

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



OPEN DATA SUPPORT

Trainingsmodul 1.4

Einführung in das Metadaten-Management

Lernziele

Am Ende dieses Trainingsmodul sollten Sie verstehen:

- was Metadaten sind;
- Terminologie und Grundsätze der Metadatenverwaltung;
- die verschiedenen Dimensionen der Qualität von Metadaten ;
- die Nutzung von kontrollierten Vokabularen für Metadaten;
- Austausch und Erstellung von Metataden;
- Metadatenverwaltung in Open Data Support.

Inhalt

Dieses Modul enthält:

- eine Erklärung, was Metadaten sind;
- einen Überblick über den Metadaten-Lebenszyklus;
- eine Einführung in die Qualität von Metadaten;
- einen Überblick über die Metadatenverwaltung und die Herangehensweise beim Austausch durch die Anwendung von Open Data Support auf der Open Data Interoperability Platform.

Was sind Metadaten?

Definition, Beispiele und Standards der Wiederverwendung

Was sind Metadaten?

“Metadaten sind strukturierte Informationen, die beschreiben, erklären, lokalisieren, oder es sonstwie einfacher machen, eine Informationsquelle abzurufen, zu verwenden, oder zu verwalten. Metadaten werden oft Daten zu bestimmten Daten oder Informationen zu bestimmten Informationen genannt.”

-- National Information Standards Organization

<http://www.niso.org/publications/press/UnderstandingMetadata.pdf>

Metadaten liefern die nötigen Informationen dafür, dass **Daten** (z.B. Dokumente, Bilder, Datensätze), **Konzepte** (z.B. Klassifikationen) und **reale Begebenheiten** (z.B. Personen, Organisationen, Standorte, Bilder, Produkte) zusammenpassen.

Arten von Metadaten

- **Beschreibende Metadaten** beschreiben eine Ressource für die Zwecke der Entdeckung und der Identifizierung.
- **Strukturelle Metadaten**, z.B. Datenmodelle und Referenzdaten.
- **Administrative Metadaten** bieten Informationen zur Verwaltung einer Ressource.

In diesem Tutorial werden wir vor allem die Konzentration auf beschreibende Metadaten für Datensätze.

Administrative Metadaten werden auch teilweise beschrieben.

Beispiele für Metadaten

Label

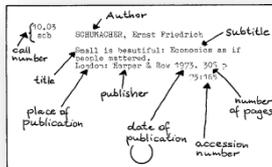


liefert Metadaten für

Can



Catalogue card



Book



Dataset description (DCAT)

```
:weather1-7 a dcat:Dataset ;
dct:title "Measurements from weather stations 1-7" ;
dct:description "Data from seven weather stations
showing temperature, humidity,
wind direction and wind speed" ;
dct:modified "2013-07-01" ;
dct:publisher <http://myweather.com/id/myweather> ;
dcat:keyword "weather" ;
dcat:landingpage <http://myweather.com/stations1-7.html> ;
dcat:distribution :weatherdata-xlsx
.

:weatherdata1-7-xlsx a dcat:Distribution ;
dct:format <http://publications.europa.eu/resource/authority/file-type/XLSX> ;
dct:licence <http://creativecommons.org/licenses/CC0> ;
dcat:downloadURL <http://myweather.com/stations1-7.xlsx>
.
```

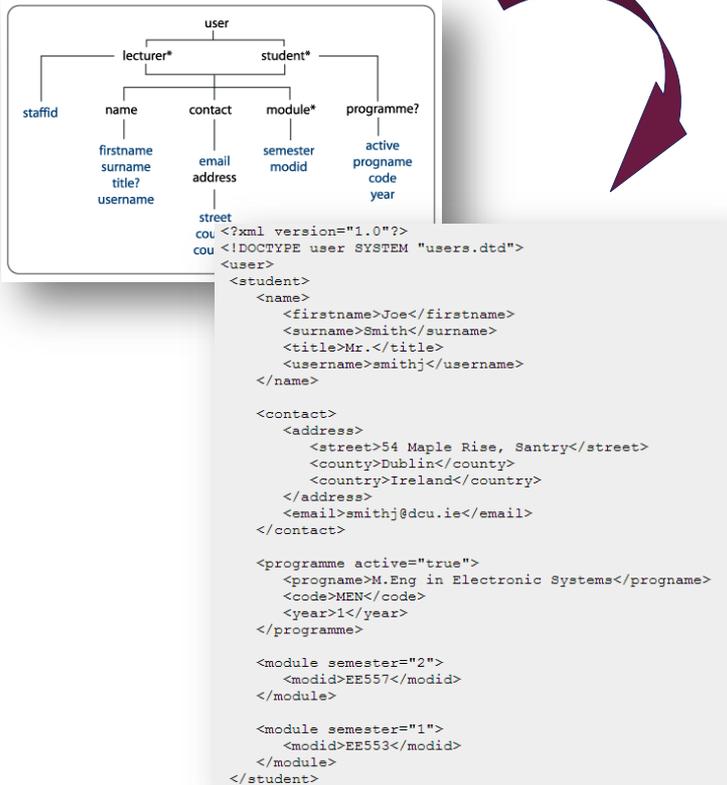
Dataset



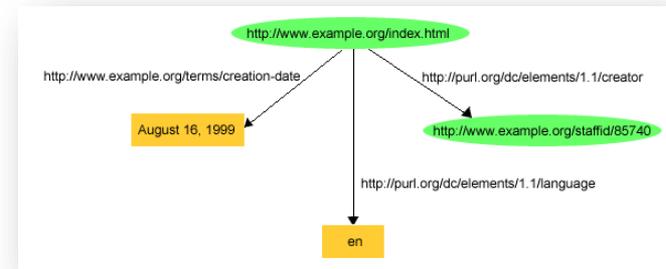
	Temp. °C	Humidity %	Wind direction	Wind speed km/h
Station 1	18.1	60	WSW	18
Station 2	17.5	59	WSW	20
Station 3	18.2	55	SW	22
Station 4	19.0	62	SW	18
Station 5	18.0	65	WSW	19
Station 6	18.2	63	SSW	21
Station 7	17.9	61	SW	22

