



OGD Schweiz

dcat:Dataset

Per 09.03.2015

1 Einleitung

Ein Metadaten-Record der Klasse dcat:Dataset (Dataset-Metadaten-Record) beschreibt *eine einzelne, thematisch geschlossene Datensammlung*. Die für den Inhalt verantwortliche Stelle (Institution, Person, usw.) bietet den Download bzw. den Zugang zur Datensammlung an.

Eine Datensammlung kann in unterschiedlichen Formen verfügbar sein (z.B. als Datei im Format text/csv oder als Datenbank mit Programmierschnittstelle). Jede Publikationsform eines Datensatzes wird mit einem eigenen Metadaten-Record der Klasse dcat:Distribution (Distribution-Metadaten-Record) beschrieben.

In der Regel gehört zu jedem Dataset-Metadaten-Record mindestens ein Distribution-Metadaten-Record. Dies gilt selbst dann, wenn die betreffende Publikationsform des Datensatzes nur auf der Webseite der anbietenden Stelle verfügbar ist und das Element dcat:landingPage im Dataset-Metadaten-Record bereits auf diese verweist.

In Ausnahmefällen kann der Katalog Dataset-Metadaten-Records enthalten, die noch nicht veröffentlichte Datensätze beschreiben. In diesen Fällen gibt es keinen Distribution-Metadaten-Record. Dies ist z.B. der Fall, wenn für die Inventarisierung Datasets beschrieben werden, die (noch) nicht veröffentlicht werden.

Die zukunftsorientierte Implementation eines Katalogs, der dem Schweizer DCAT Profil entspricht, ermöglicht die Verlinkung von Dataset-Metadaten-Records und damit die Vernetzung verwandter Datensätze (linked data). Auf diese Weise können NutzerInnen des Katalogs die thematisch verknüpften Datensätze rasch finden und Zusammenhänge zwischen den Daten entdecken. Das Schweizer DCAT Profil umfasst deshalb zusätzliche Elemente, die im Standard nicht vorgesehen sind.

2 Die Elemente des Metadaten-Records der Klasse „dcat:Dataset“

Hinweis zur Tabelle: Farbig hinterlegte Zeilen markieren zwingend notwendige Elemente, d.h. Elemente der Kardinalität 1..1 oder 1..n.

dcat:Dataset				
element	type	content	cardinality	remarks concerning the European profile
dct:title (lang="...")	rdfs:Literal	A name given to the dataset	mandatory [1..n]	"mandatory" mit Kardinalität 1..n
<p>Das Element "dct:title" kann mehrfach vorkommen, um den Titel in mehreren Sprachen auszudrücken. Auf der Benutzeroberfläche des Portals wird jeweils der Inhalt desjenigen Elements angezeigt, dessen Sprache der von der NutzerIn gewählten Anzeigesprache entspricht.</p> <p>Falls der Titel im Dataset-Metadaten-Record nicht in allen Anzeigesprachen enthalten ist, zeigt das System auf der Benutzeroberfläche den passenden Titel (falls vorhanden) oder den im ersten Element "dct:title" enthaltenen Titel an.</p> <p>Mehrsprachige Titel bedeuten nicht, dass der beschriebene Datensatz in mehreren Sprachen verfügbar ist. Falls der Datensatz in verschiedenen Sprachen verfügbar ist, muss das Element "dct:title" aber in den entsprechenden Sprachen geführt werden. Dies erlaubt es, die Distributionen auf der Benutzeroberfläche des Portals passend zu bezeichnen, auch wenn in den Distribution-Metadaten-Records keine Elemente "dct:title" enthalten sind.</p>				
dct:description (lang="...")	rdfs:Literal	Free-text account of the dataset	mandatory [1..n]	"mandatory" mit Kardinalität 1..n
<p>Das Element "dct:description" kann mehrfach vorkommen, um die Beschreibung des Datensatzes in mehreren Sprachen auszudrücken. Auf der Benutzeroberfläche des Portals wird jeweils der Inhalt desjenigen Elements angezeigt, dessen Sprache der von der NutzerIn gewählten Anzeigesprache entspricht (falls vorhanden) oder die im ersten Element "dct:description" enthaltene Beschreibung.</p>				
dct:issued	rdfs:Literal (typed as xsd:date)	Date of formal issuance (e.g., publication) of the dataset (this property should be set using the first known date of issuance)	conditional [0..1]	"optional"
<p>Das Element "dct:issued" im Dataset-Metadaten-Record stellt das Datum der erstmaligen Veröffentlichung des Datensatzes dar. Falls dieses Datum nicht bekannt ist, kann das Datum der erstmaligen Veröffentlichung des Datensatzes im Katalog eingetragen werden.</p> <p>Falls zum Datensatz, den dieser Dataset-Metadaten-Record beschreibt, (noch) keine Distribution veröffentlicht ist (z.B. bei Inventar-Datensätzen), kann das Element "dct:issued" weggelassen werden. Es kann aber durchaus ein Element "dct:issued" eingefügt werden, auch wenn noch keine Distribution veröffentlicht ist (z.B. mit dem Datum der vorgesehenen Veröffentlichung). In beiden Fällen kommt dann im Dataset-Metadaten-Record kein Element "dcat:distribution" vor.</p> <p>Auf der Benutzeroberfläche des Portals wird der Datensatz nur angezeigt, wenn ein Element "dct:issued" vorhanden ist und das Veröffentlichungsdatum nicht in der Zukunft liegt.</p>				

