

Anbindung der GDI-DE an GovData



Diese Seite erläutert die **Weiternutzung der Metadaten der GDI-DE durch GovData**. Dies betrifft insbesondere die Abläufe der Metadatenabgabe, das Mapping der zu übertragenden, relevanten Informationen (die Zuordnung der ISO-Elemente zu DCAT-AP.de-Elementen) sowie die Zuordnung von Kategorien unter DCAT-AP.de aufgrund von in den ISO-Metadaten enthaltenen Informationen.

Seitenübersicht/ Schnellzugriff

- 1 [Hintergrund](#)
 - 1.1 [Zugriff auf die ISO-Metadaten der GDI-DE](#)
 - 1.2 [Anforderungen an die ISO-Metadaten der GDI-DE](#)
 - 1.3 [Mapping der relevanten Informationen](#)
 - 1.4 [Zuordnung von Kategorien](#)

Hintergrund

Der IT-Planungsrat hat DCAT-AP.de als formalen Austauschstandard für allgemeine offene Verwaltungsdaten festgelegt. Das bedeutet, dass die Informationen über Geo-Ressourcen, die unter Open Data fallen und mit entsprechenden Lizenz- bzw. Nutzungsbedingungen versehen sind, ebenfalls in Strukturen gem. DCAT-AP.de zu dokumentieren sind. Da die Geo-Ressourcen idealerweise bereits in der GDI-DE mittels Metadaten gem. ISO 19115/19119/19139 beschrieben sind, wurde zur Vermeidung von Doppelerfassungen eine Vorgehensweise etabliert, um die Metadaten aus der GDI-DE für Open Data-Portale unter DCAT-AP.de nutzbar zu machen und eine Auffindbarkeit in Open Data-Portalen automatisch zu ermöglichen.

Zugriff auf die ISO-Metadaten der GDI-DE

Die ISO-basierten Metadaten der GDI-DE werden über die Standard-CSW-Schnittstelle des Geodatenkatalog.de bereitgestellt. Die "Übersetzung" in [DCAT-AP.de](#), d.h. das Extrahieren der relevanten Informationen aus den Metadaten der GDI-DE erfolgt durch einen Ableitungsprozess auf Seiten von GovData auf Basis des u.g. abgestimmten Mappings, d.h. auf Seiten der GDI-DE erfolgt keine aktive Ableitung mehr (Anmerkung: bis 2019 wurde durch einen Prozess beim Betrieb GDI-DE der Inhalt des Geodatenkatalog.de in die Struktur OGD abgeleitet und anschließend an GovData übergeben. Dieses Verfahren wird nicht mehr angewendet.).

Anforderungen an die ISO-Metadaten der GDI-DE

Grundsätzlich sind die Vorgaben der GDI-DE einzuhalten, die im Dokument [Konventionen zu Metadaten](#) formuliert sind und auch die Belange von INSPIRE sowie Open Data berücksichtigen. Details (insbesondere zu den nachfolgend aufgeführten Punkten) sind dort nachzulesen. An dieser Stelle erfolgt lediglich eine Zusammenfassung aufgrund der besonderen Bedeutung für die Weiternutzung der Metadaten der GDI-DE für GovData:

- **Daten-Dienste-Kopplung:** Open Data-Portale sind „Daten-orientiert“, d.h. im Fokus steht die freie Daten-Ressource. Dienste (WMS, WFS, WCS, Atom-Feed etc.) sind Werkzeuge, um auf einen Datenbestand zuzugreifen. Dienste haben in Open Data-Portalen keine eigenen Metadatensätze, sondern werden als "Zugriffswege" der Dokumentation des Datenbestands zugeordnet. Entscheidend für die Gewinnung der Informationen zu den Diensten ist daher eine funktionierende Daten-Dienste-Kopplung in den Metadaten, da ausschließlich darüber die Dienste "erkannt" und dem zugehörigen Datenbestand zugeordnet werden können. Alternativ bzw. ergänzend werden durch GovData auch in den Daten-Metadaten etwaig enthaltene unmittelbare Download-Links zu statischen Datenmengen berücksichtigt.
- **Schlüsselwort "opendata":** Zur Kennzeichnung der relevanten Daten-Metadaten sehen die Metadatenkonventionen der GDI-DE vor, das Schlüsselwort „opendata“ zu verwenden. Dieses ist das steuernde Merkmal und Indiz für GovData, dass ein relevanter Datenbestand vorliegt, dessen Metadaten nach dem u.g. Mapping abzuleiten sind und dessen Zugriffswege über Dienste in den per Verfolgung der Daten-Dienste-Kopplung auffindbaren Dienst-Metadaten (s.o.) dokumentiert sind.
- **Nutzungsbedingungen:** Die zugrunde liegenden Lizenzbedingungen sind aus technischen Gründen in den Metadaten zusätzlich zu den Angaben in menschenlesbarem Freitext in einer bestimmten Syntax anzugeben (sog. JSON-Notation), um im Ableitungsprozess der Metadaten für Open Data-Portale entsprechend berücksichtigt werden zu können. Die JSON-Notation transportiert in Parameterform die notwendigen Informationen über die verwendete Lizenz (ID und Name), den Verweis auf den Lizenztext sowie den ggf. anzubringenden Quellenvermerk.

Beispiel

```
{
  "id": "dl-by-de/2.0",
  "name": "Datenlizenz Deutschland Namensnennung 2.0",
  "url": "https://www.govdata.de/dl-de/by-2-0",
  "quelle": "Quelle: Land NRW (<Jahr des Datenbezugs>)"
}
```

Hinweis!

Hierbei ist Folgendes zu beachten: Die in der JSON-Notation anzugebende Lizenz-ID ist gem. der Liste unter <https://www.dcat-ap.de/def/licenses/> (Spalte "Lizenzcode") auszuwählen. Gegenüber der bis 2019 unter OGD verwendeten Lizenz-ID aus <https://github.com/fraunhoferfokus/ogd-metadata/blob/master/lizenzen/deutschland.json> ergeben sich dabei i. d. R. Änderungen, die der folgenden [Tabelle](#) entnommen werden können (gelb markiert):

OGD (alt)	DCAT-AP.de (neu)
apache	apache
app_commercial	other-commercial
app_freeware	other-freeware
app_opensource	other-opensource
bsd-license	bsdlicense
cc-by	cc-by
cc-by-4.0	cc-by/4.0
cc-by-nc-4.0	cc-by-nc/4.0
cc-by-nd	cc-by-nd/3.0
cc-by-nd-4.0	cc-by-nd/4.0
cc-by-sa	cc-by-sa
cc-by-sa-4.0	cc-by-sa/4.0
cc-nc	cc-by-nc
cc-pdm-1.0	ccpdm/1.0
cc-zero	cc-zero
dl-de-by-1.0	dl-by-de/1.0

dl-de-by-2.0	dl-by-de/2.0
dl-de-by-nc-1.0	dl-by-nc-de/1.0
dl-de-zero-2.0	dl-zero-de/2.0
geonutzv-de-2013-03-19	geoNutz/20130319
gfdl	gfdl
gpl-3.0	gpl/3.0
mozilla	mozilla
odc-by	odby
odc-by	odcpddl
odc-odbl	odbl
official-work	officialWork
other-closed	other-closed
other-open	other-open

Mapping der relevanten Informationen

Die Zuordnung, welche Information aus einem ISO-Metadatensatz in ein bestimmtes Element unter [DCAT-AP.de](https://dcat-ap.de) überführt wird, ist in einem abgestimmten Mapping festgehalten, das 2017 in einem Pilotprojekt zwischen Open.NRW und dem GEOkatalog NRW erarbeitet und im Jahr 2021 gemeinsam mit der GKSt. GovData präzisiert wurde. Es dient gleichermaßen für die Abbildung der Inhalte des Geodatenkatalog.de für die Verwendung bei GovData.

Das Mapping ist als [Excel-Tabelle](#) dokumentiert.