Zwei Ansätze, um Metadaten im Web zur Verfügung zu stellen

XML (Baumstruktur- / Containerstruktur-Ansatz)



RDF (Triple-basierter Ansatz)



```
ex:index.html dc:creator exstaff:85740 .
ex:index.html exterms:creation-date "August 16, 1999" .
ex:index.html dc:language "en" .
```

Verwalten Sie die Metadaten Ihrer Datensätzen

Metadaten-Management ist wichtig

Metadaten müssen verwaltet werden, um Folgendes sicherzustellen ...

- **Verfügbarkeit:** Metadaten müssen dort gespeichert werden, wo Zugriff und Registrierung möglich sind, damit man sie finden kann.
- **Qualität:** Metadaten müssen von gleichbleibender Qualität sein, so dass die Benutzer wissen, dass sie diesen vertrauen können.
- **Persistenz:** Metadaten müssen über lange Zeiträume gespeichert werden.
- **Offene Lizenz:** Metadaten sollten unter einer öffentlichen Domain-Lizenz zur Verfügung stehen, um ihre Weiterverwendungen zu ermöglichen.

*Die Metadaten-Lebenszyklus ist **größer** als der Lebenszyklus der Daten:*

- **Metadaten können erstellt werden, bevor Daten erstellt oder erfasst werden**, z.B. um über Daten zu informieren, die in der Zukunft verfügbar sein werden.
- **Metadaten müssen aufbewahrt werden, nachdem Daten entfernt worden sind**, z.B. um über Daten zu informieren, die außer Betrieb oder zurückgenommen sind.

Metadaten-Schema

“Etikettierungs-, Markierungs- oder Codierungs-Systeme werden verwendet, um Informationen der Katalogisierung zu registrieren oder beschreibende Sätze zu strukturieren. Ein Metadaten-Schema erstellt und definiert Datenelemente sowie die Bestimmungen über die Verwendung dieser Datenelemente, um eine Ressource beschreiben.

XML Schema

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE RDF>
<rdf:RDF xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:rd="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:dcam="http://purl.org/dc/dcam/" xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  xmlns:owl="http://www.w3.org/2002/07/owl#">
  <rdf:Description rdf:about="http://purl.org/dc/terms/">
    <dcterms:title xml:lang="en">DCMI Metadata Terms - other</dcterms:title>
    <dcterms:publisher rdf:resource="http://purl.org/dc/aboutdcmi#DCMI"/>
    <dcterms:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2012-06-14</dcterms:modified>
  </rdf:Description>
  <rdf:Description rdf:about="http://purl.org/dc/terms/title">
    <rdfs:label xml:lang="en">Title</rdfs:label>
    <rdfs:comment xml:lang="en">A name given to the resource.</rdfs:comment>
    <rdfs:isDefinedBy rdf:resource="http://purl.org/dc/terms/">
    <dcterms:issued rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2008-01-14</dcterms:issued>
  </rdf:Description>
  <dcterms:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2010-10-11</dcterms:modified>
  <rdf:type rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#Property"/>
  <dcterms:hasVersion rdf:resource="http://dublincore.org/usage/terms/history/#title-002"/>
  <rdfs:range rdf:resource="http://www.w3.org/2000/01/rdf-schema#Literal"/>
  <rdfs:subPropertyOf rdf:resource="http://purl.org/dc/elements/1.1/title"/>
  </rdf:Description>
  <rdf:Description rdf:about="http://purl.org/dc/terms/creator">
    <rdfs:label xml:lang="en">Creator</rdfs:label>
    <rdfs:comment xml:lang="en">An entity primarily responsible for making the resource.</rdfs:comment>
    <dcterms:description xml:lang="en">Examples of a Creator include a person, an organization, or a service.</dcterms:description>
    <rdfs:isDefinedBy rdf:resource="http://purl.org/dc/terms/">
    <dcterms:issued rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2008-01-14</dcterms:issued>
    <dcterms:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2010-10-11</dcterms:modified>
    <rdf:type rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#Property"/>
    <dcterms:hasVersion rdf:resource="http://dublincore.org/usage/terms/history/#creator-002"/>
    <rdfs:range rdf:resource="http://purl.org/dc/terms/Agent"/>
    <rdfs:subPropertyOf rdf:resource="http://purl.org/dc/elements/1.1/creator"/>
    <rdfs:subPropertyOf rdf:resource="http://purl.org/dc/terms/contributor"/>
    <owl:equivalentProperty rdf:resource="http://xmlns.com/foaf/0.1/maker"/>
  </rdf:Description>
  <rdf:Description rdf:about="http://purl.org/dc/terms/subject">
    <rdfs:label xml:lang="en">Subject</rdfs:label>
```

RDF Schema

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE RDF>
<rdf:RDF xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:rd="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:dcam="http://purl.org/dc/dcam/" xmlns:skos="http://www.w3.org/2004/02/skos/core#"
  xmlns:owl="http://www.w3.org/2002/07/owl#">
  <rdf:Description rdf:about="http://purl.org/dc/terms/">
    <dcterms:title xml:lang="en">DCMI Metadata Terms - other</dcterms:title>
    <dcterms:publisher rdf:resource="http://purl.org/dc/aboutdcmi#DCMI"/>
    <dcterms:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2012-06-14</dcterms:modified>
  </rdf:Description>
  <rdf:Description rdf:about="http://purl.org/dc/terms/title">
    <rdfs:label xml:lang="en">Title</rdfs:label>
    <rdfs:comment xml:lang="en">A name given to the resource.</rdfs:comment>
    <rdfs:isDefinedBy rdf:resource="http://purl.org/dc/terms/">
    <dcterms:issued rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2008-01-14</dcterms:issued>
  </rdf:Description>
  <dcterms:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2010-10-11</dcterms:modified>
  <rdf:type rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#Property"/>
  <dcterms:hasVersion rdf:resource="http://dublincore.org/usage/terms/history/#title-002"/>
  <rdfs:range rdf:resource="http://www.w3.org/2000/01/rdf-schema#Literal"/>
  <rdfs:subPropertyOf rdf:resource="http://purl.org/dc/elements/1.1/title"/>
  </rdf:Description>
  <rdf:Description rdf:about="http://purl.org/dc/terms/creator">
    <rdfs:label xml:lang="en">Creator</rdfs:label>
    <rdfs:comment xml:lang="en">An entity primarily responsible for making the resource.</rdfs:comment>
    <dcterms:description xml:lang="en">Examples of a Creator include a person, an organization, or a service.</dcterms:description>
    <rdfs:isDefinedBy rdf:resource="http://purl.org/dc/terms/">
    <dcterms:issued rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2008-01-14</dcterms:issued>
    <dcterms:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2010-10-11</dcterms:modified>
    <rdf:type rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#Property"/>
    <dcterms:hasVersion rdf:resource="http://dublincore.org/usage/terms/history/#creator-002"/>
    <rdfs:range rdf:resource="http://purl.org/dc/terms/Agent"/>
    <rdfs:subPropertyOf rdf:resource="http://purl.org/dc/elements/1.1/creator"/>
    <rdfs:subPropertyOf rdf:resource="http://purl.org/dc/terms/contributor"/>
    <owl:equivalentProperty rdf:resource="http://xmlns.com/foaf/0.1/maker"/>
  </rdf:Description>
  <rdf:Description rdf:about="http://purl.org/dc/terms/subject">
    <rdfs:label xml:lang="en">Subject</rdfs:label>
```

