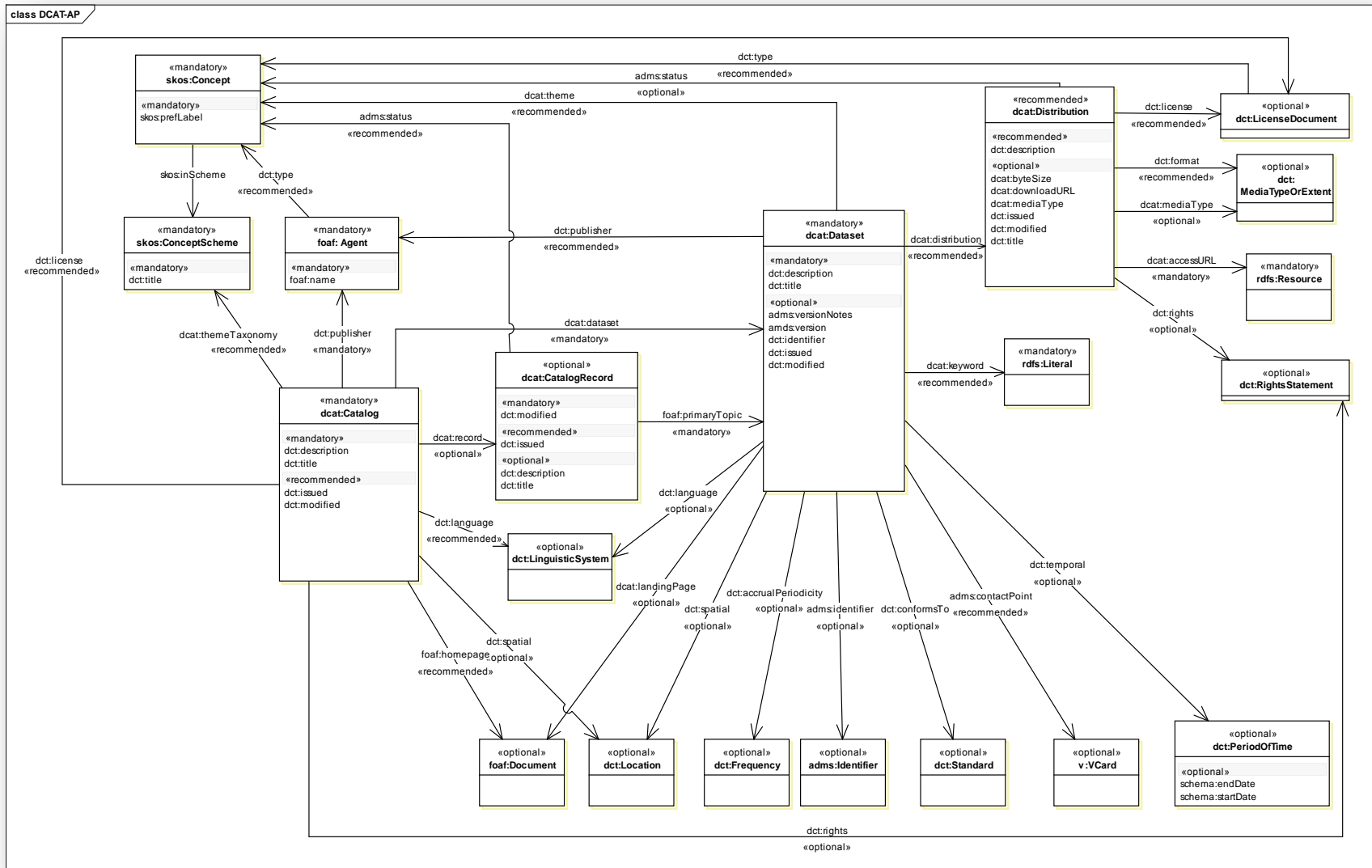
The quality and the availability of the description metadata directly affects how easily datasets can be found!

The DCAT-AP enables the exchange of description metadata between data portals



What's in the
specification?

The DCAT Application Profile data model



Usage of the DCAT Application Profile

Mandatory class: a receiver of data **MUST** be able to process information about instances of the class; a sender of data **MUST** provide information about instances of the class.

Recommended class: a receiver of data **MUST** be able to process information about instances of the class; a sender of data **MUST** provide information about instances of the class, if it is available.

Optional class: a receiver **MUST** be able to process information about instances of the class; a sender **MAY** provide the information but is not obliged to do so.

Mandatory property: a receiver **MUST** be able to process the information for that property; a sender **MUST** provide the information for that property.

Recommended property: a receiver **MUST** be able to process the information for that property; a sender **SHOULD** provide the information for that property if it is available.

Optional property: a receiver **MUST** be able to process the information for that property; a sender **MAY** provide the information for that property but is not obliged to do so.

Controlled vocabularies

Property URI	Used for Class	Proposed vocabulary
dcat:mediaType	Distribution	MDR File types Name Authority List
dcat:theme	Dataset	EuroVoc domains
dcat:themeTaxonomy	Catalog	EuroVoc
dct:accrualPeriodicity	Dataset	Dublin Core Collection Description Frequency Vocabulary
dct:format	Distribution	MDR File Type Named Authority List
dct:language	Catalog, Dataset	MDR Languages Named Authority List
dct:publisher	Catalog, Dataset	MDR Corporate bodies Named Authority List
dct:spatial	Catalog, Dataset	MDR Countries Named Authority List, MDR Places Named Authority List
adms:status	CatalogRecord	ADMS change type vocabulary
dct:type	License Document	ADMS license type vocabulary

Mapping example – data.gov.uk

Scottish Road Accident Statistics *dct:title (Dataset)*

Data about injury road accidents, accident costs, vehicles involved, drivers and riders, drink-drive accidents, drivers breath tested, casualties and international comparisons.



