



Leitfaden – Metadaten –





Bundesverwaltungsamt



KOMPETENZZENTRUM
OPEN DATA

Leitfaden Metadaten

Inhalt

Einleitung.....	3
Was sind Metadaten?.....	3
Mit Metadaten Daten auffindbar machen.....	3
Der Metadatenstandard	4
4. Metadatenfelder im Webformular auf GovData.de.....	4
Auf einen Blick – was sollten Sie beachten	12
Kontakt.....	14

Einleitung

Metadaten sind der Schlüssel für die schnelle und einfache Auffindbarkeit Ihrer Daten. Doch was verbirgt sich genau hinter diesem Begriff? Warum sind Metadaten so wichtig und wie sorgt man für eine gute Datenqualität? Zu diesen Fragen liefert dieser Leitfaden eine praktische Hilfestellung.

Was sind Metadaten?

Metadaten sind strukturierte Daten, die Informationen über andere Informationsressourcen enthalten. Ein Merkmal von Metadaten ist, dass sie maschinell lesbar und auswertbar sind. Die beschreibenden Metadaten liefern die nötigen Informationen, um den Inhalt des Datensatzes darzustellen, ohne den Datensatz selbst öffnen zu müssen. Metadatenstandards vereinheitlichen die auszufüllenden Felder, um ein einheitliches Qualitätsniveau zu erreichen.

So sehen beispielhaft die Metadaten bei einem Buch aus:

- Autor
- Titel
- Verlag
- Erscheinungsjahr
- Anzahl der Seiten
- ISBN Nummer
- Beschreibung oder Klappentext

Ein Buch ohne diese Angaben wäre im Register einer Bibliothek kaum auffindbar. Genauso verhält es sich mit Daten in einem Datenkatalog. Werden diese nicht entsprechend beschrieben, wird ein interessierter Nutzer sie nicht auffinden können.

Darum sind Metadaten der Schlüssel, um eine möglichst breite Nachnutzung zu ermöglichen. Sie enthalten die wichtigsten Informationen über den jeweiligen Datensatz sowie einen eindeutigen Verweis auf die Ressource (i.d.R. als Weblink), unter der die Daten zu finden sind.

Mit Metadaten Daten auffindbar machen

Die Metadaten von offenen Verwaltungsdaten in Deutschland sind auf dem nationalen Metadatenportal GovData katalogisiert. Damit Ihre Daten auf GovData.de gefunden werden können, bedarf es beschreibender Metadaten.

Wichtig:

Auf GovData.de werden ausschließlich Metadaten erfasst. Die Daten selbst verbleiben bei Ihnen. Ein Link in den Metadaten verweist auf die Ressource, den eigentlichen Speicherort der Daten.

Der Metadatenstandard

Um Metadaten strukturiert vorzuhalten und vergleichbar zu machen, werden durch Metadatenstandards die auszufüllenden Felder vereinheitlicht. Neben obligatorischen Feldern, wie Angaben zum Datenbereiter und dem Titel des Datensatzes, sollten empfohlene Felder ebenso befüllt werden. **Je genauer ein Datensatz beschrieben ist, desto besser ist dessen Auffindbarkeit und potentielle Weiternutzung.**

Es sollte also Ziel sein, so viele Informationen wie möglich in Form von Metadaten bereitzustellen.

Der deutsche Metadatenstandard „DCAT-AP.de“ ist eine deutsche Ableitung des europäischen Standards DCAT-AP. Eine Kompatibilität zum [Open-Data-Portal der Europäischen Union](#) ist vollständig gewährleistet.

Metadatenfelder im Webformular auf GovData.de

Die manuelle Eingabe der Metadaten erfolgt über das Webformular auf GovData.de.

Einige der dortigen Datenfelder beziehen sich auf „verpflichtend“ zu machende Angaben (Muss), diese sind mittels Sternchen markiert. Darüber hinaus besteht die Möglichkeit, weitere Informationen zum Datensatz mittels weiterer Datenfelder einzupflegen – diese entsprechen den „empfohlenen“ (Soll) und „optionalen“ (Kann)-Feldern des [DCAT-AP.de-Konventionenhandbuchs](#).

