



# Ein Metadaten-Schema für offene Daten, Dokumente und Applikationen

*Initialer Entwurf der Metadaten-Struktur für die Open Government Data Plattform  
Deutschland*

Evanela Lapi, Florian Marienfeld, Nikolay Tcholtchev,  
Jens Klessmann, Ina Schieferdecker\*

\* Fraunhofer FOKUS, Berlin, {[evanela.lapi](mailto:evanela.lapi@fokus.fraunhofer.de), [florian.marienfeld](mailto:florian.marienfeld@fokus.fraunhofer.de), [nikolay.tcholtchev](mailto:nikolay.tcholtchev@fokus.fraunhofer.de), [jens.klessmann](mailto:jens.klessmann@fokus.fraunhofer.de), [ina.schieferdecker](mailto:ina.schieferdecker@fokus.fraunhofer.de)}@fokus.fraunhofer.de

**Abstract:** *Offene Daten (auf englisch: Open Data) sind einer der wichtigsten IT-Trends der letzten Jahre, mit dem das Ziel verbunden wird, den freien Austausch und die Weiterverwendung von Daten und darauf aufbauender Mehrwert-Angebote zu erleichtern. Die bestehenden Informations- und Datenportale von Bund, Ländern und Kommunen in Deutschland sind eine gute Ausgangsbasis für den weiteren Ausbau des Angebots offener Daten in Deutschland. Politik, Verwaltung, Wirtschaft und Wissenschaft haben mit der »Dresdner Vereinbarung zum 5. Nationalen IT-Gipfel« im Dezember 2010 als einen nächsten Schritt bis 2013 den Aufbau einer zentral zugänglichen, den Interessen der NutzerInnen an einer einem einheitlichen, leichten und benutzerfreundlichen Zugriff gerecht werdenden Open-Government-Plattform mit dem Schwerpunkt auf der Bereitstellung offener Verwaltungsdaten vereinbart. Um die Aufbereitung der und den Zugang zu den Daten zu lenken und zu erleichtern, wird derzeit ein Metadaten-Schema erarbeitet, das in seinem initialen Entwurf hier vorgestellt wird.*

**Keywords:** Open Data, Open Government, Metadata, Metadata-Scheme, Categories

**Acknowledgement:** Der initiale Entwurf ist im Diskurs mit vielen Akteuren rund um Open Data in Deutschland entstanden. Wir danken dem Bundesministerium des Innern, der Bund-Länder Arbeitsgruppe Open Government, der Unterarbeitsgruppe Technik als auch den Vereinen Open Data Network, Open Knowledge Foundation Deutschland und Government 2.0 Netzwerk.

## 1. Einleitung

Bei der Bereitstellung von Open Government Data geht es um Daten der öffentlichen Verwaltung, die Dritten zur Weiterverwendung zur Verfügung gestellt werden. Verwaltungsdaten, die als offene Daten bereitgestellt werden, sollen von Dritten möglichst einfach weiterverwendet werden können. Von besonderer Bedeutung sind in diesem Zusammenhang die drei Kriterien der Maschineninterpretierbarkeit, der offenen Lizenzierung und der kostenfreie Nutzung. Ein Datensatz sollte vorrangig diese drei Merkmale erfüllen, um als offen eingestuft zu werden: Erst ein gewisser Grad an Maschineninterpretierbarkeit erlaubt die automatisierte Verarbeitung mit vergleichsweise geringem Aufwand. Sind die

Nutzungsbestimmungen nicht so formuliert, dass sie eine Weiterverwendung einfach verständlich erlauben, sind im Einzelfall aufwendige Verhandlungen zwischen DatenbereiterInnen und -nutzerInnen mit unklarem Ausgang notwendig. Ist die Weiterverwendung eines Datensatzes geldleistungspflichtig, so schränkt dies den Kreis der potenziellen NutzerInnen entsprechend ein, gesamtwirtschaftliche Effekte werden in geringerem Maße realisiert.

Vor diesem Hintergrund sollen mit der Open-Government-Plattform für Deutschland die Datenangebote öffentlicher Stellen mit den folgenden Zielen gebündelt werden:

- Aufbau eines deutschlandweit einheitlichen Zugangs zu offenen Verwaltungsdaten,
- umfangreichere Weiterverwendung offener Verwaltungsdaten durch Dritte,
- Förderung von Innovation, Transparenz und Beteiligung durch einheitlichen Datenzugang,
- verbesserte Außendarstellung teilnehmender Stellen durch erhöhte Sichtbarkeit,
- Unterstützung der Vereinheitlichung von Standards, Prozessen und Richtlinien zur Bereitstellung (offener) Verwaltungsdaten und
- Verbesserung verwaltungsinterner Abläufe.