Source agency: Scottish Government

Designation: National Statistics

Language: English *dct:language*

Alternative title: Scottish Road Accident Statistics

dct:description

Licence *dct:license*

UK Open Government Licence (OGL)

Data Resources **2**

Key statistics for 2007

2007 Volume

dct:title (Distribution)

Details Download ▾
Details Download ▾

dcat:accessURL

*dcat:downloadURL, dct:issued,
dct:format, dct:description*

Additional Information

Openness score	★★★★★
Geographic coverage	Scotland <i>dct:spatial</i>
National statistic	yes
ONS Category	Travel and Transport <i>dct:theme</i>
Temporal coverage	No value <i>dct:temporal</i>
Date added computed	No value
Date updated computed	No value

dct:publisher

Publisher

Scottish Government

Enquiries:

No details supplied

FOI Contact:

- Web:

<http://www.whatdotheyknow.com...>

adms:contactPoint

Tags

- accident
- health-well-being-and-care
- road
- road-accidents
- road-safety
- roads
- safety
- transport
- transport-accidents-and-casualties
- travel-and-transport

dcat:keyword

About this dataset

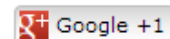
- Added to data.gov.uk: 10/12/2011
- Modified on data.gov.uk: 10/06/2013
- History of changes
- JSON, API and URI for developers

dct:issued

dct:modified

Do more with this data

- Share your app
- Share an idea
- Request new data



Example description of dataset with the DCAT-AP

```
<rdf:Description rdf:about="http://data.gov.uk/data ">
  <rdf:type rdf:resource="http://www.w3.org/ns/dcat#Catalog"/>
  < dct:title xml:lang="en">data.gov.uk</dct:title>
  < dct:description xml:lang="en">Description of the data portal</dct:description>
  < dct:license rdf:resource=" http://www.nationalarchives.gov.uk/doc/open-government-licence"/>
</rdf:Description>

<rdf:Description rdf:about="http://data.gov.uk/dataset/east-sussex-county-council-election-results"/>
  <rdf:type rdf:resource="http://www.w3.org/ns/dcat#Dataset"/>
  < dct:title xml:lang="en">East Sussex County Council election results</dct:title>
  < dct:description xml:lang="en">A list of elections to East Sussex County Council, which leads to data about candidates,
  parties, electoral divisions and votes cast. Uses the Open Election Data RDF vocabulary from http://openelectiondata.org/
  </dct:description>
</rdf:Description>

<rdf:Description rdf:about="http://www.eastsussex.gov.uk/yourcouncil/localelections/election2009/default.aspx"/>
  <rdf:type rdf:resource="http://www.w3.org/ns/dcat#Distribution"/>
  < dct:title xml:lang="en">East Sussex County Council election 4 June 2009, and subsequent bi-elections</dct:title>
  < dcat:accessURL rdf:resource="http://www.eastsussex.gov.uk/yourcouncil/localelections/election2009/default.aspx ">
  < dct:license rdf:resource="http://www.nationalarchives.gov.uk/doc/open-government-licence"/>
</rdf:Description>
```

Creating mappings to the DCAT-AP

Dataset Properties	Example Value	Harmonized Predicate	Generated SPARQL
Raw Predicate			
http://data.gov.uk/predicate/title http://data.gov.uk/predicate/unpublished	Government Major Projects data for the Foreign and Commonwealth Office 2012 FALSE	dct:title	<pre> prefix dct <http://purl.org/dc/terms/> INSERT { ?harmds dct:title ?d. } where { ?ds a <http://www.w3.org/ns/dcat#Dataset>. ?ds <http://data.gov.uk/predicate/title> ?d. ?harmrecord <http://xmlns.com/foaf/0.1/primaryTopic> ?harmds. ?harmrecord <http://data.opendatasupport.eu/ontology/harmonisation.owl#raw_?ds. } </pre>
http://data.gov.uk/predicate/update_frequency	other	dct:accrualPeriodicity	<pre> prefix dct <http://purl.org/dc/terms/> INSERT { ?harmds dct:accrualPeriodicity ?d. } where { ?ds a <http://www.w3.org/ns/dcat#Dataset>. ?ds <http://data.gov.uk/predicate/update_frequency> ?d. ?harmrecord <http://xmlns.com/foaf/0.1/primaryTopic> ?harmds. ?harmrecord <http://data.opendatasupport.eu/ontology/harmonisation.owl#raw_?ds. } </pre>
http://data.gov.uk/predicate/update_frequency-other	quarterly	dct:accrualPeriodicity	<pre> prefix dct <http://purl.org/dc/terms/> INSERT { ?harmds dct:accrualPeriodicity ?d. } where { ?ds a <http://www.w3.org/ns/dcat#Dataset>. ?ds <http://data.gov.uk/predicate/update_frequency-other> ?d. ?harmrecord <http://xmlns.com/foaf/0.1/primaryTopic> ?harmds. ?harmrecord <http://data.opendatasupport.eu/ontology/harmonisation.owl#raw_?ds. } </pre>

Where can you find it?

The screenshot shows the Joinup website interface. At the top, there is a navigation bar with links for Contact, Search, Glossary, Help, Partners, Analytics, and Disclaimer, along with a language selector set to English (en). The Joinup logo and the European Commission logo are visible. Below the logo, there is a search bar and a 'Login or Sign up' button. The main header area contains the text 'Share and reuse interoperability solutions for public administrations' and a breadcrumb trail: 'European Commission > ISA > Joinup > Semantic assets > Projects > Dcat application profile > Description'. A secondary navigation bar includes 'My Page', 'Communities', 'Semantic Assets', 'Software', 'News', 'Events', 'e-Library', and 'People'. The main content area features a sidebar on the left with a table of contents for the Semantic Asset page, including sections like Welcome, Description, Members list, Issues, Asset Releases, Metrics, Highlights, Semantic Assets, Software, Communities, Communications, News, and Events. The main content area displays the 'DCAT application profile for data portals in Europe' with a 'Download' button, a 5-star rating, and an 'Editor's choice' badge. Below this, it states the submission date (March 08, 2013) and rating (5/5), and notes that 11 people use the project. The 'Description' section explains that DCAT-AP is a specification based on the Data Catalogue vocabulary (DCAT) for describing public sector datasets in Europe. It also mentions the 'Open Data Support' service. A 'Participate in the public review' section notes that the final draft was released for public review on May 13, 2013. On the right side, there are several call-to-action boxes: 'Request to be a member of this semantic asset project', 'Export description metadata', 'I use this project', and a membership approval notice. A 'Related Content' section lists several workshops and events, including the 2nd CESAR Workshop, CESAR Workshop 2012.03.07, and the EFIR Workshop 2013.

