



**OPEN
DATA
SUPPORT**

Trainingsmodul 1.5

Förderung der
Weiterverwendung von
offenen
Regierungsdaten
durch die Open Data
Interoperability
Plattform (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.

Präsentation Metadaten

Open Data Support wird von der Europäischen Kommission finanziert, gemäß 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'(Vertrag No. 30-CE-0530965/00-17).

© 2013 European Commission

Diese Präsentation wurde von PwC erstellt

Autoren:

Michiel De Keyzer, Nikolaos Loutas and Stijn Goedertier

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.

Lernziele

Am Ende dieses Trainingsmodul sollten Sie verstehen :

- wie Sie die Barrieren der Weiterverwendung für Ihre Datensätze überwinden können.
- wie Open Data Support die Weiterverwendung von Datensätzen fördern kann.
- was das DCAT Anwendungsprofil ist und wie es verwendet werden kann.
- was die Open Data Interoperability Platform (ODIP) ist und wie sie verwendet werden kann.

Inhalt

Dieses Modul enthält...

- einen Überblick über den Kontext von offenen Regierungsdaten in Europa.
- einen Überblick über das Open Data Support Projekt.
- Informationen über das DCAT Anwendungsprofil für Datenportale in Europa als ein homogenisiertes Metadaten Modell.
- Information darüber, wie die Open Data Interoperability Platform verwendet werden kann.

Es gibt mehr als 160 Portale in Europa, die offene Regierungsdaten hosten.

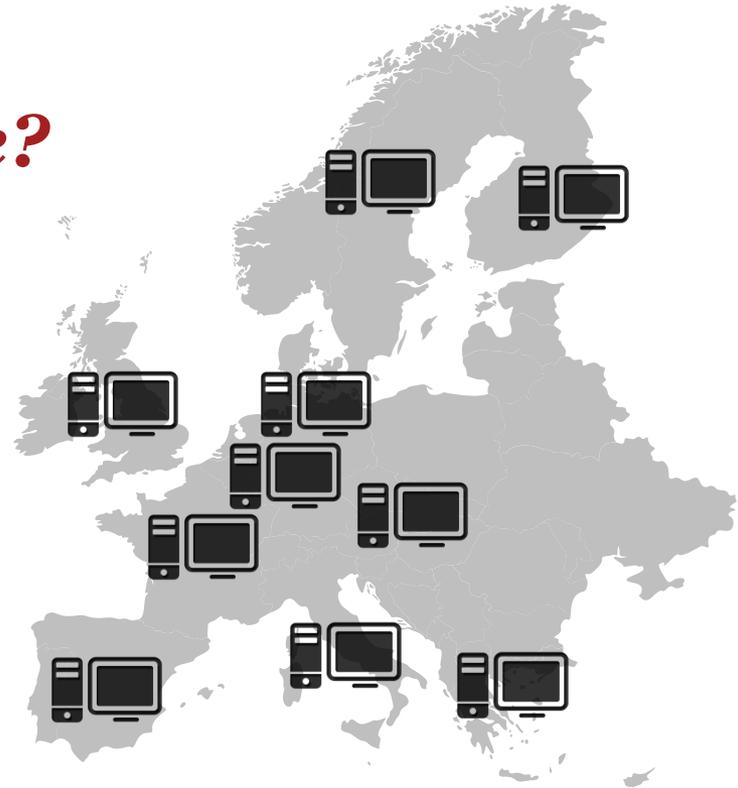
Provenance?

Licence? Persistence?

Trust? Availability?

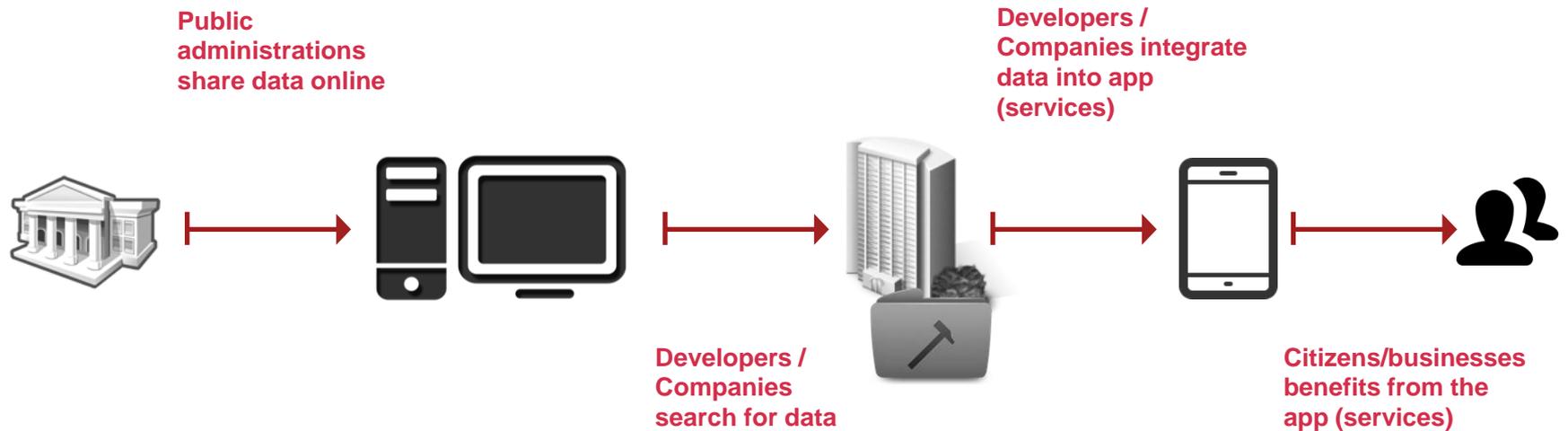
Quality?

160+



Open Data haben ein großes Potenzial, um sozialen und ökonomischen Wert zu schaffen

Publishing data



Reusing data

Hindernisse für die Veröffentlichung und Weiterverwendung von Open Data

Herausgeber

Wiederverwender

Keine Übersicht darüber, welche Daten mit höherer Wahrscheinlichkeit wiedergenutzt werden /welche ein höheres ROI-Potenzial haben.

Mangel an Überblick über vorhandene/verfügbare Datensätze

Unklare Geschäftsmodelle für die Veröffentlichung von Open Data

Unklare Geschäftsmodell für das Veröffentlichen von Open Data.

Beschränkte Toolunterstützung.

Daten sind oft von geringer Qualität, veraltet, unstrukturiert und/oder nicht maschinenlesbar

Konkurrierende Lizenzen für Datensätze

Mangel an Lizenzinformation oder inkompatible Lizenzen

Konkurrierende Vokabulare, mit denen Datensätze beschrieben werden.