Langfristig gesehen gilt es, die Angebote zu einer nachhaltigen nationalen Dateninfrastruktur zu verdichten, bei der Anforderungen des föderalen Systems berücksichtigt werden. Das Datenangebot ist fortwährend auszubauen, und DatenbereiterInnen und -nutzerInnen sind entsprechende Werkzeuge an die Hand zu geben.

Ein Kernelement dieser Plattform ist das Metadaten-Schema, entlang dessen die Datensätze, Dokumente und Applikationen beschrieben werden, zu denen die Plattform einen zentralen Zugang bietet. Das Metadaten-Schema ist ein Vorschlag für einen Standard und gemeinsame Richtlinien zur Beschreibung von Open Government Data in Deutschland. Bevor das Metadaten-Schema selber vorgestellt wird, wird auf Metadaten im allgemeinen und ihre Definition mittels eines Metadaten-Schemas eingegangen.

## 2. Metadaten und Metadaten-Schema

Metadaten sind Daten über Daten, im Falle einer Open-Government-Plattform für Deutschland sind es Beschreibungen der in der Plattform registrierten Datensätze, Dokumente und Applikationen. Insoweit ist die wichtigste Funktion der Metadaten, das effektive Auffinden der Datensätze, Dokumente und Applikationen zu ermöglichen.

Ein Standard für Metadaten legt deren Syntax und Semantik fest, d. h. die Auswahl, Struktur und Bedeutung der einzelnen Metadatenfelder. Die Wahl eines geeigneten Metadatenstandards wirkt sich langfristig auf die Funktionalität der Open-Government-Plattform für Deutschland und auf die Akzeptanz seitens der DatenbereiterInnen aus, die den Standard einhalten müssen. Daher müssen verschiedene Anforderungen und Ziele bei der Auswahl berücksichtigt werden:

- *Leichtgewichtigkeit:* Das Metadaten-Schema muss leichtgewichtig sein, d. h. es sollte nur wenige obligatorische oder schwierig zu erzeugende Felder umfassen. Außerdem sollten die Einträge mit einfachen und frei verfügbaren Werkzeugen zu erstellen sein und nicht nur mit Spezialsoftware, die u. U. nicht vorliegt oder erst aufwendig angepasst werden muss.
- *Universalität:* Das Metadaten-Schema muss domänenübergreifend nutzbar sein, da die Open-Government-Plattform für Deutschland einen zentralen Einstiegspunkt zu

verschiedenen Bereichen darstellen soll. Daher sind etablierte Standards spezieller Domänen ggfs. als Standard ungeeignet, jedoch muss das Metadaten-Schema kompatibel zu diesen gestaltet werden.

- Leichtgewichtige *Erweiterungen* sollten möglich und kontrollierbar sein. Da unterschiedlichste Domänen angebunden werden sollen, ist es wahrscheinlich, dass in regelmäßigen Abständen neue Felder oder Schlagwörter im Metadaten-Schema zu ergänzen sind. Dies sollte ohne eine langwierige Änderung des Standards möglich sein. Insbesondere wegen der föderalen Struktur Deutschlands ist es wichtig, dass Variationen bzw. Dialekte des Standards aufwandsarm genutzt werden können.
- Bereits existierende Open-Data-Plattformen im In- und Ausland sollten mit dem gewählten Standard *kompatibel* sein, um keine Insellösung zu erzeugen, sondern eine direkte Vernetzung und Föderation zu ermöglichen.

Es ist nicht Ziel der Open-Government-Plattform für Deutschland, bestehende Register wie Geodaten (GeoPortal.DE), Umweltdaten (PortalU) oder Statistikdaten (Destatis) vollständig oder teilweise nachzubilden oder gar zu ersetzen. Vielmehr geht es um ein zusätzliches, zentralisiertes Angebot, das Domänen- und Ebenen-übergreifend Informationen zu den verschiedenen Datenangeboten zusammenführt und das für die beteiligten DatenbereiterInnen mit möglichst wenig Aufwand erzeugbar sein sollte.