Weiterverwendung von vorhandenen Vokabeln, um Ihre Ressourcen mit Metadaten zu versorgen

Universelle Standards und Spezifikationen:

- **Dublin Core** für veröffentlichtes Material (Text, Bild), <http://dublincore.org/documents/dcmi-terms/>
- **FOAF** für Personen und Organisationen, <http://xmlns.com/foaf/spec/>
- **SKOS** für Konzeptsammlungen, <http://www.w3.org/TR/skos-reference>
- **ADMS** für “Interoperability Assets”, <http://www.w3.org/TR/vocab-adms/>

Spezifischer Standard für Datensätze:

- **Data Catalog Vocabulary DCAT**, <http://www.w3.org/TR/vocab-dcat/>

Spezifische Verwendung von DCAT und anderen Vokabularien, um die Interoperabilität von Datenportalen in ganz Europa zu unterstützen:

- **DCAT application profile for data portals in Europe**, http://joinup.ec.europa.eu/asset/dcat_application_profile/description

Entwerfen Sie Ihr Metadaten-Schema mit RDF Schema (RDFS) – wenn möglich wiederverwendbar

RDF Schema ist besonders gut für die Kombination von Begriffen aus verschiedenen Standards und Spezifikationen.

Erfinden Sie keine Begriffe erneut, die schon woanders definiert werden, wenn sie RDF-Schemata gestalten, **verwenden Sie** Begriffe möglichst wieder.

- Zum Beispiel verwendet das Anwendungsprofil DCAT für Datenportale in Europa (DCAT-AP) Begriffe aus DCAT, Dublin Core, FOAF, SKOS, ADMS und andere wieder.

7.3. Dataset

7.3.1. Mandatory properties for Dataset

Property	URI	Range	Usage note	Card
description	dct:description	rdfs:Literal	This property contains a free-text account of the dataset. This property can be repeated for parallel language versions of the description.	1..n
publisher	dct:publisher	foaf:Organization	This property refers to an organisation responsible for making the dataset available.	1..1
title	dct:title	rdfs:Literal	This property contains a name given to the dataset. This property can be repeated for parallel language versions of the name.	1..n

7.3.2. Recommended properties for Dataset

Property	URI	Range	Usage note	Card
contact point	adms:contactPoint	v:VCard	This property contains contact information that can be used for flagging errors in the dataset or sending comments	0..n
dataset distribution	dcat:distribution	dcat:Distribution	This property links the dataset to an available distribution.	0..n
theme/ category	dcat:theme, subproperty of dct:subject	skos:Concept	This property refers to a category of the dataset. A dataset can have multiple themes.	0..n

7.3.3. Optional properties for Dataset

Property	URI	Range	Usage note	Card
frequency	dct:accrualPeriodicity	dct:Frequency	This property refers to the frequency at which dataset is published.	0..1
identifier	dct:identifier	rdfs:Literal	This property contains the main identifier for the dataset, e.g. the URI or other unique identifier in the context of the Catalog	0..n
keyword/ tag	dcat:keyword	rdfs:Literal	This property contains a keyword or tag describing the dataset.	0..n

Beispiel: Beschreibung eines offenen Datensatzs mit DCAT-AP

Beschreibung des Katalogs

```
:catalog
  a dcat:Catalog ;
  dct:title "Imaginary Catalog" ;
  rdfs:label "Imaginary Catalog" ;
  foaf:homepage <http://example.org/catalog> ;
  dct:publisher :transparency-office ;
  dct:language <http://id.loc.gov/vocabulary/iso639-1/en> ;
  dcat:dataset :dataset-001 , :dataset-002 , :dataset-003 ;
  .
```