Neben obligatorischen Feldern, wie Angaben zum Datenbereiter und Titel des Datensatzes, sollten empfohlene Felder ebenso befüllt werden, um die Daten bestmöglich auffindbar zu machen und das volle Potenzial zu schöpfen. Wie bereits erwähnt, sollten so viele Informationen, wie möglich bereitgestellt werden. Je mehr Datenfelder Sie ausfüllen, desto höher die Qualität ihrer Metadaten und damit die Auffindbarkeit/Nutzbarkeit ihres Datensatzes. Angaben zur zeitlichen und räumlichen Abdeckung der Daten helfen Nutzenden beispielsweise ungemein bei der Suche und Auswahl der für sie relevanten Datensätze.

Zudem ist bei Teilen der empfohlenen Felder in absehbarer Zeit damit zu rechnen, dass diese im Rahmen der Weiterentwicklung von DCAT-AP.de verpflichtend werden. Sie daher heute bereits zu füllen bedeutet, weniger Nacharbeit zukünftig und optimale Voraussetzungen zur Auffindbarkeit der Daten.

Nachfolgend werden die einzelnen Felder genauer erläutert.

In der einführenden Sektion geht es um grundlegende Informationen zum Datensatz und zum Datenbereitsteller:

The screenshot shows a web form titled "Metadatenatz hinzufügen" (Add metadata set) from the GOVDATA portal. The form contains several fields:

- Datenbereitsteller ***: A dropdown menu with "Sonstige" selected.
- Ausgeblendet**: A checkbox labeled "Datensatz ist ausgeblendet und nur für Mitglieder der Organisation sichtbar" (unchecked).
- Titel ***: A text input field with the placeholder "Dieses Feld ist erforderlich".
- Beschreibung ***: A large text area with the placeholder "Dieses Feld ist erforderlich".
- GovData-ContributorID ***: A dropdown menu with "http://dcat-ap.de/def/contributors/govData" selected.

The GOVDATA logo and tagline "Das Datenportal für Deutschland" are visible in the top right corner.

Datenbereitsteller: Dieses Feld ist auf Grundlage Ihrer Anmeldung bei GovData vorausgefüllt und somit nicht änderbar.

Titel: Um was geht es genau in Ihrem Datensatz? Der Titel beschreibt den Datensatz, oder die Datei kurz und prägnant. Tragen Sie bitte einen aussagekräftigen Titel ein. Zum Beispiel nicht: „Straßennamen von K...“, sondern „Übersicht der gesamten Straßennamen der Stadt K..., 1950-1990, alphabetisch sortiert, Stand 05.2020“.

Beschreibung: Hier sind so viele Hinweise über den Inhalt und die Datei wie möglich zu machen – die Beschreibung ist die “Anleitung” zur Nutzung des Datensatzes. Die Beschreibung sollte bereits detailliert erklären, was in der Datei zu finden ist, ohne sie öffnen zu müssen. Wünschenswert sind vor allem Angaben zum Aufbau des Datensatzes, den enthaltenen Spalten, der Formatierung u.ä. ([hier](#) ein Beispiel). Handelt es sich um eine sehr umfangreiche Datei oder gibt es hinsichtlich der Aussagekraft der Daten etwas zu beachten, benennen Sie diesen Umstand.

Contributor-ID: Dieses Feld ist auf Grundlage Ihrer Anmeldung bei GovData vorausgefüllt und somit nicht änderbar.

In der folgenden Sektion geht es um genauere Angaben zur institutionellen und evtl. persönlichen Verantwortlichkeit bezüglich des Inhalts und der Bereitstellung des Datensatzes:

Kontakte	Autor		
	Name	<input type="text"/>	
	Email	<input type="text"/>	
	Webseite	<input type="text"/>	
	Adresse	Adressat	<input type="text"/>
		Zusatz	<input type="text"/>
		Straße	<input type="text"/>
		Stadt	<input type="text"/>
		PLZ	<input type="text"/>
		Land	<input type="text"/>

Autor:

Wenn es eine spezifische Person oder Stelle gibt, die die Daten erstellt hat, ist diese hier anzugeben.