Unterschiedliche Vokabulare bei der Suche nach Datensätzen

Domainspezifische Metadaten - Anforderungen

Mangel an Metadaten (von hoher Qualität)

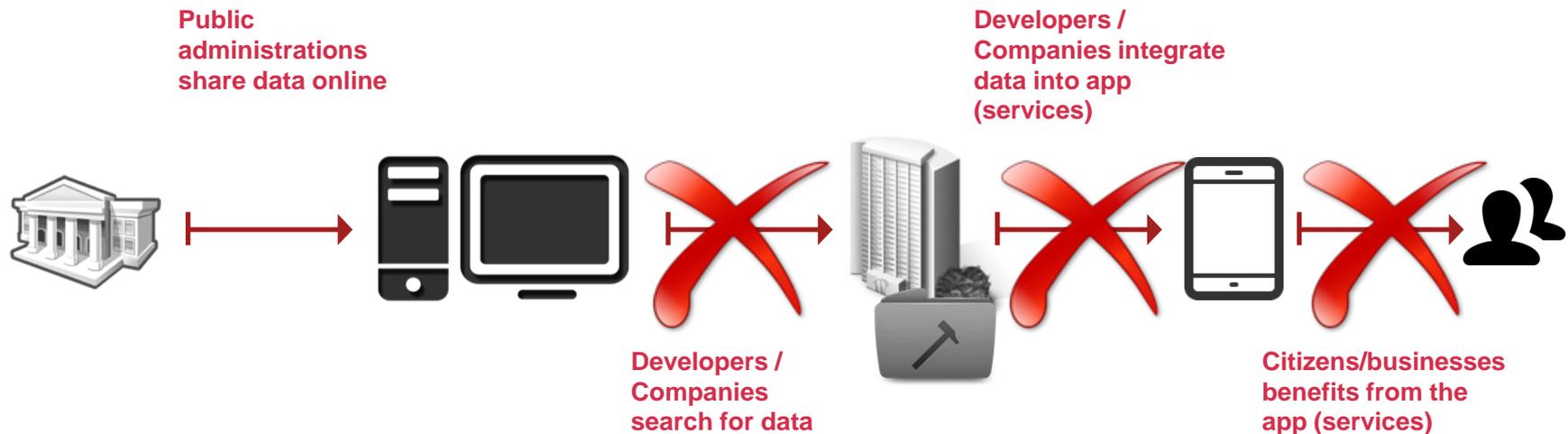
Anstrengungen, die nötig sind, um die Metadaten aktuell zu halten.

Mangel an Informationen über die Herkunft

Metadaten

Metadaten

Keine Weiterverwendung = kein sozialer und ökonomischer Wert



Open Data Support

...von der Europäischen Kommission finanziert, zielt DG CONNECT darauf ab, Zugangs- und Bewusstseinsbarrieren zu senken.

Die Aufgabe von Open Data Support...

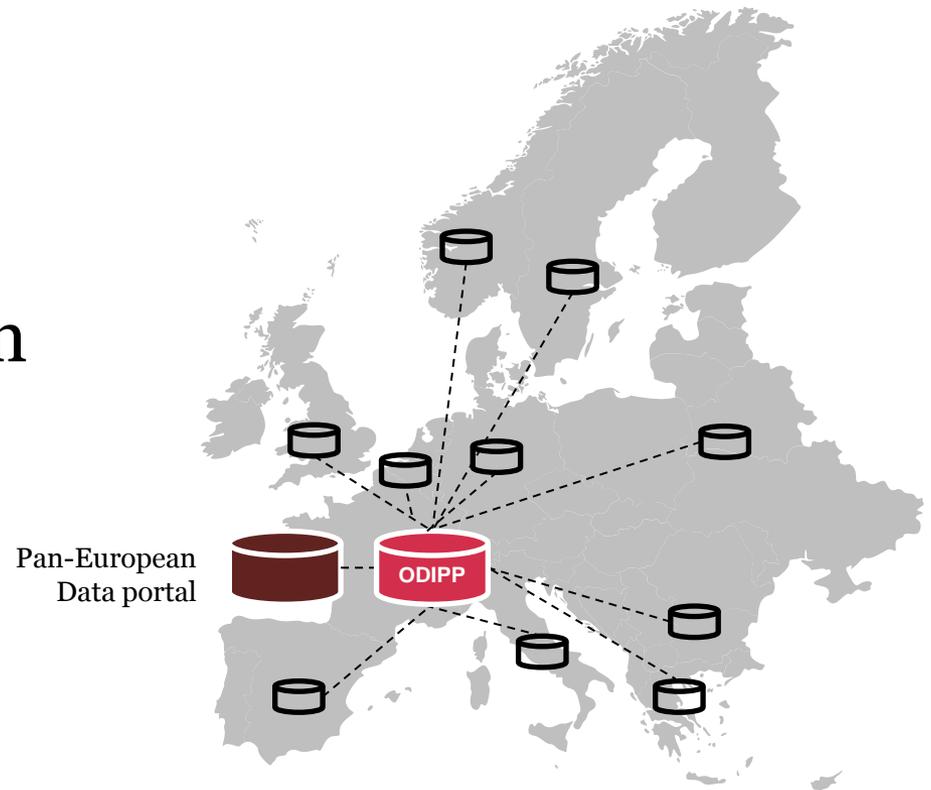
Die **Sichtbarkeit** und der **Zugang** zu Datensätzen auf lokalen und nationalen offene Datenportalen soll verbessert werden, um damit deren **Weiterverwendung** innerhalb von Grenzen und grenzüberschreitend zu steigern.

Siehe auch:

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

Mit Hilfe von...

*Einem vereinheitlichten
Zugang* zu Metadaten-
Beschreibungen von
offenen Datensätzen von
*einem einzelnen
Zugangspunkt* aus.



DCAT

Anwendungsprofil

Ein gemeinsames Vokabular für die Beschreibung von Datensätzen, die in Datenportalen gehostet sind und die auf dem „Data Catalogue vocabulary“ (DCAT) basieren.

Eine gemeinsame Initiative von...



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

Eine internationale Arbeitsgruppe von Experten

- Vorsitzender: Antonio Carneiro (Amt für Veröffentlichungen)
- 59 Arbeitsgruppenmitglieder, bestehend aus:
 - 16 europäischen Mitgliedstaaten (UK,IT,ES,DK,DE,SK,BE,AT,SE,FI,FR,IE,NL,GR,SI)
 - den USA
 - mehreren europäischen Institutionen und internationalen Organisationen
 - 40 verschiedenen Datenportalen

Siehe auch:

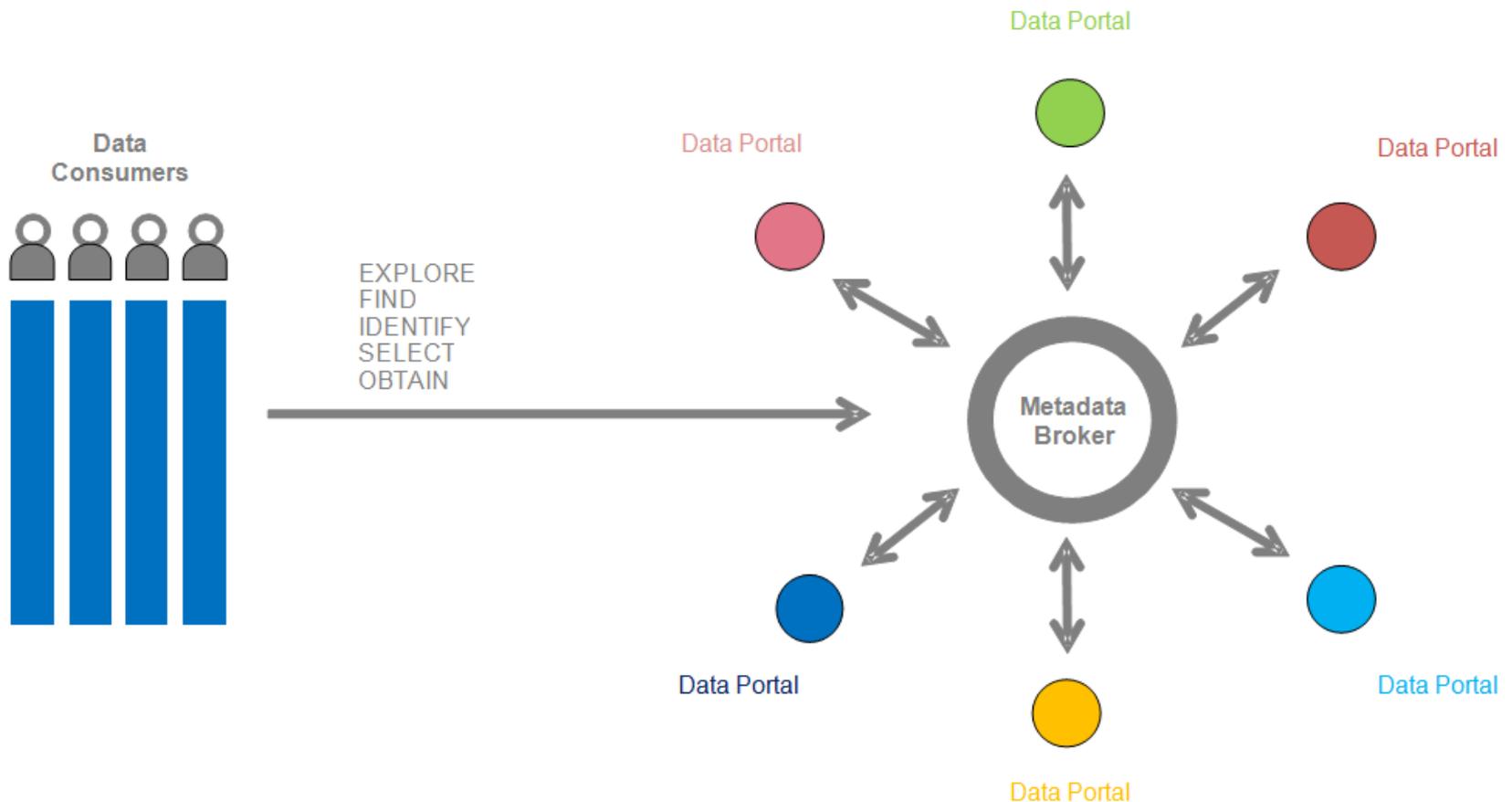
https://joinup.ec.europa.eu/asset/dcat_application_profile/description

Durch die Verwendung eines gemeinsamen Metadaten-Schema, um Datensätze zu beschreiben und Metadaten auszutauschen...

- **Datenherausgeber** erhöhen die Auffindbarkeit und so die Weiterverwendung ihrer Daten.
- **Datenverwender** können gleichmäßig über Plattformen hinweg suchen, ohne mit Schwierigkeiten konfrontiert zu sein, die durch die Verwendung von separaten Modellen oder durch Sprachunterschieden entstehen.

Die Qualität und die Verfügbarkeit der Beschreibung von Metadaten hat einen direkten Einfluss darauf, wie leicht Datensätze gefunden werden können!

Die DCAT-AP ermöglicht den Austausch von Metadaten-Beschreibungen zwischen Datenportalen



Was macht die
Spezifikation aus?

Verwendung des DCAT Anwendungsprofils

Obligatorische Klasse: Ein Empfänger von Daten MUSS in der Lage sein, Informationen über Instanzen der Klasse zu verarbeiten; ein Absender von Daten MUSS Informationen über Instanzen der Klasse liefern.

Empfohlene Klasse: Ein Empfänger von Daten MUSS in der Lage sein, Informationen über Instanzen der Klasse zu verarbeiten; ein Absender von Daten MUSS Informationen über Instanzen der Klasse liefern, wenn diese verfügbar sind.

Fakultative Klasse: Ein Empfänger von Daten MUSS in der Lage sein, Informationen über Instanzen der Klasse zu verarbeiten; ein Absender von Daten KANN die Informationen liefern, ist aber nicht verpflichtet, dies zu tun.

Obligatorische Eigenschaft: Ein Empfänger von Daten MUSS in der Lage sein, Informationen für diese Eigenschaft zu verarbeiten; ein Absender von Daten MUSS Informationen für diese Eigenschaft liefern.

Empfohlene Eigenschaft: Ein Empfänger von Daten MUSS in der Lage sein, Informationen für diese Eigenschaft zu verarbeiten; ein Absender von Daten SOLLTE Informationen für diese Eigenschaft liefern, wenn diese verfügbar sind.

Fakultative Eigenschaft: Ein Empfänger von Daten MUSS in der Lage sein, Informationen für diese Eigenschaft zu verarbeiten; ein Absender von Daten KANN Informationen für diese Eigenschaft liefern, ist aber nicht verpflichtet, dies zu tun.

Kontrollierte Vokabulare

URI Eigenschaft	Verwendet für Klasse	vorgeschlagene Vokabular
dcat:mediaType	Verteilung	MDR File types Name Authority List
dcat:theme	Datensatz	EuroVoc domains
dcat:themeTaxonomy	Katalog	EuroVoc
dct:accrualPeriodicity	Datensatz	Dublin Core Collection Description Frequency Vocabulary
dct:format	Verteilung	MDR File Type Named Authority List
dct:language	Katalog, Datensatz	MDR Languages Named Authority List
dct:publisher	Katalog, Datensatz	MDR Corporate bodies Named Authority List
dct:spatial	Katalog, Datensatz	MDR Countries Named Authority List, MDR Places Named Authority List
adms:status	Katalog, Eintragung	ADMS change type vocabulary
dct:type	Lizenz Dokument	ADMS license type vocabulary

Mapping Beispiel – data.gov.uk

Scottish Road Accident Statistics

dct:title
(Datensatz)

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



dct:description

Source agency: Scottish Government

Designation: National Statistics

Language: English *dct:language*

Alternative title: Scottish Road Accident Statistics

Licence *dct:license*

UK Open Government Licence (OGL)

Data Resources **2**

Key statistics for 2007

2007 Volume *dct:title (Distribution)*

dcat:accessURL

Details Download ▾
Details Download ▾

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:

adms:contactPoint

• Web:

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

Tags

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

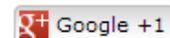
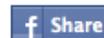
dcat:keyword