dct:modified	rdfs:Literal (typed as xsd:date)	Most recent date on which the dataset was changed, updated or modified (the value of this property indicates a change to the actual dataset, not a change to the catalog record. An absent value may indicate that the dataset has never changed after its initial publication, or that the date of the last modification is not known, or that the dataset is continuously updated).	conditional [0..1]	"optional" mit der Kardinalität 0..1
<p>Falls die Daten, die dieser Dataset-Metadaten-Record beschreibt, seit der Erstveröffentlichung auf dem Portal geändert haben (z.B. eine Berichtigung der Werte erfolgt ist), muss das Element "dct:modified" das Datum der letzten Änderung wiedergeben. Auf diese Weise kann eine Person oder Institution, die den Datensatz für eine eigene Auswertung oder Anwendung genutzt hat, erkennen, ob sie die Auswertung oder Anwendung eventuell aktualisieren muss.</p>				

dct:publisher	foaf:Agent	An entity responsible for making the dataset available.	mandatory [1..n]	"recommended"
<p>Das Element "dct:publisher" nennt die Stelle, die den Datensatz veröffentlicht, beim Namen. Bei Stellen, für die in TERMDAT ein Eintrag vorhanden ist, wird zusätzlich die betreffende Identifikation eingetragen.</p> <p>Falls der Datenlieferant nur die TERMDAT-Identifikation mitliefert, ermittelt die Portal-Software in TERMDAT den zugeordneten Namen und ergänzt das Element „dct:publisher“ entsprechend.</p> <p>Falls der Datenlieferant nur den Namen mitliefert, ermittelt die Portal-Software in TERMDAT die zugeordnete Identifikation und ergänzt das Element „dct:publisher“ entsprechend. Findet die Portal-Software den Namen in TERMDAT nicht, wird im Element „dct:publisher“ kein Verweis auf TERMDAT eingefügt; in diesem Fall soll der Name eine Kantonsangabe enthalten (z.B. „Amt für Informatik (GR)“).</p> <p>Falls der Datenlieferant sowohl den Namen als auch die TERMDAT-Identifikation mitliefert, übernimmt die Portal-Software beide Angaben unverändert.</p> <p>Wenn mehrere Stellen an der Publikation des Datensatzes beteiligt sind, können diese mit je einem Element „dct:publisher“ aufgeführt werden.</p> <p>Zur Darstellung wird systemintern die Form <dct:publisher> <rdf:Description about="Verweis auf TERMDAT-Eintrag"> <rdfs:label>Name der veröffentlichenden Stelle</rdfs:label> </rdf:Description> </dct:publisher> verwendet.</p>				