	Ansprechpartner		
	Name	<input type="text"/>	
	Email	<input type="text"/>	
	Webseite	<input type="text"/>	
	Adresse	Adressat	<input type="text"/>
		Zusatz	<input type="text"/>
		Straße	<input type="text"/>
		Stadt	<input type="text"/>
		PLZ	<input type="text"/>
		Land	<input type="text"/>

Ansprechpartner:

Hier können Sie Angaben zum Ansprechpartner oder zum Fachbereich machen, an den sich ein Interessent für fachliche Rückfragen wenden kann (Amt + Referat). Bitte geben Sie nicht Ihre persönliche, sondern eine dauerhafte Gruppen- bzw. Funktions-Email-Adresse an. Zudem kann die Webseite Ihrer Behörde benannt werden – als Beispiel: <http://www.bva.bund.de>

Veröffentlichende Stelle		
Name	<input type="text"/>	
Email	<input type="text"/>	
Webseite	<input type="text"/>	
Adresse	Adressat	<input type="text"/>
	Zusatz	<input type="text"/>
	Straße	<input type="text"/>
	Stadt	<input type="text"/>
	PLZ	<input type="text"/>
	Land	<input type="text"/>

Veröffentlichende Stelle: Tragen Sie bitte hier die Behörde, Abteilung, Referat oder allgemein die Stelle ein, die die Bereitstellung des Datensatzes verantwortet.

Sollte in Ihrem Haus an einer zentralen Stelle die Metadatenpflege bzw. die Veröffentlichung von Daten erfolgen, z.B. über eine eigenständige Open Data-Stelle, so könnte hier dieser Bereich als Kontakt genannt werden. Handelt es sich um den gleichen Bereich, der auch als „Ansprechpartner“ oder „Verfasser“ dient, so können Sie die Angaben übertragen.

Für die zugehörige Email-Adresse wählen Sie bitte eine dauerhafte Gruppen- bzw. Funktions-Email-Adresse.

Urheber		
Name	<input type="text"/>	
Email	<input type="text"/>	
Webseite	<input type="text"/>	
Adresse	Adressat	<input type="text"/>
	Zusatz	<input type="text"/>
	Straße	<input type="text"/>
	Stadt	<input type="text"/>
	PLZ	<input type="text"/>
	Land	<input type="text"/>

Zeilen mit leerem "Name"-Feld werden nicht gespeichert.

Urheber: Sollte Ihre Behörde nicht der Urheber der Daten sein bzw. möchten Sie den Urheber explizit benennen, so ist dieses Feld dafür vorgesehen.

In der folgenden Sektion geht es um eine genauere inhaltliche Kontextualisierung der bereitgestellten Daten:

The screenshot shows a form with the following sections:

- Kategorien:** A grid of 14 checkboxes for categories: Bevölkerung und Gesellschaft, Energie, Internationale Themen, Landwirtschaft, Fischerei, Forstwirtschaft und Nahrungsmittel, Regionen und Städte, Verkehr, Wissenschaft und Technologie, Bildung, Kultur und Sport, Gesundheit, Justiz, Rechtssystem und öffentliche Sicherheit, Regierung und öffentlicher Sektor, Umwelt, and Wirtschaft und Finanzen.
- Schlagwörter:** A text input field with a note: "Ein oder mehrere Werte erlaubt, mehrere Werte mit Komma getrennt angeben: Wert1, Wert2, Wert3".
- Webseite:** A text input field.
- Qualitätssicherungsprozess URI:** A text input field.
- Kalender-Daten:** Two date pickers labeled "Veröffentlicht: Datum wählen" and "Aktualisiert: Datum wählen".

Kategorie: Wählen Sie hier bitte eine oder mehrere Kategorien aus, die dem Inhalt Ihres Datensatzes am nächsten kommen. Handelt es sich bspw. um Daten der Gesundheitsfürsorge in Deutschland wählen Sie bitte die Kategorie „Gesundheit“. Datensätze können auch zwei oder mehr Kategorien zugeordnet werden, z.B.: Daten zur Ladesäuleninfrastruktur sowohl der Kategorie „Energie“ wie auch der Kategorie „Verkehr“. Die Kategorie „Regierung und öffentlicher Sektor“ bietet sich für Verwaltungsdaten an, die nicht inhaltlich präziser unter einer anderen Kategorie zu fassen sind.