About this dataset

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

Do more with this data

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



Beispiel: Beschreibung eines Datensatzes mit der 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>
```

Wo können Sie es finden?

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, the tagline reads "Share and reuse interoperability solutions for public administrations". The breadcrumb trail indicates the current page: European Commission > ISA > Joinup > Semantic assets > Projects > Dcat application profile > Description. The main navigation menu includes My Page, Communities, Semantic Assets (selected), Software, News, Events, e-Library, and People. The content area features a sidebar with a table of contents for the Semantic Asset, including sections like Welcome, Description (selected), Members list, Issues, Asset Releases, Metrics, Highlights, Semantic Assets, Software, Communities, Communications, News, and Events. The main content displays the "DCAT application profile for data portals in Europe" with a download button, a 5-star rating, and an "Editor's choice" badge. It includes submission details (by Stijn Goedertier on March 08, 2013), a rating of 5/5, and 4024 reads. A description follows, explaining that DCAT-AP is a specification based on the Data Catalogue vocabulary for describing public sector datasets in Europe. It also mentions the "Open Data Support" service and a public review period. A "Related Content" section lists several workshops and events from 2012 and 2013.

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

Erstellen von Mappings zur DCAT-AP

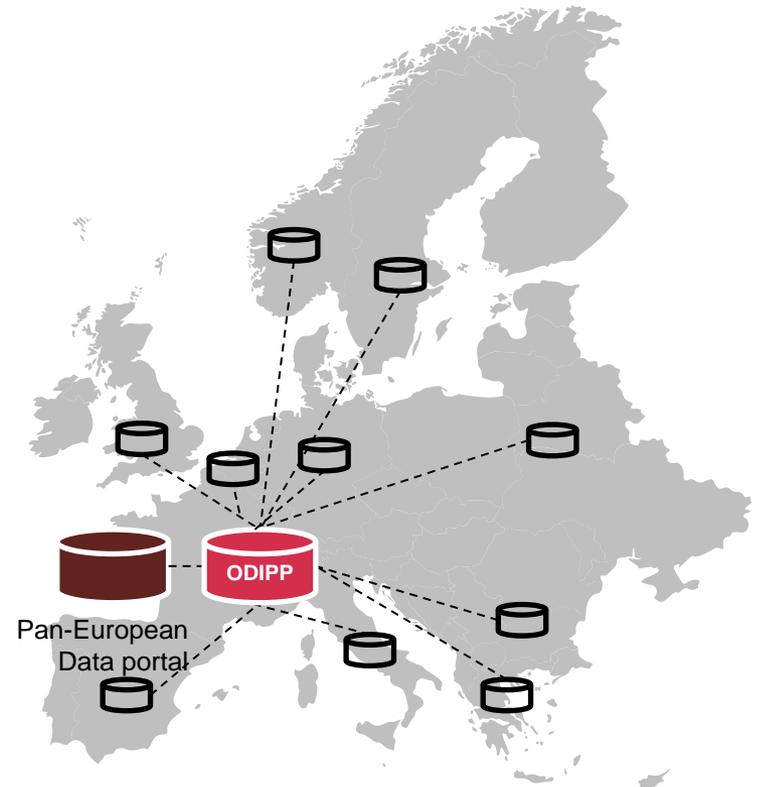
Dataset Properties	Example Value	Harmonized Predicate	Generated SPARQL
Raw Predicate			
http://data.gov.uk/predicate/title http://data.gov.uk/predicate/unpublished	Government Major Projects data for the Foreign and Commonwealth Office 2012 FALSE	dct:title	<pre> prefix dct <http://purl.org/dc/terms/> 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>

Teilen Sie Ihre Metadaten auf ODIP

Die Open Data Interoperability Platform (ODIP) ermöglicht Ihnen, Metadaten Datensätze mit DCAT-AP zu teilen, um so die Auffindbarkeit und Sichtbarkeit Ihrer Datensätze zu verbessern, was schließlich zu breiterer Weiterverwendung führen kann.

Was kann ODIP machen?

- **Sammeln** Sie Metadaten von einem offene Datenportal.
- **Verwandeln** Sie die Metadaten in RDF.
- **Stimmen** Sie die RDF-Metadaten, die in den vorherigen Schritten mit DCAT-AP produziert wurden, aufeinander ab.
- **Bestätigen** Sie die abgestimmten Metadaten in der DCAT-AP.
- **Veröffentlichen** Sie die Beschreibung der Metadaten als Linked Open Data.
- **Übersetzen** Sie Metadaten automatisch auf English



Wie kann ODIP Ihnen helfen, Ihre Metadaten zu verbessern?

- ODIP bildet Ihre Metadaten auf ein Standardmodell ab, d.h. die DCAT-AP.
- ODIP hilft Ihnen, standardisierte, mehrsprachige, kontrollierte Vokabulare in Ihren Metadaten wiederzuverwenden und ersetzt fehleranfällige Textwerte oder maßgeschneiderte Listen.
- Mittels seines Validierungsservices erlaubt Ihnen ODIP, Unstimmigkeiten und Fehler in Ihren Metadaten zu entdecken.
- ODIP ordnet Ihren Metadaten dauernde URIs zu.
- ODIP verknüpft Ihre Metadaten mit anderen Metadaten und bereichert sie durch zusätzlichen Kontext, dadurch gewinnen diese an Bedeutung.