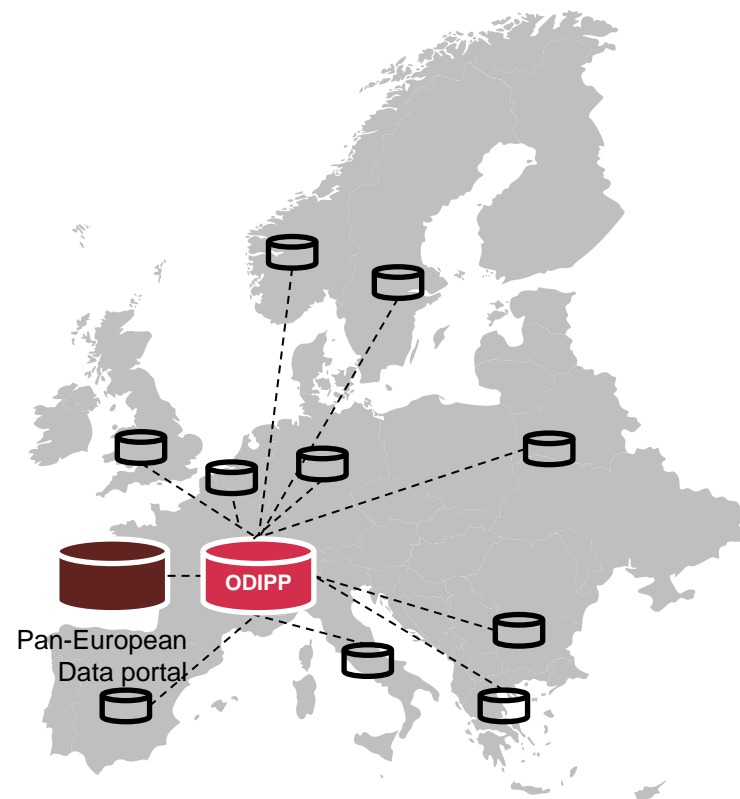
https://joinup.ec.europa.eu/asset/dcat_application_profile/description

Share the metadata of your datasets on ODIP

The Open Data Interoperability Platform (ODIP) enables you to share metadata of datasets described using the DCAT-AP, thus improving the discoverability and visibility of your datasets, eventually leading to wider reuse.

What can ODIP do?

- **Harvest** metadata from an Open Data portal.
- **Transform** the metadata to RDF.
- **Harmonise** the RDF metadata produced in the previous steps with DCAT-AP.
- **Validate** the harmonised metadata against the DCAT-AP.
- **Publish** the description metadata as Linked Open Metadata.
- **Translate** metadata automatically in English



How can ODIP help you improve your metadata?

- ODIP maps your metadata to a standard model, i.e. the DCAT-AP.
- ODIP helps you reuse standardised multilingual controlled vocabularies in your metadata, replacing error-prone text values or tailor-made lists.
- By means of its validation services, ODIP allows you to detect inconsistencies and errors in your metadata.
- ODIP assigns persistent URIs to your metadata.
- ODIP links your metadata with other metadata, thus adding context to it and enriching its meaning.
- ODIP automatically translates the title and description of the metadata to English.

How does ODIP look like?

The screenshot displays the Open Data Interoperability Platform (ODIP) interface. The main area shows a table of jobs with columns for NAME, INTERVAL, STATUS, SCHEDULE, RUN, and DELETE. The 'odp raw harvesting' job is selected and highlighted in blue. To the right, a detailed view of this job is shown, including its creation date, description, interval, name, and schedule type. The 'odp raw harvesting' job is scheduled to run every 4 days.

NAME	INTERVAL	STATUS	SCHEDULE	RUN	DELETE
Ireland: harmonization				Run	X
Ireland: raw	0 0 3 ***	30	Cancel	Run	X
odp: harmonization				Run	X
data.gov.uk: raw harvest	0 0 0 ***	30	Cancel	Run	X
odp raw harvesting	0 0 4 ***	30	Cancel	Run	X
data.gov.uk: harmonization				Run	X

Job Details: odp raw harvesting

- Chained
- Created: Fri Jun 14 16:05:43 CEST 2013
- Description: sdfsdfsdf
- Interval: 0 0 4 ***
- Name: odp raw harvesting
- Previous Job Id:
- Schedule Type: interval
- Scheduled
- Next run in 0 days, 16 hours, 35 minutes and 0 seconds
- Edit Job

SELECTED EXTRACTORS

CKAN Extractor

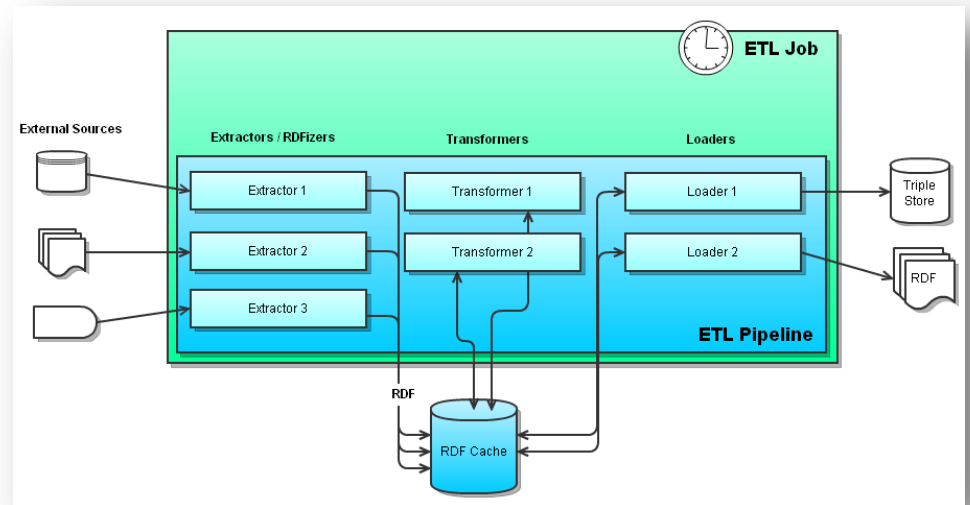
<http://odip.opendatasupport.eu>

An ODIP Job

The ODIP job consists of three possible phases which need to be ran in order and that are composed of several plug-ins :