Schlagwörter: Hier sollten Schlagworte, Schlüsselbegriffe oder synonyme Begriffe eingepflegt werden.

Eine gewissenhafte Verschlagwortung ermöglicht es Nutzern, Daten schneller und gezielter zu finden. Nutzen Sie Schlagwörter, um den Inhalt des bereitgestellten Datensatzes über die Einordnung als „Kategorie“ hinaus möglichst genau zu beschreiben. Dabei ist es empfehlenswert, auch den Titel der Ressource (bzw. die charakteristischen Bestandteile daraus) zusätzlich als Schlüsselwort zu erfassen.

Synonyme sollten ebenfalls als Schlagwort hinterlegt sein, um die bereitgestellten Daten einem möglichst großen Publikum zugänglich zu machen. Für „Kindertagesstätten“ wären beispielsweise als zusätzliche Schlagworte denkbar: „Kita“, „Hort“, „Kinderladen“, „Kinderhort“ u.ä.

Ein Blick in bereits bestehende Datensätze auf GovData unter einem gewählten Schlagwort kann für die eigene Verschlagwortung Anregungen für weitere Schlagworte bieten, auch der [Musterdatenkatalog](#) kann Anhaltspunkte liefern.

Webseite: Wenn Sie eine Webseite anbieten können, die weitere Informationen zu den Daten liefert, geben Sie diese bitte hier an (bei Bevölkerungsdaten z.B eine Website der entsprechenden Abteilung innerhalb eines statistischen Amtes).

Qualitätssicherungsprozess URI: Dieses Feld bitte nicht berücksichtigen.

Kalenderdaten: Bitte tragen Sie hier das Datum der ersten Veröffentlichung der Daten (nicht auf GovData) ein. Sollten Sie den Datensatz einmal aktualisieren, tragen sie dies bitte im rechten Feld ein. Dort sollte sich stets das Datum der letzten Aktualisierung finden.

In der folgenden Sektion werden technische Angaben zur Ressource selbst abgefragt:

The screenshot shows a form titled "Ressourcen" with a sub-section "Ressource 1". The form contains the following fields and options:

- Webseite ***: A text input field containing the placeholder text "Dieses Feld ist erforderlich".
- Name**: A text input field.
- Beschreibung**: A text input field.
- Format**: A text input field.
- Sprache**: A text input field. Below it, a note reads: "Ein oder mehrere Werte erlaubt, mehrere Werte mit Komma getrennt angeben: Wert1, Wert2, Wert3".
- Lizenz ***: A dropdown menu with the selected option "Datenlizenz Deutschland Namensnennung 2.0".
- Namensnennungstext für "By"-Clauses**: A text input field.
- Aktualisierungsdatum**: A date picker field with the text "Datum wählen" and a calendar icon.
- Verfügbarkeit**: A dropdown menu with the selected option "Keine Angabe".

At the bottom of the form is a button labeled "Ressource hinzufügen".

Webseite: An dieser Stelle ist die URL zur Website zu hinterlegen, über die die Ressource abgefragt werden kann.

Name / Beschreibung / Format / Sprache: Hier ist der Name der Datei anzugeben. Weitere, die Ressource beschreibende Informationen sowie Angaben zum Format (bspw. CSV) und zur Sprache des Datensatzes erhöhen die Nutzbarkeit.

Tipp: Stellen Sie Daten in verschiedenen Formaten zur Verfügung. Das Feld „Ressource“ ist mehrfach vorhanden. Diese Eigenschaft kann für parallele Sprachversionen wiederholt werden, falls die Daten der Datei in verschiedenen Sprachen bereitgestellt werden.

Lizenz: GovData empfiehlt die Verwendung der „Datenlizenz Deutschland Namensnennung 2.0“ (voreingestellt) oder der „Datenlizenz Deutschland – Zero – Version 2.0“ sowie der „Creative Commons Namensnennung – 4.0 International (CC BY 4.0)“. Soweit dies nicht möglich sein sollte, wird die Verwendung einer der übrigen zur „freien Nutzung“ ausgewiesenen Lizenzen in der jeweils neuesten Version empfohlen.