Bei der Interpretation der ISO-Metadaten finden darüber hinaus folgende Grundsätze Anwendung:

1. aus Geo-Diensten werden Distributionen:
 - a. Geo-Dienste werden genauso wie Download-Links zu statischen Datenmengen aus den Daten-Metadaten behandelt und unter [DCAT-AP.de](https://dcat-ap.de) zu sog. „Distributionen“, d.h. ergänzend zur Beschreibung der Daten werden Zugriffsmöglichkeiten zugeordnet. Im GovData-Portal werden in der Folge die zugehörigen Links präsentiert.
 - b. Daten-Metadaten aus der GDI-DE, die weder Download-Links zu statischen Datenmengen aufweisen noch verknüpfte Dienst-Metadaten haben, sind mangels Bereitstellbarkeit von Distributionen sinnlos für eine Präsentation im GovData-Portal und werden daher in der Ableitung ignoriert.
2. Interpretation der Lizenzangaben:
 - a. die in den Daten-Metadaten zwingend zu dokumentierende Open Data-Lizenz gilt gemeinsam für die Daten selbst und alle zuzuordnenden Distributionen;
 - b. wenn die zugehörigen Dienst-Metadaten (Daten-Dienste-Kopplung, s.o.) eigene Lizenzinformationen beinhalten, werden diese Informationen als nur für die daraus generierte einzelne Distribution geltend berücksichtigt. Die Dokumentation der Open Data-Lizenz für die Daten sowie für alle Distributionen ohne eigene Aussagen bleiben gem. a) dadurch unberührt.

Zuordnung von Kategorien

Zur Gruppierung bzw. Filterung der Suchergebnisse in OpenData-Portalen wie z.B. GovData werden unter [DCAT-AP.de](https://dcat-ap.de) die sog. "[MDR Data Themes](#)" genutzt. Da die dort verwendeten Begriffe nicht 1:1 die in den ISO-Metadaten der GDI-DE übliche Kategorisierung (Element *topicCategory*) abbilden, liegt der Ableitung von ISO-Metadaten zu [DCAT-AP.de](https://dcat-ap.de)-Strukturen ebenso ein präzises Mapping der Kategorien zugrunde. Dadurch wird sichergestellt, dass jeder Kategorie der ISO-Metadaten mindestens eine Kategorie für Open Data zugeordnet wird. In vielen Fällen werden automatisch sogar mehrere MDR Data Themes ausgehend von einer einzigen Kategorie der ISO-Metadaten zugeordnet, um die Auffindbarkeit in OpenData-Portalen zu erhöhen. Auch dieses Mapping wurde im Pilotprojekt für Open.NRW erarbeitet und GovData zur Verfügung gestellt.

Die Zuordnung einer OpenData-Kategorie (MDR Data Theme) kann sich durch verschiedene Kriterien ergeben:

1. Durch das ISO-Element *topicCategory*: systematische Zuordnung von OpenData-Kategorien entsprechend der unten aufgeführten Tabelle. Da eine multiple Ausprägung des ISO-Elements *topicCategory* möglich ist, ergeben sich u.U. auch entsprechend viele OpenData-Kategorien.
2. Durch das Vorhandensein von Informationen bzgl. INSPIRE im ISO-Metadatensatz (INSPIRE-Thema als Schlüsselwort mit entsprechender Thesaurus-Angabe (GEMET)): Dadurch erfolgt eine Zuordnung zusätzlicher OpenData-Kategorien, die als fachlich sinnvoll bewertet worden sind. In vielen Fällen ergibt ein evtl. dokumentiertes INSPIRE-Thema jedoch „keine neuen Erkenntnisse“, d.h. es erfolgt keine weitere Zuordnung über diejenige hinaus, die sich aus dem ISO-Element *topicCategory* ohnehin ergibt.
3. Durch die direkte Ablage von OpenData-Kategorien (MDR Data Themes) in den Schlüsselwörtern im ISO-Metadatensatz (Element *keyword*): hierdurch besteht für den Metadatenerfasser die Möglichkeit, ergänzend zum systematischen Mapping aus Nr.1 und ggf. Nr.2 eine bewusste Zuordnung zu weiteren OpenData-Kategorien zu bewirken. Dazu muss das gewünschte MDR Data Theme als „4-Buchstaben-Code“ ohne Thesaurus-Angabe erfasst werden.