- ODIP übersetzt automatisch die Titel und Beschreibung der Metadaten auf English

Wie sieht ODIP aus?

The screenshot displays the Open Data Interoperability Platform (ODIP) interface. The top navigation bar includes links for 'New Job', 'Manage Jobs', 'Error Reports', 'Get Support', and 'About'. The main content area is divided into two sections: a table of jobs and a detailed view of the selected job.

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

The detailed view on the right shows the following information for the selected job '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: [empty]
- Schedule Type: interval
- Scheduled
- Next run in 0 days, 16 hours, 35 minutes and 0 seconds
- Buttons: Edit Job

At the bottom, there are tabs for 'Extractors', 'Transformers', and 'Loaders'. Under 'SELECTED EXTRACTORS', the 'CKAN Extractor' is listed.

<http://odip.opendatasupport.eu>

Eine ODIP Aufgabe

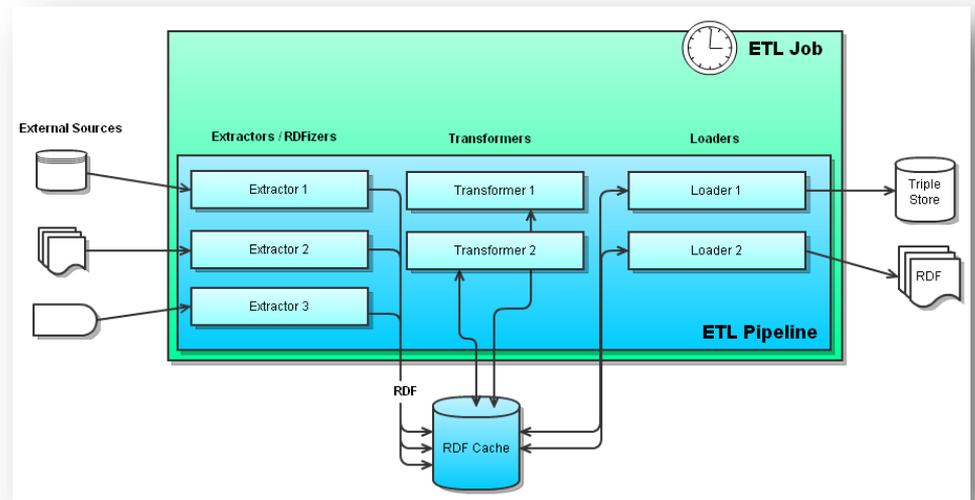
Die ODIP Aufgabe besteht aus drei möglichen Phasen, die in einer bestimmten Reihenfolge gemacht werden müssen, und die aus mehreren Plug-Ins zusammengesetzt sind.

1. Extraktion

2. Transformation

3. Laden

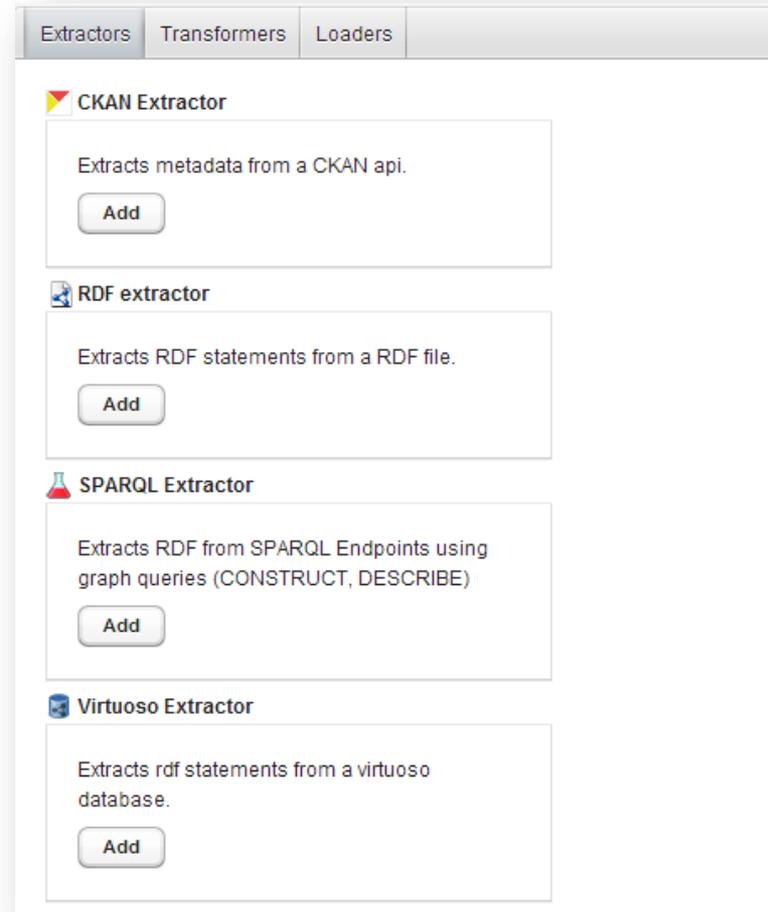
Außerdem können diese Aufgaben geplant werden, so dass sie periodisch, nacheinander oder manuell gestartet werden können.



Überblick über den
ODIP Prozess
Extrahieren-
Transformieren-Laden

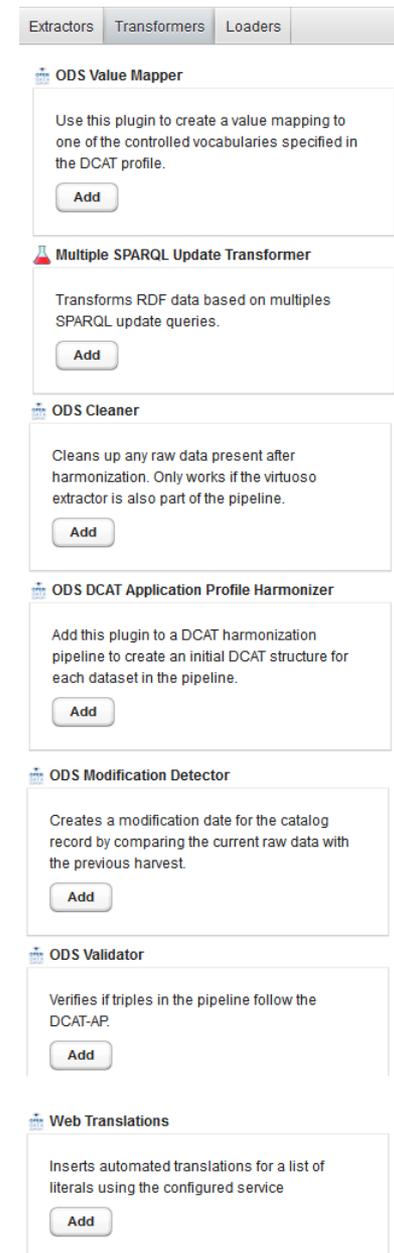
1. Extraktion

- In der Extraktionsphase werden Rohdaten aus einem gegebenen Herkunfts- Open Data Portal mit Hilfe des entsprechende Plug-Ins abgefragt, abhängig von der Technologie des Herkunftsortes.
- Mögliche Extraktoren:
 - CKAN Extractor
 - RDF Extractor
 - SPARQL Extractor
 - Virtuoso Extractor
 - CSV Extractor



2. Transformation (1/3)

- Ziel der Transformationsphase ist es, die von den Open Data Portalen gesammelten Metadaten abzugleichen, zu bereinigen und darauf vorzubereiten, auf ODIP gespeichert zu werden.
- Vorhandene Transformatoren:
 - ODS Value Mapper.
 - SPARQL Update Query Transformer.
 - ODS Cleaner.
 - ODS DCAT Application Profile Harmoniser.
 - ODS Modification Detector.
 - ODS Validator.
 - Web Translations.

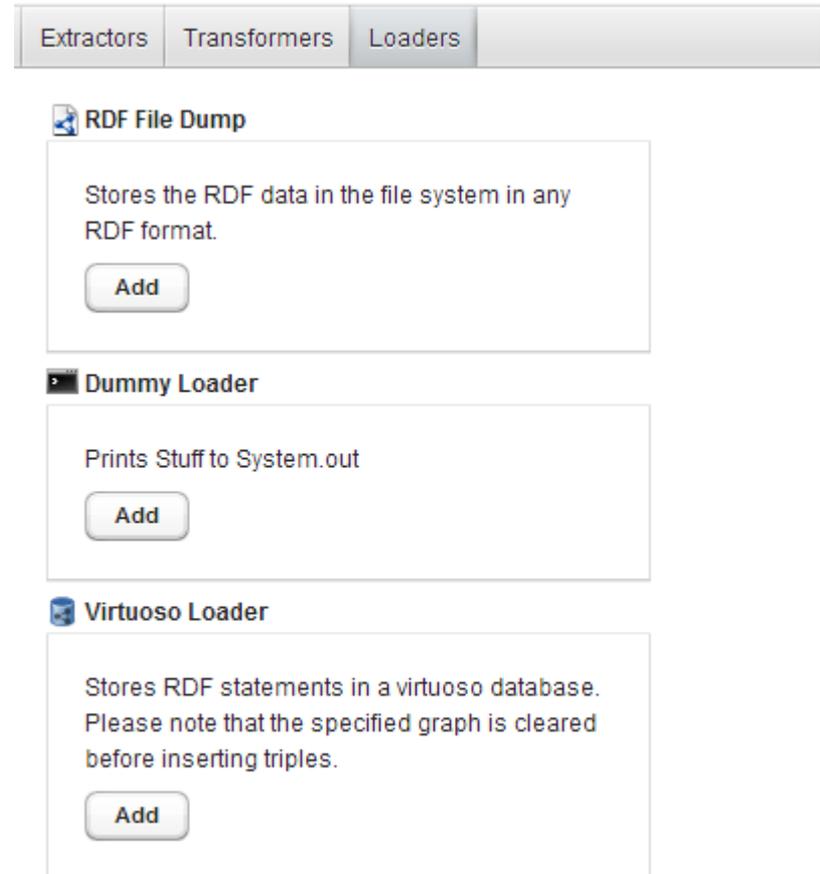


The screenshot displays a web interface for configuring data transformations. At the top, there are three tabs: 'Extractors', 'Transformers', and 'Loaders'. The 'Transformers' tab is active. Below the tabs, several transformer plugins are listed, each with a description and an 'Add' button:

- 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.

Laden

- In der Ladephase werden die gesammelten und abgeglichenen Metadaten mit Hilfe des Virtuoso Loaders auf dem Virtuoso RDF Speicher gespeichert.



The screenshot shows the 'Loaders' tab in the Virtuoso configuration interface. It lists three loader options, 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.

Beispiel

Die Übernahme eines CKAN-basierten offenen Datenportals

1. Erstellen Sie einen neuen Job auf ODIP
2. Extraktionphase
 - Fügen und konfigurieren Sie eines CKAN Extractor hinzu, um Daten von einem CKAN API zu übernehmen.
3. Transformationsphase
 - Fügen Sie ODS Value mapper hinzu
 - Fügen Sie ein SPARQL Update Query Transformer mit den relevanten Fragen hinzu
 - Fügen Sie ODS Cleaner hinzu