1. Extraction
2. Transformation
3. Loading

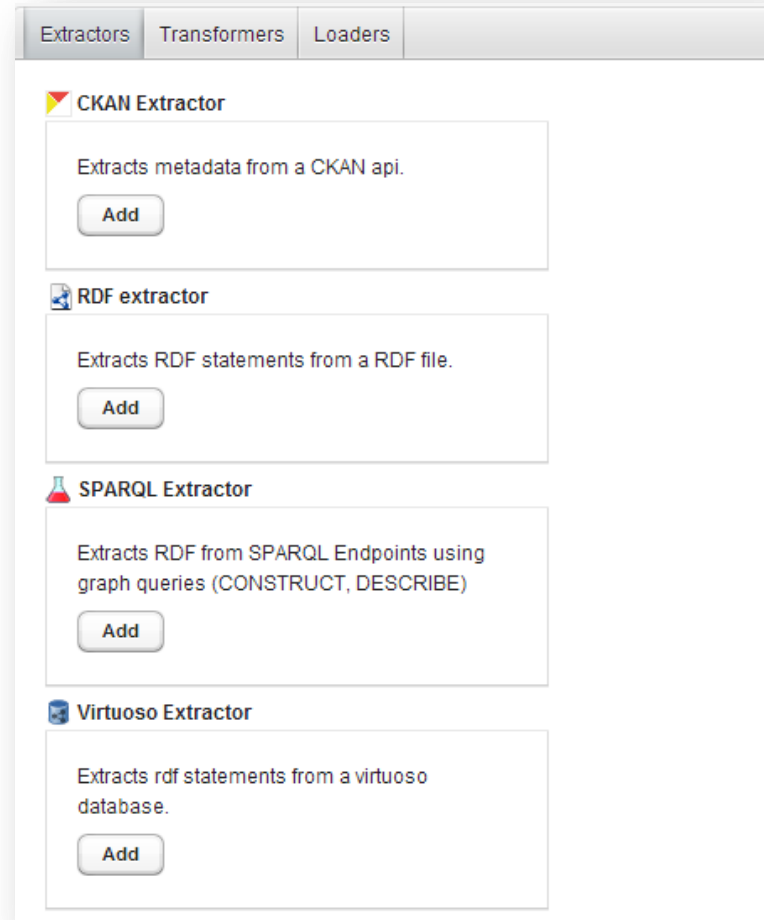
Furthermore these jobs can be scheduled to be launched periodically, in succession or manually.



Overview of ODIP's Extract-Transform- Load process

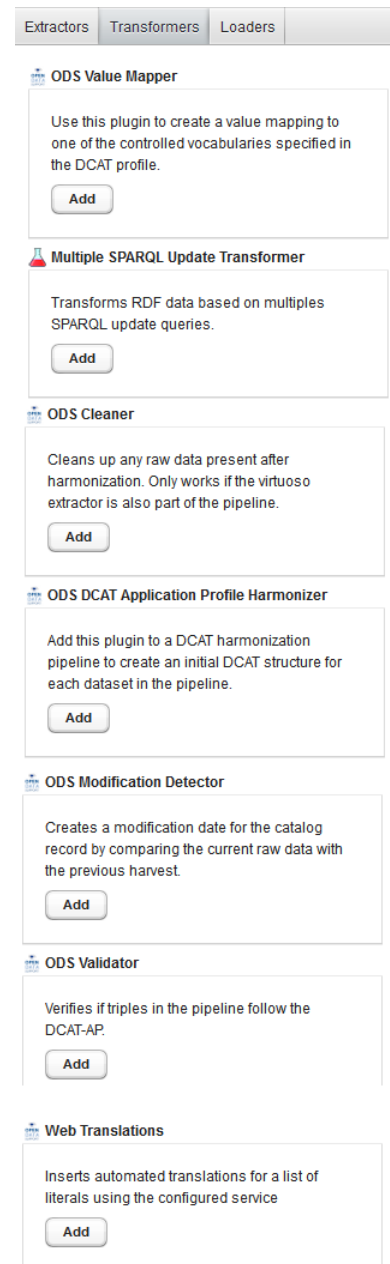
1. Extraction

- The extraction phase entails retrieving (extracting) raw data from a given source Open Data portal using the appropriate plug-in, depending on the technology of the source.
- Available extractors:
 - CKAN Extractor
 - RDF extractor
 - SPARQL Extractor
 - Virtuoso Extractor
 - CSV Extractor



2. Transformation (1/3)

- The goal of the transformation phase is to harmonise, cleanse and prepare for storing on ODIP metadata harvested from Open Data portals.
- Available transformers:
 - ODS Value Mapper.
 - SPARQL Update Query Transformer.
 - ODS Cleaner.
 - ODS DCAT Application Profile Harmoniser.
 - ODS Modification Detector.
 - ODS Validator.
 - Web Translations.