In Vorbereitung des Metadaten-Schemas wurden Best-Practices zu Open Data analysiert, wobei vor allem zwei Gruppen von Standards anzutreffen sind. Die erste Gruppe kommt aus dem Bereich der Geodaten und baut vor allem auf dem ISO (International Standards Organisation)-Standard 19115 auf. Dem gegenüber stehen die Portale, die als Datenkatalog das CKAN (Comprehensive Knowledge Archive Network) der Open Knowledge Foundation nutzen, das wiederum auf dem DCAT-Vokabular bzw. dem Dublin Core – einem Semantic-Web-Ansatz – basiert.

Die ISO 19115 »Geographic Information – Metadata« wurde 2003 verabschiedet und ist weit verbreitet. Prominente Vertreter in Deutschland sind naheliegenderweise die Betreiber von Geoportalen, vor allem Geoportal.DE, sowie PortalU und Destatis. Diese verwenden den Catalogue Service for the Web (CSW) – ein Schnittstellen-Standard, der auf ISO 19115 aufbaut. Das Informationsmodell von ISO 19115 sowie die darin festgelegten Begrifflichkeiten sind sehr umfangreich: Es wurden nahezu alle potenziell für Geodaten relevanten Metadatenfelder und -begriffe festgehalten. Zudem können standardkonforme Erweiterungen und Profile erstellt werden. Dem CSW liegt der ISO 19115 zugrunde, wobei nur eine geringe Anzahl von Metadatenfeldern des ISO-Standards verwendet werden (circa 40). Die im CSW enthaltenen Metadatenfelder sind konform mit den Vorgaben des Dublin Core. Unter anderem mit GeoNetwork und InGriDCATalog® steht der Verwaltung auch Datenkatalog-Software zur Verfügung, die diesen Standard umsetzt.

Wichtigster Vertreter der DCAT/CKAN-Gruppe ist das britische Portal data.gov.uk. Die für dieses Portal entwickelte Software CKAN ist im Open-Data-Umfeld weit verbreitet. Dementsprechend hat sich das dort eingesetzte Metadaten-Format als Quasistandard etabliert. Dieser ist nicht eindeutig festgelegt, sondern basiert lose auf den im DCAT-Vokabular festgelegten Einträgen, die das gut etablierte Dublin-Core-Vokabular verwenden. CKAN-Metadaten verfügen außerdem über einen als »Extras« bezeichneten Bereich, in dem beliebige weitere Schlüssel-Wert-Paare hinterlegt werden können. In der Praxis orientieren sich neue Portale an den bestehenden Portalen; nur in Ausnahmefällen wird vom Vorbild data.gov.uk abgewichen. Im Gegensatz zu ISO 19115 wurde bei DCAT/CKAN nicht versucht, alle erdenklichen

Einsatzmöglichkeiten vorab zu standardisieren. Vielmehr wird hier ein minimaler Satz an Feldern festgelegt, der gerade ausreicht, um Datensätze hinreichend beschreibbar und auffindbar zu machen. Wollen DatenbereitstellerInnen zusätzliche Metadaten angeben, können sie Informationen ohne aufwendige Absprachen hinzufügen und dezentral definieren, wie die Angabe zu verstehen ist, beispielsweise mit Hilfe des Semantic Webs bzw. mit Linked Data.

Es ist Aufgabe der Open Government Data-Plattform für Deutschland (OGPD), Metadaten beider Standards von Datenbereitstellern zu akzeptieren, insbesondere um ein Harvesting (einen regelmäßigen automatischen Import von Metadaten) zu ermöglichen. Für CKAN existiert eine entsprechende INSPIRE-Import-Erweiterung, mit der bspw. Geo-Metadaten nach data.gov.uk geharvestet werden.

Im Ergebnis der Analyse wurde für den OGPD-Prototypen empfohlen, den DCAT/CKAN-Metadaten-Standard zu nutzen. Dieser hat sich speziell in domänenübergreifenden und verwaltungsrechtlich heterogenen Szenarien bewährt und ist im Open-Data-Umfeld weit verbreitet. Mit Blick auf die vielen unterschiedlichen Datenbereitsteller, für die die Anpassung an ein starres, domänenspezifisches Schema sehr aufwendig wäre, eignet sich ein schlanker, flexibler Standard wie DCAT/CKAN besser.

### **3. Die initiale Version des Metadaten-Schemas**

Dieser Abschnitt beschreibt das Metadaten-Schemata für Datensätze, Dokumente und Apps. Dieses Schema erweitert das Kernschema von CKAN mit zusätzlichen Feldern, welche OGPD-Prototyp spezifisch sind. Um den Nutzen der CKAN-Features zu maximieren, wird ein gemeinsames Schema definiert, welches sowohl für Datensätze als auch für Dokumente und Apps verwendet werden kann.