 - Fügen Sie den DCAT Application Profile Harmoniser hinzu
 - Fügen Sie einen Änderungsdetektor hinzu
 - Fügen Sie ODS Validator hinzu
 - Fügen Sie Web Translations hinzu
4. Ladephase
 - Laden Sie die extrahierten Daten mit Hilfe des Virtuoso Loader in einem Virtuoso RDF Store hoch
5. Erstellen Sie den Aufgabenplan auf ODIP

Beispiel – 1. Erstellen Sie eine Job auf ODIP

- Um einen neuen Job zu erstellen, klicken Sie auf “New Job”.
- Am unteren Teil des Bildschirms können Sie die aktuelle Aufgabenstellung innerhalb jeder der drei Phasen konfigurieren, indem Sie einen Tab auswählen.
- Für jede Phase können Sie die entsprechenden Module hinzufügen und konfigurieren.

Geben Sie einen Namen für den Job

Präsentieren Sie den Job mit einer kurzen Beschreibung

Drücken Sie auf den “Add”-Knopf, um die Plug-Ins, die zu deployen sein sollte, zu bestimmen

Beispiel – 2.Extraktion : Das Hinzufügen und das Konfigurieren eines CKAN Extraktors, um Daten von einem CKAN API zu übernehmen

Nach dem Hinzufügen des CKAN Extraktor Plug-Ins werden Sie aufgefordert, das folgende Formular auszufüllen:

Herausgeber, Lizenz, Titel und Beschreibung:
Verwendet im gespeicherten Katalog für die Eigenschaften von dct:publisher, dct:license, dct:title und dct:description.

Configuration: CKAN Extractor

CKAN Uri *

Publisher *

Title *

Description *

License *

Predicate Prefix *

Subject Prefix *

Ignored Keys

harvest all datasets

Configure

Der Web Standort des CKAN-Portals, das Sie übernehmen möchten.
Das Portal sollte API Version 3 unterstützen, und API muss aktiviert sein.

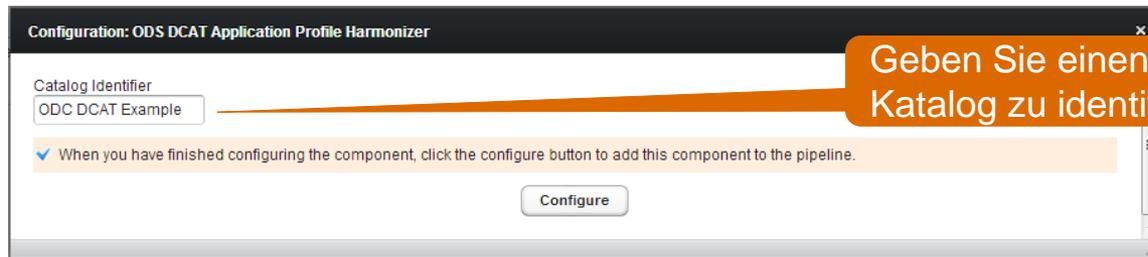
Themenpräfix: Das Präfix wird genutzt, um einen URI für die Metadaten der übernommenen Datensätze zu erstellen.
Das Thema wird erstellt als `<subjectprefix>/dataset/<datasetid>`

Prädikat Präfix: JSON-Attribute werden in Prädikate umgewandelt, indem sie an das Prädikat-Präfix angehängt werden.
Die CKAN-API-Antwort ist in JSON, die wir dann in RDF umwandeln.

Ignorierte Tasten: Eine durch Komma getrennte Liste von JSON-Attributen, die nicht in RDF-Triple umgewandelt werden sollten.

Beispiel – 3. Transformation : Hinzufügen und Konfigurieren von Plug-Ins, um Daten abzugleichen (1/3)

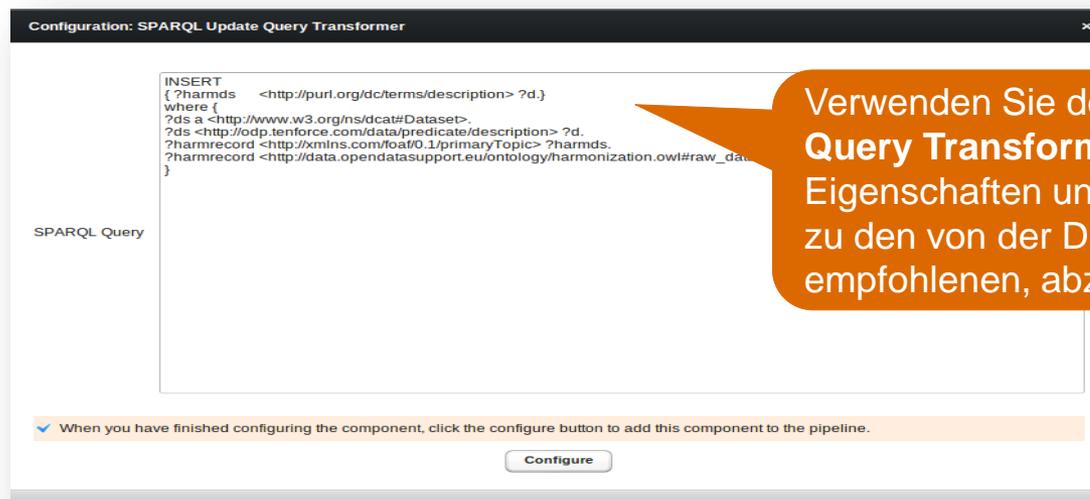
- Beginnen Sie, indem Sie den **ODS DCAT Application Profile Harmonizer** hinzufügen.
 - ✓ Dieses Plugin erstellt die harmonisierten Katalogdaten und ein Grundgerüst für jeden Datensatz, den es identifiziert.



- Verwenden Sie die **Modification Detector (Änderungsdetektor)**, um die Herkunft von Daten, die vom CKAN-Extraktor zwischen der gegenwärtigen und vorherigen Version der Rohdaten erzeugt wurde, zu vergleichen und das `dct:modified` Feld für Datensätze in Katalogen zu setzen.
 - ✓ Es ist keine Konfiguration erforderlich.

Beispiel – 3. Transformation : Hinzufügen und Konfigurieren von Plug-Ins, um Daten abzugleichen (2/3)

- Das Mapping der Beschreibung des Datensatzes zu dct:description, wie es die DCAT-AP erfordert.



Verwenden Sie den **SPARQL Update Query Transformer**, um vorhandene Eigenschaften und Werte, zusätzlich zu den von der DCAT-AP empfohlenen, abzubilden.

- Verwenden Sie die **ODS Cleaner Plugin**, um die in den aktiven Datensatz geladenen Rohdaten zu entfernen, bevor Sie diese in einer abgeglichenen Grafik speichern.
 - ✓ Keine Konfiguration ist erforderlich.