Beschreibung des Datensatzs

```
:dataset-001
  a dcat:Dataset ;
  dct:title "Imaginary dataset" ;
  dcat:keyword "accountability","transparency" ,"payments" ;
  dct:issued "2011-12-05"^^xsd:date ;
  dct:modified "2011-12-05"^^xsd:date ;
  dct:publisher :finance-ministry ;
  dct:language <http://id.loc.gov/vocabulary/iso639-1/en> ;
  dcat:distribution :dataset-001-csv ;
  .
```

Beschreibung der Verteilung

```
:dataset-001-csv
  a dcat:Distribution ;
  dcat:downloadURL <http://www.example.org/files/001.csv> ;
  dct:title "CSV distribution of imaginary dataset 001" ;
  dcat:mediaType "text/csv" ;
  dcat:byteSize "5120"^^xsd:decimal ;
  .
```

Kontrollierte Vokabulare

Thesauren, Taxonomien und standardisierte Listen von Begriffen können für die Zuweisung von Werten zu Metadaten Eigenschaften verwendet werden.

Was sind kontrollierte Vokabulare?

Ein kontrolliertes Vokabular ist eine vordefinierte Liste von Werten, die als Werte für eine bestimmte Eigenschaft in Ihrem Metadaten Schema verwendet werden.

- Zusätzlich zur sorgfältigen Gestaltung von Schemata sind die Wertebereiche von Metadaten Eigenschaften wichtig für den Informationsaustausch und so für die Interoperabilität.
- Gemeinsame kontrollierte Vokabulare für Wertebereiche machen Metadaten systemübergreifend verständlich.

Welches kontrollierte Vokabular eignet sich für welche Objektart?

- Verwendung der **Codelisten** als kontrolliertes Vokabular für freien Text oder “String” Eigenschaften.
- Verwenden Sie **Konzepte**, die von einem **URI** in Bezug auf “Dinge” indentifiziert werden.

Beispiel DCAT-AP Eigenschaft:

keyword/ tag	dcat:keyword	rdfs:Literal	This property contains a keyword or tag describing the dataset.
--------------	--------------	--------------	---

Beispiel DCAT-AP Eigenschaft:

theme/ category	dcat:theme, subproperty of dct:subject	skos:Concept	This property refers to a category of the dataset.
--------------------	--	--------------	--

Beispiel Codelist - ObjectInCrimeClass (ListPoint)

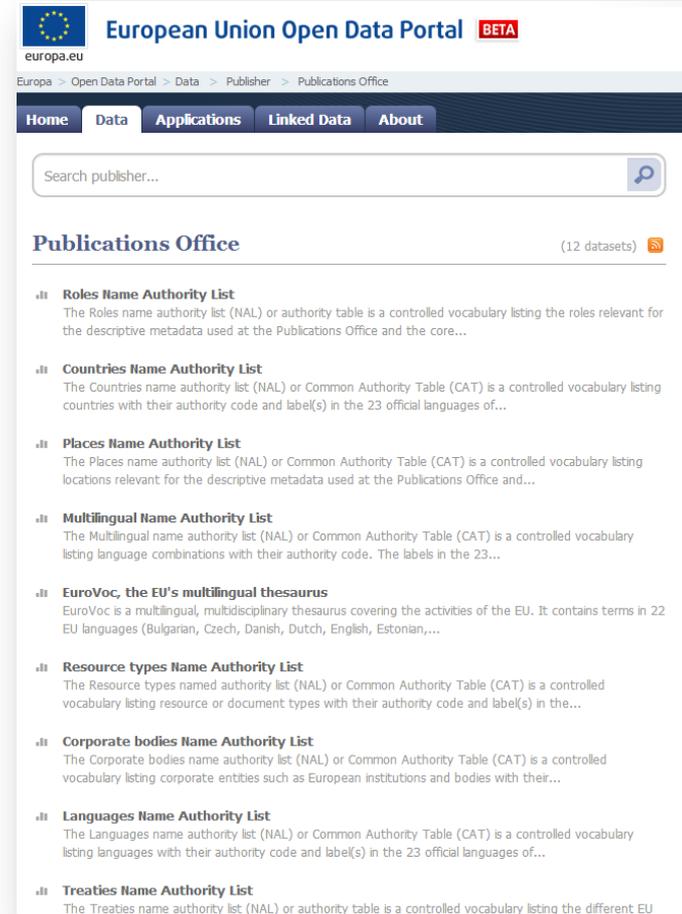
Code	Meaning
1	Used
2	Taken Without Consent
3	Eliminated
4	Not involved
5	Involved
6	Crime Weapon
7	Handled / Received
11	Exhibit

- Beispiel Taxonomie mit Begriffen, die ein URI – EuroVoc – haben

```
<rdf:Description rdf:about="http://eurovoc.europa.eu/300">
  <xl:altLabel rdf:resource="http://eurovoc.europa.eu/415040"/>
  <s04:prefLabel xml:lang="da">international kredit</s04:prefLabel>
  <s04:prefLabel xml:lang="sv">internationell kredit</s04:prefLabel>
  <s04:prefLabel xml:lang="en">international credit</s04:prefLabel>
  <s04:prefLabel xml:lang="de">internationaler Kredit</s04:prefLabel>
  <s04:prefLabel xml:lang="nl">internationaal krediet</s04:prefLabel>
</rdf:Description>
```

Beispiel – Auswahllisten (Named Authority Lists) des Amts für Veröffentlichungen

- Die Named Authority Lists bieten wiederverwendbare kontrollierte Vokabulare für:
 - Länder
 - Unternehmen
 - Dateitypen
 - interinstitutionelle Verfahren
 - Sprachen
 - Mehrsprachigkeit
 - Ressourcentypen
 - Rollen
 - Verträge