dcat:contact-Point	vcard:Kind	Link a dataset to relevant contact information	mandatory [1..n]	"optional"
<p>Dieses Element soll den Verweis auf eine E-Mail-Adresse enthalten, die von der publizierenden Stelle auch wirklich bedient wird.</p> <p>Wenn mehrere Stellen an der Publikation des Datensatzes beteiligt sind, kann das Element „dcat:contactPoint“ mehrfach vorkommen.</p>				

dct:language	rdfs:Literal (ISO 639-1 two-letter code)	The language of the dataset (if the dataset is available in multiple languages, use multiple values for this property. If each language is available separately, define an instance of dcat:Distribution for each language and describe the specific language of each distribution using dct:language.	conditional [0..n]	"optional"
<p>Falls die Daten in einer sprachunabhängigen Form vorliegen (oder gar kein Distribution-Metadaten-Record vorhanden ist), kann das Element "dct:language" weggelassen werden.</p> <p>Für alle in den Distributionen vorkommenden Sprachen muss je ein Element „dct:language“ vorgesehen werden. (Eine Distribution kann mehrsprachig sein, oder nur in einer einzelnen Sprache vorliegen.)</p>				
dcat:theme	skos:Concept	The main category of the dataset (a dataset can have multiple themes; themes are part of the knowledge organization system (KOS) used to classify the catalog's datasets)	mandatory [1..n]	"recommended"
<p>Der Inhalt eines Elements "dcat:theme" muss einem Begriff aus der Klassifikation des skos:ConceptScheme des DCAT Application Profile for Swiss data portals (Categories) entsprechen.</p>				
dcat:distribution	dcat:Distribution	Connects a dataset to its available distributions	conditional [0..n]	"recommended"
<p>In der Regel zeigt ein Element "dcat:distribution" auf einen entsprechenden Distribution-Metadaten-Record.</p> <p>In Ausnahmefällen kann mit dem Dataset-Metadaten-Record ein Dataset beschrieben werden, zu dem (noch) keine Distributionen existieren. In diesem Fall wird das Element "dcat:distribution" weggelassen.</p>				
dct:relation	rdf:Property	Recommended best practice is to identify the related resource by means of a string conforming to a formal identification system	optional [0..n]	Im Europäischen Profil nicht vorgesehen
<p>Wenn Datensätze direkt (nicht über eine landing page) zur Verfügung gestellt werden, sind eventuell ausführlichere Angaben zu Struktur und Inhalt der Daten nötig. In diesen Fällen kann auf ein (oder mehrere) Dokument(e) mit den erläuternden Angaben verwiesen werden.</p> <p>Zur Darstellung wird systemintern die Form <dct:relation> <rdf:Description about="Verweis auf Dokument"> <rdfs:label>Bezeichnung des betreffenden Dokuments</rdfs:label> </rdf:Description> </dct:relation> verwendet.</p>				
dcat:keyword	rdfs:Literal	A keyword or tag describing the dataset	optional [0..n]	"recommended"
<p>Als "keywords" empfehlen sich die Begriffe aus TERMDAT. Zusätzlich nötige Begriffe sollen nach einem noch zu definierenden kontrollierten Prozess TERMDAT hinzugefügt werden.</p> <p>Das System erleichtert den Nutzer/innen die Auswahl von Begriffen, indem es bei der Eingabe dynamisch Ausschnitte aus der Begriffsliste anzeigt. Ausserdem fügt das System bei der Suche automatisch die in TERMDAT definierten Synonyme der eingegebenen Begriffe ein.</p>				

dcat:landingPage	foaf:Document	A Web page that can be navigated to in a Web browser to gain access to the dataset, its distributions and/or additional information; if the distribution(s) are accessible only through a landing page (i.e. direct download URLs are not known), then the landing page link should be duplicated as accessURL on a distribution.	conditional [0..1]	"optional"
-------------------------	---------------	---	-------------------------------------	------------