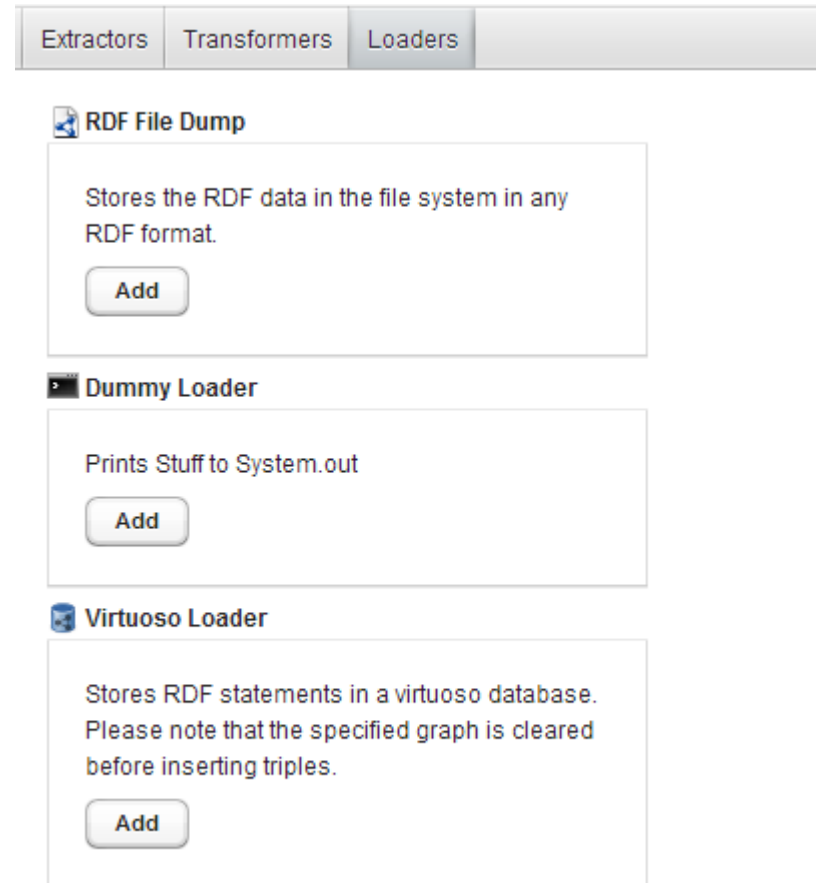


The screenshot displays a web interface for configuring data transformers. At the top, there are three tabs: 'Extractors', 'Transformers', and 'Loaders'. The 'Transformers' tab is active. Below the tabs, a list of transformer plugins is shown, each with a description and an 'Add' button:

- ODS Value Mapper**: Use this plugin to create a value mapping to one of the controlled vocabularies specified in the DCAT profile.
- Multiple SPARQL Update Transformer**: Transforms RDF data based on multiples SPARQL update queries.
- ODS Cleaner**: Cleans up any raw data present after harmonization. Only works if the virtuoso extractor is also part of the pipeline.
- ODS DCAT Application Profile Harmonizer**: Add this plugin to a DCAT harmonization pipeline to create an initial DCAT structure for each dataset in the pipeline.
- ODS Modification Detector**: Creates a modification date for the catalog record by comparing the current raw data with the previous harvest.
- ODS Validator**: Verifies if triples in the pipeline follow the DCAT-AP.
- Web Translations**: Inserts automated translations for a list of literals using the configured service.

Loading

- In the loading phase, the harvested and harmonised metadata is stored on Virtuoso's RDF repository using the Virtuoso Loader.



The screenshot displays a web interface for configuring data loaders. At the top, there are three tabs: 'Extractors', 'Transformers', and 'Loaders', with 'Loaders' being the active tab. Below the tabs, three loader options are listed, each with a description and an 'Add' button:

- RDF File Dump**: Stores the RDF data in the file system in any RDF format.
- Dummy Loader**: Prints Stuff to System.out.
- Virtuoso Loader**: Stores RDF statements in a virtuoso database. Please note that the specified graph is cleared before inserting triples.

Example

Harvesting a CKAN-based Open Data portal

1. Create a new job on ODIP
2. Extraction phase
 - Add and Configure a CKAN Extractor to harvest data from a CKAN API.
3. Transformation phase
 - Add ODS Value mapper
 - Add a SPARQL Update Query Transformer with the pertinent queries
 - Add ODS Cleaner
 - Add and configure DCAT Application Profile Harmoniser
 - Add Modification detector
 - Add ODS Validator
 - Add Web Translations
4. Loading phase
 - Load the extracted data in a Virtuoso RDF Store via the Virtuoso Loader
5. Scheduling the job on ODIP

Example – 1. Create Job : Creating a job on ODIP

- To create a new job, click on “New Job”.
- At the bottom part of the screen you can configure the actual tasks within each of the three phases by selecting a tab.
- For each phase you can add and configure modules accordingly.

Open Data Interoperability Platform

New Job Manage Jobs Error Reports Get Support Logout

Name *

Description

Schedule type interval

Scheduled

Execution Interval

Save Cancel

Extractors Transformers Loaders

CKAN Extractor
Extracts metadata from a CKAN api.
Add

RDF extractor
extracts rdf statements from a rdf file on a uri
Add

SPARQL Extractor
Extracts RDF from SPARQL Endpoints using graph queries (CONSTRUCT, DESCRIBE)
Add

Provide a name for the Job.

Present the job with a short description.

Press the “Add” button to determine the plug-ins to deploy.

Example – 2.Extraction : Adding and Configuring a CKAN Extractor to harvest data from a CKAN API

After adding the CKAN extractor plugin you will be prompted to fill out the following form:

Publisher, license, title and description: Used in the stored catalog for the dct:publisher, dct:license, dct:title and dct:description properties.

Subject prefix: The prefix used to create a URI for each the metadata of harvested dataset.
The subject is created as `<subjectprefix>/dataset/<datasetid>`

Ignored keys: A comma seperated list of JSON attributes that should not be converted to RDF triples.

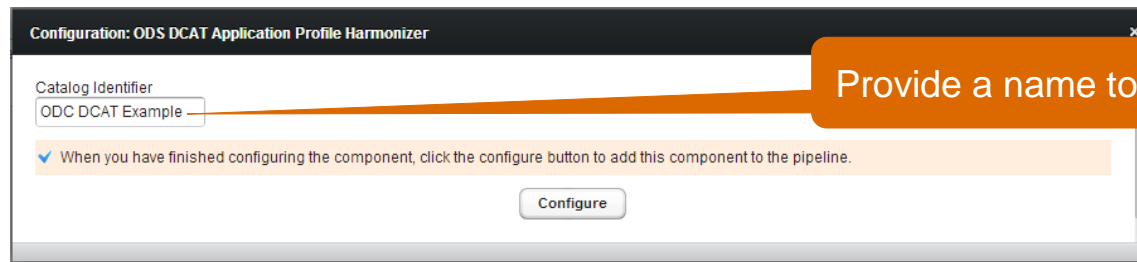
The Web location of the CKAN portal you wish to harvest.
The portal should support API version 3 and the API must be enabled.

Predicate prefix: JSON attributes are converted to predicates by appending them to the predicate prefix.
The CKAN API response is in JSON, we then convert this into RDF.

The screenshot shows a web form titled "Configuration: CKAN Extractor". The form contains several input fields and a checkbox. The fields are: "CKAN Uri" (http://odp.tenforce.com/data/), "Publisher" (bert@tenforce.com), "Title" (ODP EU), "Description" (The european open data portal), "License" (http://ec.europa.eu/geninfo/), "Predicate Prefix" (http://odp.tenforce.com/data/predicate/), "Subject Prefix" (http://odp.tenforce.com/data/dataset/), and "Ignored Keys" (rdf). There is a checkbox labeled "harvest all datasets" which is checked. At the bottom right, there is a "Configure" button. A yellow callout bubble points to the "Ignored Keys" field.