Beispiel – 3. Transformation : Das Hinzufügen und das Konfigurieren von Plug-Ins zum Abgleichen von Daten (3/3)

Das endgültige Ergebnis Ihrer Abgleichungspipeline sollte ähnlich wie das folgenden aussehen:

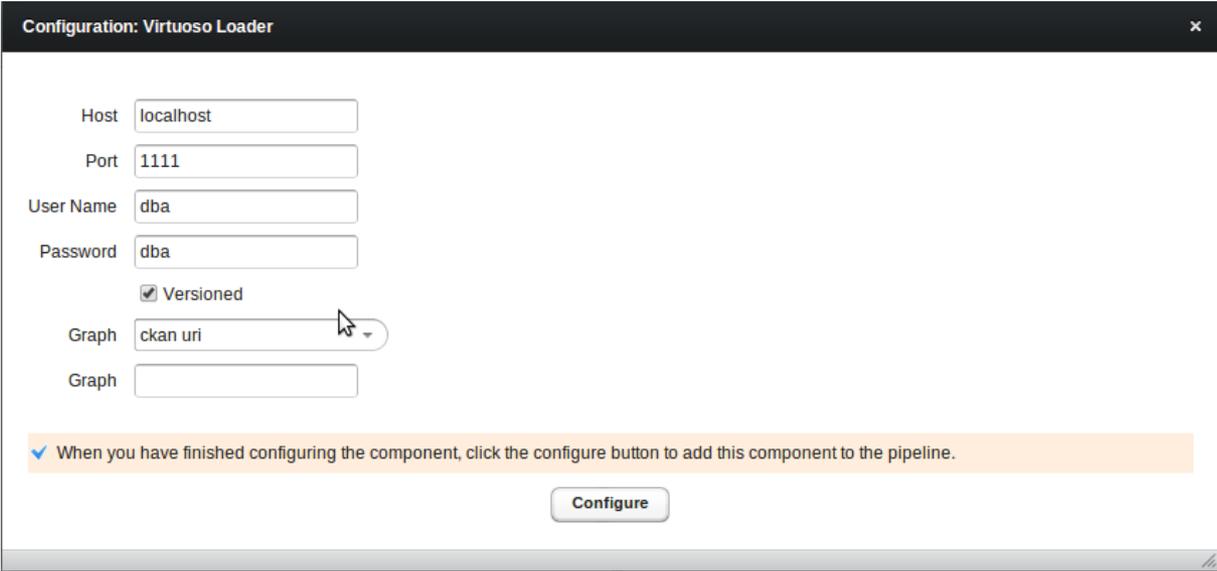
Extractors	Transformers	Loaders
SELECTED TRANSFORMERS		
	ODS DCAT Application Profile Harmonizer	
	ODS Modification Detector	
	SPARQL Update Query Transformer [INSERT { ?harmds <http://purl.org/dc/terms/p]	
	SPARQL Update Query Transformer [INSERT { ?harmds <http://purl.org/dc/terms/t]	
	SPARQL Update Query Transformer [INSERT { ?harmds <http://purl.org/dc/terms/d]	
	ODS Cleaner	

Konfigurieren Sie den Virtuoso Loader, um die abgeglichenen Daten in Virtuoso hochzuladen.

Beispiel– 4. Laden: Laden Sie die extrahierten Daten mit Virtuoso Loader in den Virtuoso RDF Store hoch

Das Virtuoso Loader wird die generierten Tripel in den Virtuoso RDF-Speicher speichern. Die Tripel werden in eine Grafikdatei Ihrer Wahl eingefügt.

Das Virtuoso Loader braucht einen Host, einen Port und die Benutzeranmeldedaten, um sich mit Ihrem Virtuoso-Server zu verbinden.



The image shows a configuration window titled "Configuration: Virtuoso Loader". It contains several input fields and a checkbox:

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

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

5. Planen Sie einen Job auf ODIP

Ein Job kann so geplant werden, dass er in einem festgelegten Intervall läuft, oder sich einem anderen Job anschließt:

- **Intervallplanung:**

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

□ Beispiel:

- 0 0 4 * * * - jeden Tag um 4 Uhr
 - 0 0 0 * * 1 - jeden Montag um Mitternacht
 - 0 30 * * * - jede Hälfte jeder Stunde vorbei
- **Verkettete Planung:** Wählen Sie einen Job, nach dem dieser Job ausgeführt werden soll.

ODIP Reporting-Tool

Jedes Mal wenn ein “Job” gelaufen ist, wird ein Bericht erstellt, den man einsehen kann, wie man im folgenden Screenshot sehen kann:

The screenshot displays the ODIP Reporting-Tool interface. On the left, there is a list of jobs with columns for ID and NAME. The main area shows a 'Last Executions' table with columns for DATE, DURATION, and COMPONENT ERRORS. Below this, there is a detailed execution report with columns for EXTRACTOR, TRANSFORMER, and LOADER, each with sub-columns for RESULT and WARNINGS.

ID	NAME
c81f979a-44a8-4385-a89a-42b93f56fd01	ireland: harmonization
88a42eaa-94e0-4a3c-95d9-f4459544d0f	ireland: raw
3b870748-571b-49c2-87d1-7ea91eba36	odp: harmonization
dc6fccb3-3cfe-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			

Wählen Sie den entsprechenden Job

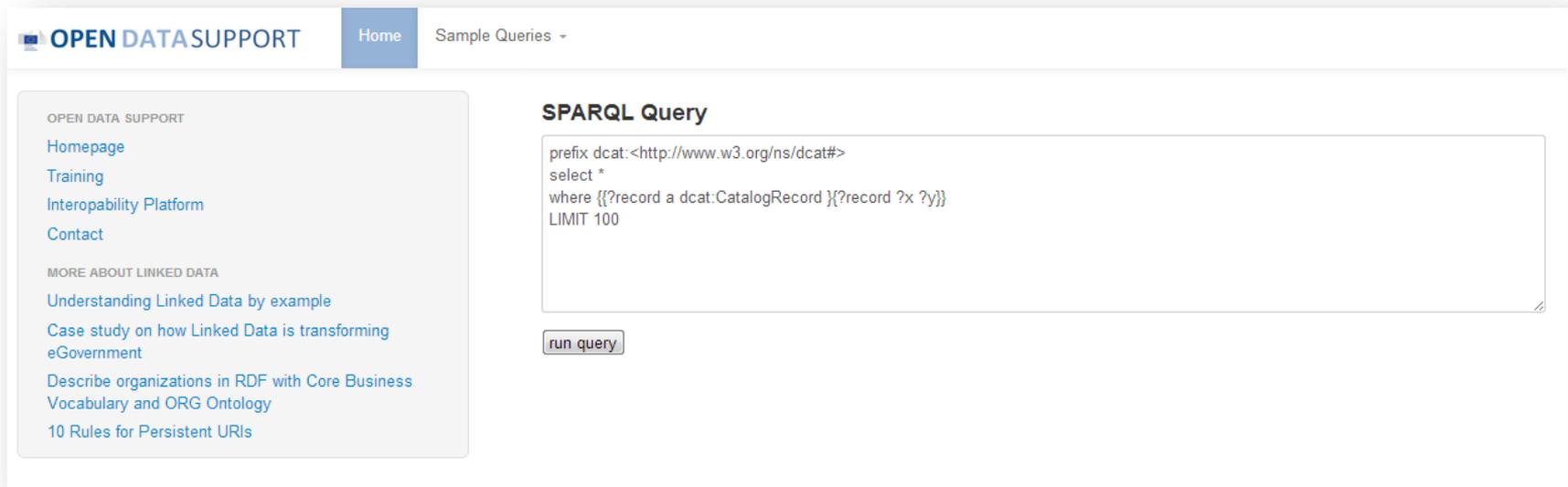
Informiert Benutzer, ob ein Plug-In funktioniert oder nicht

Entdecken Sie Datensätze durch ODIP

Die Open Data Interoperability Platform (ODIP) ermöglicht Ihnen, mit Hilfe der DCAT-AP Metadaten von beschriebenen Datensätzen zu teilen, und so die Auffindbarkeit und Sichtbarkeit Ihrer Datensätze zu verbessern, was schließlich zu einer breiteren Weiterverwendung führt.

Der öffentliche SPARQL-Endpunkt von ODIP

Abfrageschnittstelle