Grundsätzlich wird die Verwendung von offenen und standardisierten Lizenzen für die Bereitstellung von Daten als offene Daten empfohlen. Hierbei handelt es sich um Nutzungsbedingungen, die die rechtlichen Bedingungen, unter denen ein Werk zur Verfügung gestellt wird, regelt. Grundsätzlich begünstigt eine Lizenzierung nach international/national gängigen Standards eine einfache Weiterverwendung von Daten, insbesondere die Kombination von Datensätzen: Je offener die bereitgestellten Daten, desto besser die Nachnutzbarkeit. Die genannten Lizenzen mit Namensnennung sind sowohl in Deutschland als auch international verbreitet und entsprechen den [Kriterien der „Offenheit“ der Open Knowledge Foundation](#). Die Nutzung einer der beiden Lizenzen würde, durch zwingende Quellenangabe und Veränderungskennzeichnung, einer missbräuchlichen Veränderung der Daten vorbeugen.

Aktualisierungsdatum: Hier ist das Datum der letzten Aktualisierung der Datei auf Ihrem Server anzugeben: Wann ist die Datei bei Ihnen zuletzt bearbeitet worden?

Verfügbarkeit: Hier ist anzugeben, wie dauerhaft die Daten auf GovData bereitgestellt werden. Erklärungen zu den unterschiedlichen Angaben finden Sie im Auswahlménü selbst. Wünschenswert ist allgemein eine Bereitstellung von Daten als „stable“ (Daten bleiben langfristig erhalten).

Zum Abschluss besteht die Möglichkeit, bestimmte inhaltliche Aspekte des bereitgestellten Datensatzes zu präzisieren:

The screenshot shows a web form with the following sections:

- Zeitbezug:** Two date pickers labeled 'Von: Datum wählen' and 'Bis: Datum wählen'.
- Geopolitische Abdeckung URI:** A text input field with a note: 'Ein oder mehrere Werte erlaubt, mehrere Werte mit Komma getrennt angeben: Wert1, Wert2, Wert3'.
- Rechtsgrundlage für Zugangseröffnung:** A text input field with the same note as above.
- Ebene der geopolitischen Abdeckung URI:** A dropdown menu with 'Keine Angabe' selected.
- Abdeckung Text:** A text input field with the same note as above.
- Raumbezug:** A large empty text area. To its right, there is a note: 'Mit Geojison.io kann das aktuelle Polygon bearbeitet oder ein neues erzeugt werden. Bitte folgendes Schema einhalten:' followed by a JSON schema example:

```
{  "type": "Polygon",  "coordinates": [    [      [11.598079, 51.29537],      [11.598079, 53.89386],      [6.654667, 53.89386],      [6.654667, 51.29537],      [11.598079, 51.29537]    ]  ]}
```

Zeitbezug: Geben Sie hier bitte den Zeitraum an, für den die bereitgestellten Daten erhoben wurden: Über welchen Zeitraum treffen die bereitgestellten Daten eine Aussage?

Geopolitische Abdeckung URI: Hier können Sie mittels eines eindeutigen Codes die verwaltungsgeographische Abdeckung der bereitgestellten Daten festlegen. Listen der URIs für [Bundesländer](#), [Regierungsbezirke](#), [Landkreise](#) und [Kommunen](#) stehen hierfür zur Verfügung.

Wenn hier keine Angabe sinnvoll ist, bietet „Abdeckung Text“ die Möglichkeit der Freitext-Eingabe.

Rechtsgrundlage für Zugangseröffnung: Wenn die Veröffentlichung aufgrund einer gesonderten Rechtsgrundlage wie länderspezifischen Transparenz- oder Open-Data-Gesetzen erfolgt, die über die bundesgesetzlichen Regelungen hinausgehen, kann dies hier angegeben werden. Ansonsten bleibt das Feld frei.