The screenshot displays the 'European Union Open Data Portal' interface. At the top, there is a search bar labeled 'Search publisher...'. Below the search bar, the 'Publications Office' section is highlighted, showing '(12 datasets)'. A list of Named Authority Lists (NALs) is provided, each with a brief description:

- Roles Name Authority List**: The Roles name authority list (NAL) or authority table is a controlled vocabulary listing the roles relevant for the descriptive metadata used at the Publications Office and the core...
- Countries Name Authority List**: The Countries name authority list (NAL) or Common Authority Table (CAT) is a controlled vocabulary listing countries with their authority code and label(s) in the 23 official languages of...
- Places Name Authority List**: The Places name authority list (NAL) or Common Authority Table (CAT) is a controlled vocabulary listing locations relevant for the descriptive metadata used at the Publications Office and...
- Multilingual Name Authority List**: The Multilingual name authority list (NAL) or Common Authority Table (CAT) is a controlled vocabulary listing language combinations with their authority code. The labels in the 23...
- EuroVoc, the EU's multilingual thesaurus**: EuroVoc is a multilingual, multidisciplinary thesaurus covering the activities of the EU. It contains terms in 22 EU languages (Bulgarian, Czech, Danish, Dutch, English, Estonian,...
- Resource types Name Authority List**: The Resource types named authority list (NAL) or Common Authority Table (CAT) is a controlled vocabulary listing resource or document types with their authority code and label(s) in the...
- Corporate bodies Name Authority List**: The Corporate bodies name authority list (NAL) or Common Authority Table (CAT) is a controlled vocabulary listing corporate entities such as European institutions and bodies with their...
- Languages Name Authority List**: The Languages name authority list (NAL) or Common Authority Table (CAT) is a controlled vocabulary listing languages with their authority code and label(s) in the 23 official languages of...
- Treaties Name Authority List**: The Treaties name authority list (NAL) or authority table is a controlled vocabulary listing the different EU

Der Metadaten- Lebenszyklus

*Erstellung, Pflege, Aktualisierung, Speicherung,
Veröffentlichung von Metadaten und der Umgang mit
Datenlöschung*

Das Erstellen Ihrer Metadaten

Metadaten Erstellung kann durch (semi-)automatische Prozesse unterstützt werden.

- Dokumenteigenschaften, in (Office)-Tools generiert, z.B. Erstellungsdatum.
- Räumliche und zeitliche Information, erfasst von Kameras, Sensoren ...
- Information aus Publikationsworkflow, z.B. der Speicherort der Datei oder URL

Allerdings erfordern andere Eigenschaften menschlichen Eingriff:

- Was ist die Ressource (z.B. Verknüpfung zu einem Fachvokabular)?
- Wie kann die Ressource verwendet werden (z.B. Verknüpfung zu einer Lizenz)?
- Wo kann ich weitere Informationen über diese Ressource finden (z.B. Verknüpfung zu einer Website oder Dokumentation, die die Ressource beschreibt)?
- Wie kann hochwertige Information eingeschlossen werden?

Das Pflegen Ihrer Metadaten

Ansätze für die Pflege von Metadaten müssen für den Datentyp, der veröffentlicht wird, geeignet sein.

- Wenn die **Daten** sich nicht **ändern**, können **Metadaten** relativ **stabil** sein. Änderungen (Massenumwandlung) können offline stattfinden, wenn nötig.
- Wenn **Daten** sich **häufig ändern** (z.B. Echtzeitsensorendaten), müssen **Metadaten** mit dem Datenworkflow eng gekoppelt und **Änderungen** praktisch **unmittelbar** sein.

Das Aktualisieren Ihrer Metadaten – Änderungen planen

Metadaten stehen in einem globalen Kontext, der ständig Änderungen ausgesetzt ist!

- **Organisation** – Abteilungen werden gebildet, mit anderen zusammengeführt, Verantwortlichkeiten werden weitergegeben.
- **Die Verwendung der Daten** – rund um die Daten entstehen neue Anwendungen.
- **Referenzdaten** – kontrolliertes Vokabulare werden entwickelt und miteinander verknüpft.
- **Datenstandards und Technologien** – Der Lebenszyklus von Technologien wird immer kürzer; wie wird das Web von morgen sein?
- **Werkzeuge und Systeme** – die Entwicklung von Speicher, Bandbreite, Mobile...

Metadaten sollten so aktuell wie möglich gehalten werden, wobei die verfügbare Zeit und das Budget berücksichtigt werden sollten.

Speichern Sie Ihre Metadaten – was sind die Optionen?

Je nach Betriebserfordernissen können Metadaten in die Daten eingebettet oder gesondert von den Daten gespeichert werden.

- Die Einbettung der Metadaten in die Daten (z.B. Office Dokumente, MP3, JPG, RDF Daten) macht den Datenaustausch leichter.
- Die Trennung der Metadaten von den Daten (z.B. In einer Datenbank) mit Links zu den entsprechenden Dateien macht die Verwaltung einfacher.

Metadaten können in einer **“klassischen” relationalen Datenbank** oder einem **RDF Triple Speicher** gespeichert werden, dies hängt von der Verfügbarkeit von Werkzeugen und den Anforderungen an Leistung und Kapazität ab.

Das Löschen von Daten

In vielen Fällen müssen Metadaten sogar nach der Löschung der Daten, die sie beschreiben, erhalten bleiben.

Stillegung oder Löschung von Daten geschieht zum Beispiel:

- Wenn Daten nicht mehr erforderlich sind.
- Wenn Daten nicht mehr gültig sind.
- Wenn Daten falsch sind.
- Wenn Daten vom Eigentümer/Herausgeber abgezogen werden.