The screenshot shows the Open Data Support website interface. At the top left, there is a logo for 'OPEN DATASUPPORT' and a navigation menu with 'Home' and 'Sample Queries'. Below the logo, there is a sidebar with links for 'Homepage', 'Training', 'Interoperability Platform', and 'Contact'. Under the heading 'MORE ABOUT LINKED DATA', there are several articles listed, including '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 contains a text input field 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 input field is a button labeled 'run query'.

<http://data.opendatasupport.eu>

Der öffentliche SPARQL-Endpunkt von ODIP

Ergebnis-Set

OPEN DATASUPPORT

Home
Sample Queries -

OPEN DATA SUPPORT

- [Homepage](#)
- [Training](#)
- [Interoperability Platform](#)
- [Contact](#)

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](#)

SPARQL Query

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

run query

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

Mehr über ODIP



- ODIP basiert auf der [LOD Management Suite](#), die ursprünglich von der [Semantic Web Company](#) im Kontext des [LOD2](#) FP7-Projekts erstellt wurde.



- Die LOD Manager Suite wurde von TenForce im Kontext von Open Data Support für den Einsatz von ODIP erweitert.
- Es wird auf GitHub unter [GPLv2](#) verfügbar gemacht werden.

Schlussfolgerungen

- Gute Qualität von Beschreibungsmetadaten kann die Auffindbarkeit von offenen Datensätzen verbessern.
- DCAT-AP kann für das Abgleichen von Metadaten der Datensätze, die auf verschiedenen offenen Datenportalen gehostet sind, verwendet werden. Es ermöglicht deren Abfrage mit einem gleichförmigen Vokabular.
- ODIP kann die Sammlung und die Harmonisierung nach dem DCAT-AP sowie die Veröffentlichung als Linked Data der Metadaten von Datensätzen, die auf verschiedenen offenen Datenportalen veröffentlicht werden, unterstützen .
- Durch seinen öffentlichen SPARQL-Endpunkt liefert ODIP einen einzelnen Zugangspunkt zu Datensätzen aus ganz Europa
- Leichter Zugang zu Datensätzen bedeutet höhere Wiederverwendung von Datensätzen.

Gruppenfragen



<http://www.visualpharm.com>

Wie viele offene Regierungsdatenportale kennen Sie in Ihrem Land?



<http://www.visualpharm.com>

Kennen Sie in Ihrem Land Anwendungen oder Dienste, die auf offenen Regierungsdaten aufbauen?



<http://www.visualpharm.com>

Wie würden Sie die Sichtbarkeit von offenen Regierungsdatenportalen mit denjenigen von traditionellen Datenanbietern wie den nationalen statistischen Ämtern vergleichen?



<http://www.visualpharm.com>

Haben Sie von der Initiative der Open Government Data der Europäischen Kommission gehört?

Nehmen Sie auch den Online-Test!

Vielen Dank!

...und jetzt IHRE Fragen?

Referenzen

Folien 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>

Folie 12:

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

Folien 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

Folien 23-35:

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

Folie 29:

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

Verwandte Projekte und Initiativen



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/>

Werden Sie Teil unseres Teams...

*Finden Sie
uns auf*



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

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



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

<http://goo.gl/y9ZZI>

*Folgen Sie
uns auf*



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

*Begleiten Sie
uns auf*



joinup

<http://www.opendatasupport.eu>

*Kontaktieren
Sie uns unter*

contact@opendatasupport.eu