Ebene der geopolitischen Abdeckung URI: Wählen Sie hier aus, ob sich die bereitgestellten Daten administrativ auf die Ebene „International“, „EU“, „Bund“, „Bundesland“, „Kreis“ oder „Kommune“ beziehen. Wenn Sie eine Ebene unterhalb der Ebene „Bund“ wählen, sollte eine Ergänzung in „Geopolitische Abdeckung URI“ (s.o) erfolgen.

Bei Bundesbehörden wird die Abdeckung in der Regel der Ebene „Bund“ entsprechen.

Abdeckung Text: Wenn der verwaltungsgographische Bezug nicht durch einen URI angegeben werden kann, bietet dieses Feld optional die Möglichkeit, die Abdeckung als Freitext anzugeben (z.B. bei komplexeren Bund-Länder-Kooperationen oder auf kommunaler Ebene).

Raumbezug: Dieses Feld bezieht sich auf rein geographische Bezüge, welche der Datensatz abdeckt.

Zur Hilfestellung ist ein Link zur Erzeugung eines GeoJSON-Schemas hinterlegt, mit dem Sie auf einer Landkarte den geographischen Bezugsrahmen einzeichnen und anschließend den erzeugten Code kopieren und in das Freifeld des Formulars einfügen können.

Tipp: Das mit dem Link erzeugte Schema muss nach Übernahme in das Eingabefeld des Formulars noch einmal gegen das rechts angezeigte Beispiel geprüft und der zu viel generierte Code (inkl. zu vieler Klammern am Ende) händisch entfernt werden.

Auf einen Blick – was sollten Sie beachten

- Verwenden Sie einen kurzen, prägnanten Titel für Ihre Daten: Erläuterungen und Beschreibungen sind dadurch schneller verständlich und der Datensatz kann inhaltlich schneller zugeordnet werden.
- Die Freitextfelder können zur Beschreibung der Beschaffenheit Ihrer Daten genutzt werden. Das können Erklärungen zum Inhalt, zur Spaltenaufteilung, Erhebungsmethode oder ähnliches sein.
- Beachten Sie, dass Formatierungen wie Zeilenumbrüche, Einrückungen und Spiegelstriche etc. bei der Weitergabe der Metadaten in angeschlossene Systeme verloren gehen können.
- Vermeiden Sie unnötige und selbst erdachte Abkürzungen:
 - Abkürzungen nach „hausinternen“ Regeln sind nach außen i. d. R. niemandem bekannt.
 - Folglich wird kaum jemand auf diese Weise danach suchen. Schlimmstenfalls wird man Ihre Ressource überhaupt nicht finden, wenn diese Abkürzung die einzige Benennung ist und eine langschriftliche Beschreibung fehlt.
- Ebenso sind regionale Dialekte nicht geeignet, um Datensätze zu beschreiben. Versuchen Sie Formulierungen in Hochdeutsch, einfach und kompakt zu halten.
- Verwenden Sie möglichst ausgeschriebene Begriffe, da Abkürzungen nicht immer eindeutig und für jeden verständlich sind.
- So sorgen Sie dafür, dass jeder – vom Laien bis zum Experten – Ihre Daten „entdecken“ kann.

- Suchen Sie nach allgemeinen Ausdrücken für Ihre Fachbegriffe:
- Natürlich kommen Fachthemen nicht ohne feste Spezialtermini aus. Aber durch den Einsatz von alltagssprachlichen Synonymen (zusätzlich zum Fachvokabular) können Sie erreichen, dass auch jemand, der fachfremd ist, Ihre Daten findet.
- Ihre Aussagen müssen „global“ aussagekräftig und brauchbar sein:
- Bedenken Sie, dass andere Institutionen ähnliche Ressourcen haben könnten und daher der räumliche Geltungsbereich eine sehr wichtige Rolle spielt. Auch textlich sollte daher die Aussage (z. B. in der Kurzbeschreibung) „...daten im Kreisgebiet“ vollständig sein und auch benennen, um welchen Kreis oder welche Gemeinde es sich genau handelt.

Metadaten sollten fachlich aktuell und stets inhaltlich richtig und zutreffend sein. Sie sollten regelmäßig überprüft und bei relevanten Änderungen aktualisiert werden.

Kontakt

Kompetenzzentrum Open Data (CCOD)

OpenData@bva.bund.de

www.bva.bund.de