Das Element "dcat:landingPage" wird benötigt, wenn der Datensatz, den der Dataset-Metadaten-Record beschreibt, nur indirekt (über eine Webseite der publizierenden Stelle) verfügbar ist. In diesem Fall muss im Katalog trotzdem ein Distribution-Metadaten-Record zugeordnet sein, wobei der Inhalt des Elements "dcat:accessURL" im Distribution-Metadaten-Record mit dem Inhalt des Elements "dcat:landingPage" übereinstimmen muss.

dct:spatial	dct:Location	Spatial coverage of the dataset	optional [0..n]	"optional"
--------------------	--------------	---------------------------------	----------------------------------	------------

Das Element "dct:spatial" ermöglicht die geografische Zuordnung des Datensatzes. Als Werte für das Element kommen vorwiegend politische Begriffe wie "CH" oder Kantons-, Gemeinde- oder verwandte Namen in Frage. Es sind aber auch die Angabe der Koordinaten eines Punktes oder die Definition einer rechteckigen "Bounding-Box" zulässig.

Das System erleichtert den Nutzer/innen die Nennung der Begriffe, indem es die Begriffsliste bei Bedarf automatisch erweitert (z.B. mit der Eingabe der Kantonsbezeichnung automatisch die Gemeindennamen hinzufügt).

Der Datentyp "dct:Location" lässt keine Beschreibung der geografischen Abdeckung in Form eines Polygons zu; deshalb soll dafür das Element "dct:Coverage" verwendet werden.

dct:coverage	dct:Location-PeriodOr-Jurisdiction	The Coverage element describes the spatial (and temporal) characteristics of the object or resource and is the key element for supporting spatial or temporal range searching on document-like objects that are spatially referenced or time-referenced.	optional [0..n]	im Europäischen Profil nicht vorgesehen
---------------------	------------------------------------	--	----------------------------------	---

Das Element "dct:coverage" soll nur dann verwendet werden, wenn die geografische Abdeckung in Form eines Polygons beschrieben werden muss (z.B. name="dc.coverage.polygon" scheme="WGS84.DD" content="9.547 46.737 9.550 46.739 9.556 46.740 9.554 46.743 9.548 46.742 9.547 46.737"; name="dc.coverage.placeName content="Heidsee").

In diesem Fall muss zur Wahrung der Kompatibilität mit dem DCAT Application Profile for data portals in Europe ein Element "dct:spatial" mit einer passenden "Bounding Box" eingefügt werden. Dies wird vom System automatisch erledigt.

dct:temporal	dct:PeriodOfTime (an interval of time that is named or defined by its start and end dates)	The temporal period that the dataset covers.	optional [0..n]	"optional" mit der Kardinalität 0..n
---------------------	--	--	----------------------------------	--------------------------------------

Für Zeitangaben wird zur Wahrung der Kompatibilität mit dem DCAT Application Profile for data portals in Europe immer das Element "dct:temporal" verwendet.

dct:accrualPeriodicity	dct:Frequency	The frequency at which the dataset is published	optional [0..1]	"optional"
Dieses Element kann verwendet werden, um eine allfällige Periodizität der Aktualisierung des Datensatzes anzugeben. Der Wert, den das Element "dct:accrualPeriodicity" annehmen kann, muss dem entsprechenden kontrollierten Vokabular entnommen sein (z.B. "halbmonatlich").				