Example – 3. Transformation : Adding and configuring plug-ins to harmonise data(1/3)

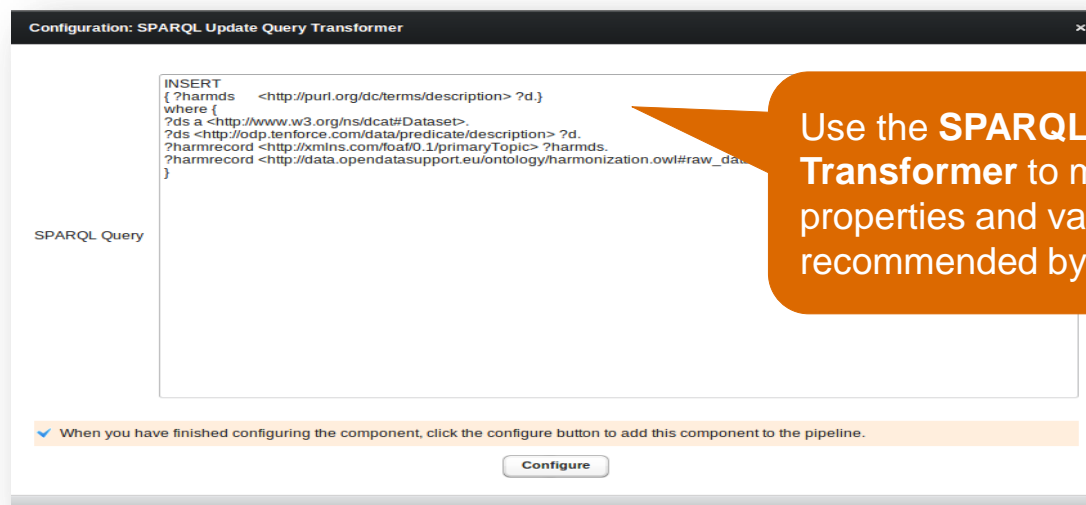
- Start by adding the **ODS DCAT Application Profile Harmonizer**.
 - ✓ This plugin will create the harmonized catalog data and a basic skeleton for each dataset it identifies.



- Use the **Modification Detector** to compare provenance data generated by the CKAN extractor between the current and previous version of the raw data to set the dct:modified field of the catalog records.
 - ✓ No configuration is required.

Example – 3. Transformation : Adding and configuring plug-ins to harmonise data (2/3)

- Mapping the description of dataset to dct:description as required by the DCAT-AP.



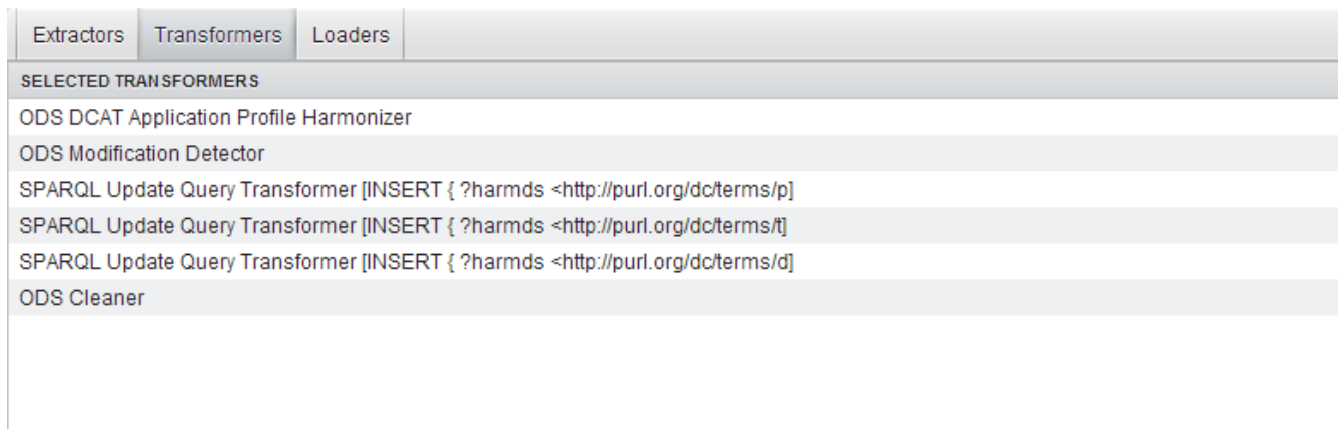
Use the **SPARQL Update Query Transformer** to map existing properties and values to the ones of recommended by the DCAT-AP.

- Use the **ODS Cleaner Plugin** to remove raw data loaded into the working set before storing it into a harmonized graph.
 - ✓ No configuration is required.

Example – 3. Transformation : Adding and configuring plug-ins to harmonise data (3/3)

Result

The final result of your harmonisation pipeline should look similar to the following :