Die folgende [Tabelle](#) dokumentiert die Ableitung gem. der Nr.1 und 2:

Themenkategorie ISO		DCAT-AP.de (MDR Data Themes)				
MD_TopicCategoryCode	Name DE	feste Zuordnung		INSPIRE-Thema (sofern vorhanden) ...	... führt zusätzlich zu MDR Data Theme	
farming	Landwirtschaft	• E N VI • A G RI	• Umwelt • Landwirtschaft, Fischerei, Forstwirtschaft und Nahrungsmittel	---	---	---
biota	Biologie	• E N VI	• Umwelt	---	---	---
boundaries	Grenzen	• R E GI • G O VE	• Regionen und Städte • Regierung und öffentlicher Sektor	---	---	---
climatologyMeteorologyAtmosphere	Atmosphäre	• E N VI • TE CH	• Umwelt • Wissenschaft und Technologie	---	---	---
economy	Wirtschaft	• E C ON	• Wirtschaft und Finanzen	Energiequellen	• E N ER • E N VI • TE CH	• Energie • Umwelt • Wissenschaft und Technologie

				Mineralische Bodenschätze	• E • N • VI • TE • CH	• Umwelt • Wissenschaft und Technologie
elevation	Höhenangaben	• E • N • VI • G • O • VE • TE • CH	• Umwelt • Regierung und öffentlicher Sektor • Wissenschaft und Technologie	---	---	---
environment	Umwelt	• E • N • VI	• Umwelt	---	---	---
geoscientificInformation	Geowissenschaften	• R • E • GI • E • N • VI • TE • CH	• Regionen und Städte • Umwelt • Wissenschaft und Technologie	---	---	---
health	Gesundheitswesen	• H • E • AL	• Gesundheit	---	---	---
imageryBaseMapsEarthCover	Oberflächenbeschreibung	• E • N • VI • G • O • VE • TE • CH • R • E • GI • A • G • RI	• Umwelt • Regierung und öffentlicher Sektor • Wissenschaft und Technologie • Regionen und Städte • Landwirtschaft, Fischerei, Forstwirtschaft und Nahrungsmittel	---	---	---
intelligenceMilitary	Militär und Aufklärung	• JU • ST	• Justiz, Rechtssystem und öffentliche Sicherheit	---	---	---

inlandWaters	Binnengewässer	• E • N • VI • T • R • AN • A • G • RI	• Umwelt • Verkehr • Landwirtschaft, Fischerei, Forstwirtschaft und Nahrungsmittel	---	---	---
location	Ortsangaben	• R • E • GI • G • O • VE	• Regionen und Städte • Regierung und öffentlicher Sektor	---	---	---
oceans	Meere	• E • N • VI • T • R • AN • A • G • RI	• Umwelt • Verkehr • Landwirtschaft, Fischerei, Forstwirtschaft und Nahrungsmittel	---	---	---
				---	---	---
planningCadastre	Planungsunterlagen, Kataster	• R • E • GI • G • O • VE	• Regionen und Städte • Regierung und öffentlicher Sektor	Bewirtschaftungsgebiete / Schutzgebiete / geregelte Gebiete und Berichterstattungseinheiten	---	---
				Bodennutzung	---	---
				Flurstücke/Grundstücke	• JU • ST	• Justiz, Rechtssystem und öffentliche Sicherheit
society	Gesellschaft	• S • O • CI • E • D • UC	• Bevölkerung und Gesellschaft • Bildung, Kultur und Sport	---	---	---
structure	Bauwerke	• R • E • GI • T • R • AN	• Regionen und Städte • Verkehr	Gebäude	---	---
				Produktions- und Industrieanlagen	• E • C • ON	• Wirtschaft und Finanzen

				Umweltüberwachung	• E N VI	• Umwelt
transportation	Verkehrswesen	• T R AN	• Verkehr	---	---	---
utilitiesCommunication	Ver- und Entsorgung, Kommunikation	• E N ER • E N VI • G O VE	• Energie • Umwelt • Regierung und öffentlicher Sektor	---	---	---