dct:identifier	rdfs:Literal	a unique identifier of the dataset (the identifier might be used as part of the URI of the dataset, but still having it represented explicitly is useful)	mandatory [1..1]	"optional"
<p>Die eindeutige Kennzeichnung jedes Dataset-Metadaten-Records ist Voraussetzung für die Verlinkung von Dataset-Metadaten-Records, die verwandte Datensätze beschreiben.</p> <p>Der Datenlieferant kann im Element "dct:identifier" eine eigene Kennzeichnung liefern, um die Verwaltung der Datensätze auf dem OGD-Portal Schweiz und seinem eigenen Portal zu vereinfachen. Damit die Kennzeichnung im OGD-Portal Schweiz tatsächlich eindeutig ist, ergänzt die Portal-Software die Kennzeichnung mit einem Zusatz nach folgenden Regeln:</p> <ul style="list-style-type: none"> Falls der Datenlieferant eine Kennzeichnung liefert, ergänzt die Portal-Software diese Kennzeichnung mit dem Organisationsteil der (ersten) Kontakt-E-Mail-Adresse. (Z.B. wird die Kennzeichnung "BBT_025.37.1" zu "BBT_025.37.1@bbt.admin.ch".) Falls der Datenlieferant die Ergänzung bereits mitliefert (gekennzeichnet durch das Zeichen "@"), wird die Kennzeichnung unverändert übernommen. Falls der Dataset-Metadaten-Record eines Datenlieferanten kein Element "dct:identifier" enthält, fügt die Portal-Software das Element mit einer eindeutigen Kennzeichnung, die aus einer fortlaufenden Nummer und dem Organisationsteil der (ersten) Kontakt-E-Mail-Adresse besteht, ein. (Z.B. "7089@bbt.admin.ch") <p>Die Verlinkung der Dataset-Metadaten-Records (über das Element "rdfs:seeAlso") erfolgt jeweils durch die Angabe des entsprechenden Elements "dct:identifier".</p> <p>Das Ersetzen und Löschen von Dataset-Metadaten-Records (und den zugehörigen Distribution-Metadaten-Records) erfordert weitere Automatismen von Seiten der Portal-Software:</p> <ul style="list-style-type: none"> Ersetzt ein Dateneigner einen bestehenden Dataset-Metadaten-Record durch einen neuen mit gleicher Kennzeichnung, bleiben die Elemente "rdfs:seeAlso" der verlinkten Dataset-Metadaten-Records unverändert. Ersetzt ein Dateneigner einen bestehenden Dataset-Metadaten-Record durch einen neuen mit neuer Kennzeichnung, ändert die Portal-Software die Werte aller Elemente "rdfs:seeAlso" der verlinkten Dataset-Metadaten-Records entsprechend. Die Verlinkung bleibt somit bestehen. Löscht ein Dateneigner einen bestehenden Dataset-Metadaten-Record, entfernt die Portal-Software in allen verlinkten Dataset-Metadaten-Records die Elemente "rdfs:seeAlso". 				

rdfs:seeAlso	rdfs:Resource	links to a resource that might be interesting to users interested in the subject of this dataset	optional [0..n]	Im Europäischen Profil nicht vorgesehen
<p>Mit dem Element "rdfs:seeAlso" kann ein Dataset-Metadaten-Record mit einem anderen Dataset-Metadaten-Record verknüpft werden. Die Verknüpfungen ermöglichen den NutzerInnen, zu einem gefundenen Datensatz verwandte Datensätze zu entdecken.</p> <p>Die Elemente "rdfs:seeAlso" in einem Dataset-Metadaten-Record werden vom Dateneigner bei Bedarf eingefügt. Auf diese Weise können verwandte Datensätze (z.B. Bevölkerungszahlen von 2013 mit den Bevölkerungszahlen von 2014) einfach miteinander verknüpft werden.</p>				

Zur Darstellung wird systemintern die Form `<rdfs:seeAlso> <rdf:Description rdf:about="Verweis auf den betreffenden Dataset-Metadaten-Record (dct:identifizier)"> <rdfs:label>Titel des betreffenden Datensatzes</rdfs:label> </rdf:Description> </rdfs:seeAlso>` verwendet.

3. Elemente des Dataset-Metadaten-Record, die im europäischen Profil vorgesehen sind, im Schweizer Profil jedoch nicht verwendet werden sollen.

adms:version adms:versionNotes adms:identifier dct:conformsTo				Im Europäischen Profil „optional“, in der DCAT-Spezifikation aber nicht enthalten.
--	--	--	--	--