Zur Definition des Metadaten-Schemas wurde die JSON-Schemasprache (JavaScript Object Notation) genutzt, obwohl diese noch kein offizieller IETF-Standard ist, jedoch insbesondere JSON-basierte Webdienste unterstützt, wie es bei OGPD mit den CKAN-basierten REST-Diensten (Representational State Transfer) der Fall sein wird. Zudem vereinfacht ein solches JSON-Schema eine Validierung der Metadaten.

Das allgemeine Metadaten-Schema für den OGPD-Prototypen ist ein selbstbeschreibendes Schema, welches anzeigt, wenn ein Feld nur für Datensätze, Dokumente, Apps oder alle genutzt werden kann. Das JSON-Schema Definition für das Metadaten-Schema für den OGPD-Prototypen kann bei den AutorInnen angefragt werden. Details über das Kernmetadaten-Schema von CKAN sind über die Referenzen nachvollziehbar.

Kurzgefasst enthält ein Metadatum des OGPD-Prototypen die folgenden Elemente: Die wichtigsten Eigenschaften, die den gesamten Datensatz, ein Dokument oder eine App betreffen, werden auf oberster Ebene abgelegt. Dazu gehören: Titel, Bezeichner, Beschreibung, Verantwortliche und Nutzungsbestimmungen. Der andere essentielle Bereich ist die Liste der Ressourcen, also bei Datensätzen die eigentlichen Datendateien (oder Datendienste). Wichtigste Eigenschaft einer Ressource ist wiederum deren URL. Außerdem können je Ressource u.a. Beschreibung und Format vermerkt werden. Dieser Aufbau ermöglicht es, inhaltliche zusammengehörende Dateien als einen Datensatz zusammenzufassen, beispielsweise für verschiedene Zeitabschnitte, Sprachen, Formate oder zugehörige Schlüsselwörter. Bei Dokumenten und Anwendungen verweist die Ressource auf das eigentliche Dokument bzw. auf die herunterladbare ausführbare Datei.

Im Detail enthält ein Metadatum des OGPD-Prototypen die folgenden Elemente: Die Identifikation eines Datensatzes, eines Dokuments oder einer Applikation erfolgt über einen

Identifizier, einen Titel und einen Namen:

- **id:** Der Identifizier kann entweder ein Universal Unique Identifier (UUID) oder eine Unique Resource Location (URL) sein. Jedes existiert jeweils nur einmal und kann somit Datensätze eindeutig identifizieren.
- **title:** Der Kurztitel beschreibt den Datensatz, das Dokument oder die App prägnant und wird z. B. in Suchergebnissen und Listen angezeigt.
- **name:** Mit Name ist ein für Menschen lesbarer Bezeichner des Datensatzes, des Dokuments oder der Apps gemeint, der für eine Identifizierung genutzt werden kann

Jeder Datensatz, jedes Dokument und jede Applikation kann individuell beschrieben werden:

- **notes:** Dieses Feld enthält eine Beschreibung und weitere Informationen zum Datensatz, zum Dokument oder zur App.

Datensätze, Dokumente und Applikationen werden in Gruppen (siehe Tabelle 1) und Untergruppen kategorisiert:

- **groups:** Hier werden die Kategorien für die Datensätze und Dokumente aufgelistet. Jede Kategorie erhält einen vordefinierten Identifizier gemäß nachstehender Tabelle. Die Kategorien sind statisch und werden vom Portalbetreiber langfristig gepflegt. Zusätzlich können mit subgroups fachspezifische Unterkategorien beispielsweise aus INSPIRE oder EVAS genutzt werden.
- **subgroups:** Hier werden die Unterkategorien für die Datensätze und Dokumente aufgelistet. Im Gegensatz zu den Oberkategorien werden diese genau den fachlichen Kategorien der Datenbereitsteller angeglichen, d.h. jeder neu angeschlossene Datenbereitsteller kann hier benötigte Unterkategorien hinzufügen.