In diesem Fall sollten die Metadaten **Informationen enthalten**, dass die Daten **gelöscht** wurden, und wenn diese **archiviert** wurde, wie und wo eine **Archivierungskopie angefordert** werden kann.

Veröffentlichung von Metadaten – Welche Optionen haben Sie?

- ‘Offene’ Veröffentlichung: direkter Zugang auf URIs:
 - Dies ist die Option, die am meisten im Einklang mit der Vision von Linked Open Data steht. Sie ermöglicht das “der-Nase-nach”-Prinzip.
- Machen Sie Ihre Metadaten durch einen **SPARQL-Endpunkt** verfügbar:
 - Dies ermöglicht, dass externe Systeme Anfragen zu einem RDF Tripel Speicher senden.
 - Dies erfordert Kenntnisse über das in dem Tripel Speicher verwendete Schema.
- Verzögerte Veröffentlichung: Zugang zu exportierten Dateien in RDF
 - Erzeugt durch die Umwandlung von Nicht-RDF-Daten in RDF-Daten.
 - Ermöglicht das offline Herunterladen von Großvolumen und das Zwischenspeichern von Daten Sammlungen.
 - Ermöglicht die Einführung einer Zu...

Siehe auch:

<http://www.slideshare.net/OpenDataSupport/licence-your-data-metadata>

Metadaten Qualität

Die Qualität und Vollständigkeit der Metadaten-Beschreibung Ihrer Datensätze, haben einen direkten Einfluss auf ihre Auffindbarkeit und Weiterverwendung.

Bei der Metadaten Qualität geht es um... (1/3)

- Die **Genauigkeit** Ihrer Metadaten: Werden die Eigenschaften der Ressourcen richtig wiedergegeben?
 - *Wenn man z.B. den richtigen Titel, die richtige Lizenz, den richtigen Herausgeber angibt, ermöglicht dies den Benutzern, diejenigen Ressourcen zu entdecken, die sie brauchen.*
- Die **Verfügbarkeit** Ihrer Metadaten: Kann man jetzt und in Zukunft auf die Metadaten zugreifen?
 - *z.B. sie für das Katalogisieren und Herunterladen verfügbar machen und sie einem regelmäßigen Backup-Prozess unterziehen.*
- Die **Vollständigkeit** Ihrer Metadaten; Sind alle relevanten Eigenschaften der Ressource genannt (soweit praktisch und wirtschaftlich möglich und für die Anwendung notwendig)?
 - *z.B. Angabe der Lizenz, die die Weiterverwendung regelt oder das Format der Verteilung, das Filter auf solche Aspekte ermöglicht.*

Bei der Metadaten Qualität geht es um... (2/3)

- Die **Konformität** Ihrer Metadaten mit anerkannten Standards: Stimmen die Metadaten mit spezifischen Metadaten-Standards oder einem Anwendungsprofil überein?
 - z.B. Die Konformität der Beschreibung eines Datensatz mit dem DCAT-AP.
- Die **Konsistenz** Ihrer Metadaten: Enthalten die Daten keine Widersprüche?
 - Sie sollten z.B. keine mehrfachen und widersprüchlichen Lizenzklärungen für dieselben Daten haben.
- Die **Glaubwürdigkeit** und **Herkunft** Ihrer Metadaten: Beruhen die Metadaten auf vertrauenswürdigen Quellen?
 - z.B. Verknüpfung zu veröffentlichten Referenzdaten und Verwaltung von einer soliden Organisation (z.B. dem EU-Amt für Veröffentlichung).

Bei der Metadaten Qualität geht es um... (3/3)

- Die **Verarbeitbarkeit** der Metadaten: Sind die Metadaten richtig maschinenlesbar?
 - *z.B. Werden Sie die Metadaten eines Datensatzes in RDF und/oder XML, und nicht als Freitext machen?*
- Die **Relevanz** der Metadaten: Enthalten die Metadaten die richtige Menge an Information für die Aufgabe?
 - *Begrenzen sie z.B. die Informationen, um die Bedürfnisse der Benutzer optimal zu bedienen?*
- Die **Rechtzeitigkeit** Ihrer Metadaten: Entsprechen die Metadaten, den tatsächlichen (aktuellen) Eigenschaften der Ressourcen und werden sie früh genug veröffentlicht?
 - *z.B. Anzeigen Sie das letzte Änderungsdatum der Ressource, um so sicher machen, dass die Metadaten frisch sind, so dass Benutzer die neuesten Informationen sehen werden.*

Austausch von Metadaten von Datensätzen

Ordnen Sie Ihre Metadaten einem gemeinsamen Metadaten Vokabular zu und tauschen Sie die Metadaten über Plattformen aus.