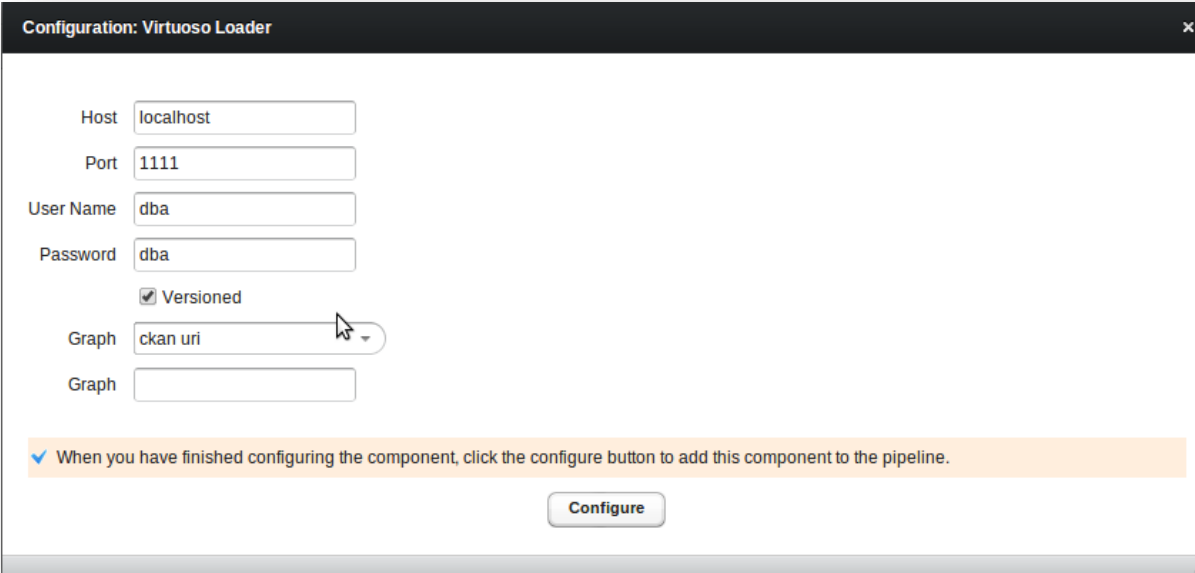


Configure the Virtuoso Loader to load the harmonized data into Virtuoso.

Example – 4. Loading: Load the extracted data in a Virtuoso RDF Store via the Virtuoso Loader

The Virtuoso Loader will store the generated triples in the Virtuoso RDF store. The triples will be inserted into a graph of your choice.

The Virtuoso Loader needs host, port and user credentials to connect to your Virtuoso server.



The screenshot shows a configuration window titled "Configuration: Virtuoso Loader". It contains the following fields and options:

- Host: localhost
- Port: 1111
- User Name: dba
- Password: dba
- Versioned
- Graph: ckan uri (with a dropdown arrow)
- Graph: (empty text box)

At the bottom, there is a blue checkmark icon followed by the text: "When you have finished configuring the component, click the configure button to add this component to the pipeline." Below this is a "Configure" button.

5. Scheduling a job on ODIP

A job can be scheduled to run at a set interval or chained after another job:

- **Interval Scheduling:**

<sec> <min> <hour> <day-of-month> <month> <day-of-week>

□ Example:

- 0 0 4 * * * - each day at 4 am
 - 0 0 0 * * 1 - each Monday at midnight
 - 0 30 * * * - every half past the hour
- **Chained scheduling:** Select a job after which this job should be executed.

ODIP Reporting tool

Whenever a “job” is ran, a report is created and can be reviewed as can be seen in the following screenshot:

The screenshot displays the ODIP reporting tool interface. On the left, there is a table of jobs with columns for ID and NAME. The selected job is 'Adms minimum example : Test' with ID 'd58dd6fa-01d3-453d-act3-8a6e434ebdb'. The top right section shows 'Last Executions' with columns for DATE, DURATION, and COMPONENT ERRORS. The execution date is '19.08.13 - 16:55:16', duration is '1 second', and component errors are '0'. The bottom section shows a detailed execution report with columns for EXTRACTOR, RESULT, WARNINGS, TRANSFORMER, RESULT, WARNINGS, LOADER, RESULT, and WARNINGS. The report shows that the RDF extractor and Virtuoso Loader both completed successfully with 'OK' results and '0' warnings. The transformers (SPARQL Update Query Transformer, ODS DCAT Application Profile Harm, and ODS Cleaner) also completed successfully with 'OK' results and '0' warnings.

ID	NAME
c81f979a-44a8-4385-a89a-42b93f56fd01	ireland: harmonization
88a42eaa-94e0-4a3c-95d9-f4459544d0f	ireland: raw
3b870748-571b-49c2-87d1-7ea91eba36	odp: harmonization
dc6fcc83-3cfe-4338-acbd-2162fb07d43b	ADMS_Test_2nd_Wave
48e7d981-7d0c-4303-896d-e1a73d46fbc	data.gov.uk: raw harvest
e5601b42-0023-4995-a2b0-e1c5622f5dc	odp raw harvesting
2f9a64a1-75e8-4f05-ab95-fec1a01d15dc	data.gov.uk: harmonization
5c614323-db87-4b7e-a224-17e4922a76	CKAN_test_2nd_wave
d58dd6fa-01d3-453d-act3-8a6e434ebdb	Adms minimum example : Test

DATE	DURATION	COMPONENT ERRORS
19.08.13 - 16:55:16	1 second	0

EXTRACTOR	RESULT	WARNINGS	TRANSFORMER	RESULT	WARNINGS	LOADER	RESULT	WARNINGS
RDF extractor	OK	0	SPARQL Update Query Transformer	OK	0	Virtuoso Loader	OK	0
			SPARQL Update Query Transformer	OK	0			
			SPARQL Update Query Transformer	OK	0			
			ODS DCAT Application Profile Harm	OK	0			
			ODS Cleaner	OK	0			

Select the appropriate job

Informs user whether or not a plug-in functioned correctly or not.

Discover datasets through ODIP

The Open Data Interoperability Platform (ODIP) enables you to share metadata of datasets described using the DCAT-AP, thus improving the discoverability and visibility of your datasets, eventually leading to wider reuse.

The public SPARQL endpoint of ODIP

Query interface



The screenshot shows the Open Data Support website's SPARQL Query interface. The header includes the logo and name 'OPEN DATASUPPORT', a 'Home' button, and a 'Sample Queries' dropdown menu. On the left, there is a navigation menu with links for 'Homepage', 'Training', 'Interoperability Platform', and 'Contact'. Below this, a section titled 'MORE ABOUT LINKED DATA' contains links for 'Understanding Linked Data by example', 'Case study on how Linked Data is transforming eGovernment', 'Describe organizations in RDF with Core Business Vocabulary and ORG Ontology', and '10 Rules for Persistent URIs'. The main content area is titled 'SPARQL Query' and features a text input field containing the following query:

```
prefix dcat: <http://www.w3.org/ns/dcat#>
select *
where {{?record a dcat:CatalogRecord }}{{?record ?x ?y}}
LIMIT 100
```

Below the input field is a 'run query' button.