Vereinheitlichung von Metadaten

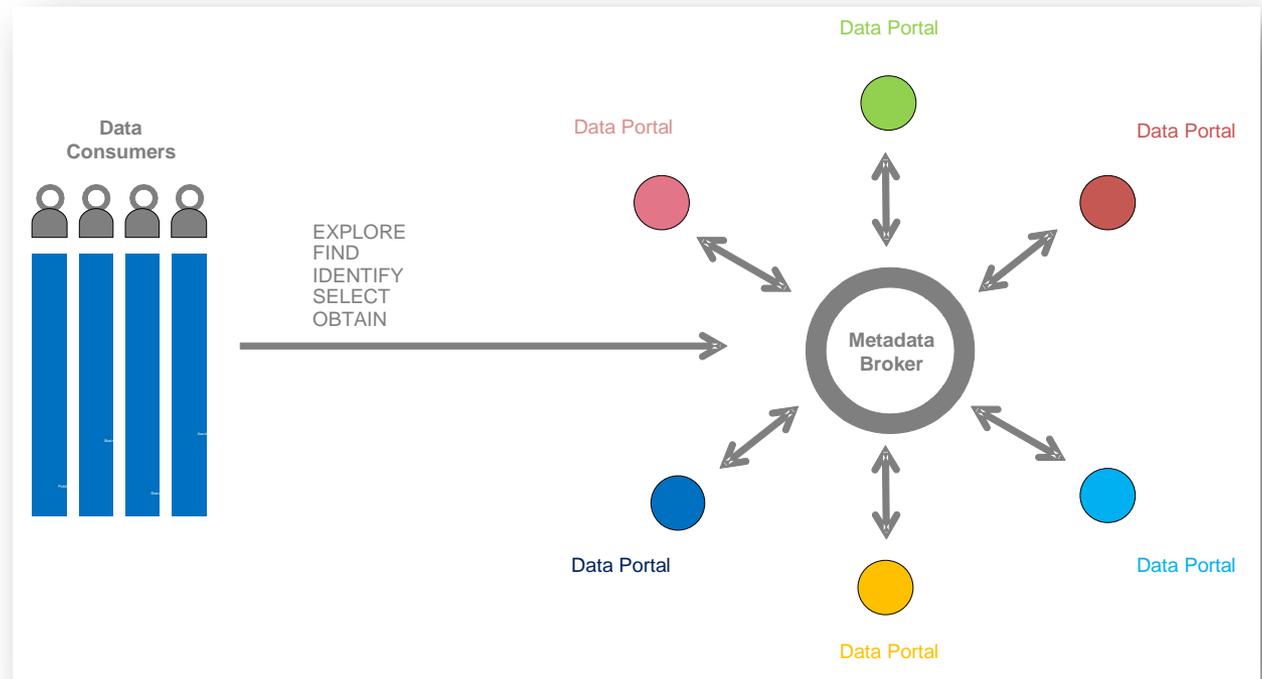
Metadaten sollten nach einem gemeinsamen Vorbild zugeordnet werden, wenn sie innerhalb von Systemen ausgetauscht werden. Damit haben Absender und Empfänger ein gemeinsames Verständnis von der Bedeutung der Metadaten.

- Auf der Schema-Ebene können Metadaten, die aus verschiedenen Quellen kommen, **auf verschiedenen Metadaten Schemata basieren**, z.B. DCAT, schema.org, CERIF, eigene interne Modell...
- Auf der Daten-(Wert)**Ebene** sollten die Metadaten-Eigenschaften den Werten von **verschiedenen kontrollierten Vokabularen** oder **Syntaxen** zugewiesen werden, z.B.:
 - Sprache: Englisch kann ausgedrückt werden als:
<http://publications.europa.eu/resource/authority/language/ENG> oder als
<http://id.loc.gov/vocabulary/iso639-1/en>
 - Datum: ISO8601 (“20130101”) versus W3C DTF (“2013-01-01”)

Beispiel: Das Homogenisieren von Metadaten zu Datensätzen

Das DCAT Anwendungsprofil für Datenportale in Europa

Das DCAT-AP kann als gemeinsames Modell für den Austausch von Metadaten mit offenen Daten-Plattformen in Europa verwendet werden und/oder mit einer Daten-Vermittlung (z.B. die Open Data Interoperability Platform - ODIP).



Siehe auch:

http://joinup.ec.europa.eu/asset/dcat_application_profile/home

Abbildungsbeispiel- 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.



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) [OPEN DATA](#)

Data Resources **2**

Key statistics for 2007

2007 Volume

dct:title (Distribution)



Dcat:accessURL

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

Additional Information

Openness score	☆☆☆☆☆
Geographic coverage	Scotland dct:spatial
National statistic	yes
ONS Category	Travel and Transport dct:theme
Temporal coverage	No value dct:temporal
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



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



Tweet



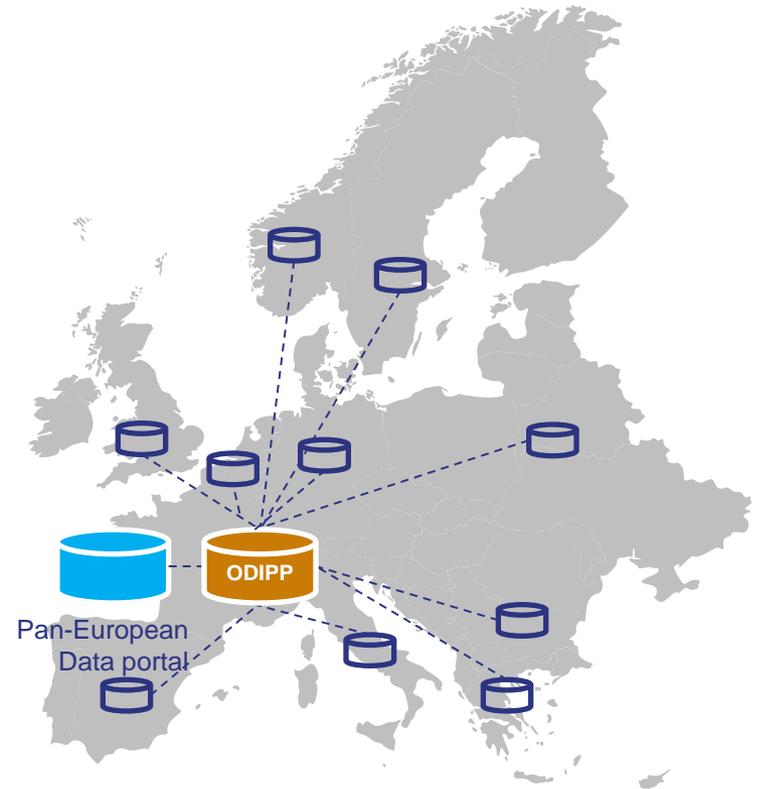
Share



Google +1

Was kann die Open Data Interoperability Platform?

- Sie **sammelt** Metadaten von einem offenen Datenportal.
- Sie **verwandelt** die Metadaten in RDF.
- Sie **vereinheitlicht** die RDF-Metadaten, die in den vorherigen Schritten mit DCAT-AP produziert worden sind.
- Sie **bestätigt** die harmonisierten Metadaten mit Hilfe der DCAT-AP.
- Sie **veröffentlicht** die Beschreibung der Metadaten als Linked Open Data.



Siehe auch:

<http://www.slideshare.net/OpenDataSupport/promoting-the-re-use-of-open-data-through-odip>

Schlussfolgerungen

- Metadaten liefern Informationen über Ihre Daten und Ressourcen. Die Qualität der Metadaten beeinflusst direkt die Auffindbarkeit und Weiterverwendung Ihrer Ressourcen.
- Für die Metadatenverwaltung sollte ein strukturierter Ansatz verfolgt werden.
- Der Metadaten-Lebenszyklus verlängert die Lebensdauer von Datensätzen (Metadaten vor der Veröffentlichung und nach dem Löschen).
- Vereinheitlichte Metadaten ermöglichen den Einsatz von Metadaten-Brokern, die die Zugangsbarrieren zu Ihren Ressourcen wiederum senken, was zu einer verbesserten Sichtbarkeit und Auffindbarkeit führen sowie das Wiederverwendungspotenzial steigern kann.

Gruppenübung und Fragen



<http://www.visualpharm.com>

Wählen Sie in Zweiergruppen einen Datensatz aus Ihrem Land aus und beschreiben Sie ihn mit dem DCAT Anwendungsprofil.



<http://www.visualpharm.com>

Hat Ihre Organisation einen minimalen Satz von Metadaten, die zusammen mit Open Data zur Verfügung gestellt werden können?



<http://www.visualpharm.com>

Was sind Ihrer Meinung nach die größten Hindernisse in Ihren Metadaten für die (Wieder-) Verwendung von kontrollierten Standardvokabularen?



<http://www.visualpharm.com>

Haben Sie auf der betrieblichen Ebene eine Methodik für die Verwaltung von Daten und/oder Metadaten?

Nehmen Sie auch den Online-Test!

Vielen Dank!

... Und jetzt IHRE Fragen?

Diese Präsentation wurde von Open Data Support erstellt

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.

Autoren:

Maxx Dekkers, Michiel De Keyzer, Nikolaos Loutas and Stijn Goedertier

Referenzen

Folie 6,7:

- NISO. Understanding Metadata. <http://www.niso.org/publications/press/UnderstandingMetadata.pdf>

Folie 9:

- Dublin City University. Chapter 3: Introduction to XML. <http://wiki.eeng.dcu.ie/ee557/q2/326-EE.html>
- W3C. RDF Primer. <http://www.w3.org/TR/rdf-primer/>

Folie 12:

- http://gondolin.rutgers.edu/MIC/text/how/catalog_glossary.htm
- Dublin Core. Example XML Schema. <http://dublincore.org/schemas/xmls/qdc/dc.xsd>
- Dublin Core, Example RDF Schema. <http://dublincore.org/2012/06/14/dcterms.rdf>

Folien 14, 33:

- The ISA Programme. DCAT Application Profile for Data Portals in Europe - Final Draft. https://joinup.ec.europa.eu/asset/dcat_application_profile/asset_release/dcat-application-profile-data-portals-europe-final-draf

Folie 18:

- ListPoint. ObjectInCrimeClass. <http://www.listpoint.co.uk/CodeList/details/ObjectInCrimeClass/1.2/1>

Folien 19:

- Publications Office. Countries Name Authority List. <http://open-data.europa.eu/en/data/dataset/2nM4aG8LdHG6RBMumfkNzQ>

Weiterführende Lektüre



Understanding Metadata, NISO.

<http://www.niso.org/publications/press/UnderstandingMetadata.pdf>



Ben Jareo and Malcolm Saldanha. The value proposition of a metadata driven data governance program. Best Practices Metadata. May 2012.

https://community.informatica.com/mpresources/Communities/IW2012/Docs/bos_30.pdf



John R. Friedrich, II. Metadata Management Best Practices and Lessons Learned. The 10th Annual Wilshire Meta-Data Conference and the 18th Annual DAMA International Symposium. April 2006.

<http://www.metaintegration.net/Publications/2006-Wilshire-DAMA-MetaIntegrationBestPractices.pdf>

Verwandte Initiativen



Metadata Management. Trainer screencasts, <http://managemetadata.com/screencasts/msa/>



MIT Libraries. Data Management and Publishing. Reasons to Manage and Publish Your Data, <http://libraries.mit.edu/guides/subjects/data-management/why.html>



ISA Programme. DCAT Application Profile for European Data Portals, https://joinup.ec.europa.eu/asset/dcat_application_profile/description



Generating ADMS-based descriptions of assets using Open Refine RDF, <https://joinup.ec.europa.eu/asset/adms/document/generate-adms-asset-descriptions-spreadsheet-refine-rdf>

The Dublin Core Metadata Initiative, <http://dublincore.org/>



Werden Sie Teil unseres Teams...

*Finden Sie
uns auf*



Open Data Support

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



Open Data Support

<http://goo.gl/y9ZZI>

*Folgen Sie
uns auf*



@OpenDataSupport

*Begleiten Sie
uns auf*



joinup

<http://www.opendatasupport.eu>

*Kontaktieren
Sie uns unter*

contact@opendatasupport.eu