<http://data.opendatasupport.eu>

The public SPARQL endpoint of ODIP

Result set

The screenshot shows the ODIP SPARQL query interface. At the top, there is a navigation bar with 'OPEN DATASUPPORT', 'Home', and 'Sample Queries'. A sidebar on the left contains links for 'OPEN DATA SUPPORT' (Homepage, Training, Interoperability Platform, Contact) and 'MORE ABOUT LINKED DATA' (Understanding Linked Data by example, Case study on how Linked Data is transforming eGovernment, Describe organizations in RDF with Core Business Vocabulary and ORG Ontology, 10 Rules for Persistent URIs).

The main area is titled 'SPARQL Query' and contains a text box with the following query:

```
prefix dcat: <http://www.w3.org/ns/dcat#>
select *
where {{?record a dcat:CatalogRecord }}{?record ?x ?y}}
LIMIT 100
```

Below the query box is a 'run query' button. The result set is displayed as a table with three columns: 'record', 'x', and 'y'. The 'record' column contains various URIs, the 'x' column contains URIs for related concepts, and the 'y' column contains URIs for related concepts and dates.

record	x	y
http://data.opendatasupport.eu/id/catalog/test/	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://www.w3.org/ns/dcat#CatalogRecord
http://data.opendatasupport.eu/id/catalog/test/	http://xmlns.com/foaf/0.1/primaryTopic	http://data.opendatasupport.eu/id/catalog/test/
http://data.opendatasupport.eu/id/catalog/test/	http://opendatasupport.eu/ontology/harmonisa	http://joinup.ec.europa.eu/asset/adms/release
http://data.opendatasupport.eu/id/catalog/irela-quarterly-	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://www.w3.org/ns/dcat#CatalogRecord
http://data.opendatasupport.eu/id/catalog/irela-quarterly-	http://purl.org/dc/terms/modified	2013-08-18T03:00:00.850+02:00
http://data.opendatasupport.eu/id/catalog/irela-quarterly-	http://xmlns.com/foaf/0.1/primaryTopic	http://data.opendatasupport.eu/id/catalog/irela-quarterly-
http://data.opendatasupport.eu/id/catalog/irela-quarterly-	http://data.opendatasupport.eu/ontology/harm	http://ie.ckan.net/dataset/deaths--quarterly-
http://data.opendatasupport.eu/id/catalog/irela-quarterly-	http://www.w3.org/ns/adms#status	:updated
http://data.opendatasupport.eu/id/catalog/irela-living-conditions-and-poverty	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://www.w3.org/ns/dcat#CatalogRecord
http://data.opendatasupport.eu/id/catalog/irela-living-conditions-and-poverty	http://xmlns.com/foaf/0.1/primaryTopic	http://data.opendatasupport.eu/id/catalog/irela-living-conditions-and-poverty
http://data.opendatasupport.eu/id/catalog/irela-living-conditions-and-poverty	http://data.opendatasupport.eu/ontology/harm	http://ie.ckan.net/dataset/income-living-conditions-and-poverty
http://data.opendatasupport.eu/id/catalog/irela-for-people-with-disabilities	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://www.w3.org/ns/dcat#CatalogRecord
http://data.opendatasupport.eu/id/catalog/irela-for-people-with-disabilities	http://xmlns.com/foaf/0.1/primaryTopic	http://data.opendatasupport.eu/id/catalog/irela-for-people-with-disabilities

More about ODIP



- ODIP is based on the [LOD Management Suite](#), originally created by the [Semantic Web Company](#) in the context of [LOD2](#) FP7 project.
- The LOD Manager Suite was further extended by TenForce in the context of Open Data Support for the deployment of ODIP.
- It will be made available on GitHub under [GPLv2](#).

Conclusions

- Good quality description metadata can improve the discoverability of open datasets.
- DCAT-AP can be used for homogenising metadata of datasets hosted on different Open Data portals and allows for querying them using a uniform vocabulary.
- ODIP can support harvesting, harmonising according to the DCAT-AP and publishing as linked data metadata of datasets published on different Open Data portals.
- ODIP, through its public SPARQL endpoint, provides a single point of access to datasets from all over Europe.
- Easier access to datasets means higher reuse of datasets.

Group questions



How many Open Government Data portals do you know in your country?

<http://www.visualpharm.com>



In your country, are you aware of any applications or services that were built upon Open Government Data?

<http://www.visualpharm.com>



How would you compare the visibility of Open Government Data portals with that of traditional data providers such as national statistics offices?

<http://www.visualpharm.com>



Have you heard about the Open Government Data initiatives of the European Commission?

<http://www.visualpharm.com>

Take also the online test [here!](#)

Thank you!

...and now YOUR questions?

References

Slide 4, 6, 9, 10, 11 & 12:

- Open Data Support: How can we help you?. Open Data Support.
<http://www.slideshare.net/OpenDataSupport/open-data-support-service-description>

Slide 12:

- Data Catalogue Vocabulary. <http://www.w3.org/TR/vocab-dcat/>

Slide 13-21:

- DCAT Application Profile for data portals in Europe Community. ISA Programme.
https://joinup.ec.europa.eu/asset/dcat_application_profile/description
https://joinup.ec.europa.eu/asset/dcat_application_profile/asset_release/all

Slide 23-35:

- LODMS User Manual for Open Data Support. Open Data Support

Slide 29:

- Figure from <http://www.semantic-web.at/linked-open-data-management-suite-lodms>

Related projects and initiatives



DCAT Application Profile for Data Portals in Europe, https://joinup.ec.europa.eu/asset/dcat_application_profile/description



Publicdata.eu, http://www.w3.org/2011/gld/wiki/Main_Page



LOD2 FP7 Project, <http://lod2.eu/>



The Semantic Web Company, <http://www.semantic-web.at/>



Linked Open Data Management Suite, <http://www.semantic-web.at/linked-open-data-management-suite-lodms>



OpenLink Virtuoso, <http://virtuoso.openlinksw.com/>

Data Catalog Interoperability Protocol, <http://spec.datacatalogs.org/>

Be part of our team...

Find us on



[Open Data Support](http://www.slideshare.net/OpenDataSupport)

<http://www.slideshare.net/OpenDataSupport>



[Open Data Support](http://goo.gl/y9ZZI)

<http://goo.gl/y9ZZI>

Follow us



[@OpenDataSupport](https://twitter.com/OpenDataSupport)

Join us on



joinup

<http://www.opendatasupport.eu>

Contact us

contact@opendatasupport.eu