*Tabelle 1: Mögliche Werte für das Feld Kategorie*

Name	Id
Wirtschaft	wirtschaft
Transport und Verkehr	transport
Umwelt und Klima	umwelt
Ver- und Entsorgung	verentsorgung
Geografie	geografie
Gesundheit	gesundheit
Verbraucherschutz	verbraucher
Öffentliche Sicherheit	sicherheit
Wohnen und Immobilien	wohnen
Bildung	bildung
Öffentliche Verwaltung, Haushalt und Steuern	verwaltung
Gesetze und Justiz	justiz
Demografie	demografie
Arbeitsmarkt	arbeit
Wahlen	wahl
Sozialleistungen	sozial
Kunst und Kultur	kultur
Sport und Erholung	erholung
Tourismus	tourismus
Jugend	jugend
Sonstiges	sonstiges
Politik	politik
Stadtplanung	stadtplanung
Geologie	geologie

Datensätze, Dokumente und Applikationen können verschlagwortet werden. Links auf Thesauri unterstützen auch spezialisierte Suchen:

- **tags:** Mit frei wählbaren Schlagwörtern können in diesem Feld Datensätze, Dokumente und Apps beschrieben werden, um diese leichter mit der Suche ausfindig machen zu können.
- **tag\_sources:** Datensätze, Dokumente und Apps können in diesem Feld durch URLs für weitere Schlagwortquellen beschrieben werden.
- **tag\_thesauri:** Die Thesauri – URLs für Datensätze, Dokumente und Apps finden hier Platz.

Je Datensatz, Dokument und Applikation wird zwischen SchöpferIn bzw. UrheberIn, Pflegebeauftragten und AnsprechpartnerInnen unterschieden:

- **author:** Der/Die SchöpferIn oder UrheberIn des Datensatzes, des Dokuments oder der App wird in diesem Feld als AutorIn beschrieben.
- **author\_address:** In dieses Feld wird die Adressinformation (E-Mail, Kontaktformular oder Adresse) der/s Autorin/s des Datensatzes, Dokuments oder der App zur erleichterten Kontaktaufnahme eingestellt.
- **maintainer:** Hier wird der/die aktuelle Pflegebeauftragte der Metadaten des Datensatzes, des Dokuments oder der App eingetragen.
- **maintainer\_address :** In diesem Feld ist die Adressinformation (E-Mail, Kontaktformular oder Adresse) des/r Pflegebeauftragte/n der Metadaten des Datensatzes, Dokuments oder Apps zu finden.
- **point\_of\_contact:** AnsprechpartnerIn zu Fragen zum Datensatz, zum Dokument oder zur App.
- **point\_of\_contact\_address:** In diesem Feld ist die Adressinformation (E-Mail, Kontaktformular oder Adresse) des/der Ansprechpartners/in des Datensatzes, Dokuments oder Apps zu finden.

Des weiteren wird die ursprüngliche Webadresse, die Zeitangaben als auch Links auf die Quell-Metadaten und das Quell-Portal geführt, so dass einfacher auf die ursprüngliche Quelle zugreifen zu können:

- **url:** Die ursprüngliche Webadresse des Datensatzes, des Dokuments oder der App wird hier gezeigt, um zu weiteren Informationen über den Datensatz oder die App zu gelangen.
- **dates:** Hier können Erstellungs-, Veröffentlichungs- und Aktualisierungsdaten von Daten und Metadaten angegeben werden.
- **metadata\_original:** Der ursprüngliche URL–Metadateneintrag der Datensätze oder der Dokumente kann hier eingetragen und nachverfolgt werden. Damit kann schnell die Originalposition von geharvesteten Datensätzen oder Dokumenten nachvollzogen und Dubletten erkannt werden.
- **metadata\_original\_source:** Dieses Feld ist für URL des Portals vorgesehen, von dem der Metadateneintrags des Datensatzes oder des Dokuments geharvestet wurde.

Zudem werden je Datensatz die geographische und zeitliche Abdeckung und Granularität geführt:

- **geographical\_coverage:** Die geographische Abdeckung eines Datensatzes kann hier eingetragen werden.
- **geographical\_granularity:** Ebenso kann die geographische Granularität des Datensatzes in diesem Feld erfasst werden.
- **temporal\_coverage\_from:** Der Zeitpunkt, von dem der Datensatz oder das Dokument einschließlich Daten enthält, wird hier eingetragen.
- **temporal\_coverage\_to:** Der Zeitpunkt, bis zu dem der Datensatz oder das Dokument einschließlich Daten enthält, wird hier eingetragen.
- **temporal\_granularity:** Weiterhin wird ein Feld zur Verfügung gestellt, um die zeitliche Auflösung der enthaltenen Daten zu beschreiben.

Für die Nutzungsbestimmungen eines Datensatzes, eines Dokuments oder einer Applikation wird die Kostenfreiheit festgestellt und allgemein bekannte (siehe Tabelle 1) bzw. proprietäre Lizenzen geführt:

- **is\_free\_of\_charge:** Die Information, ob der Datensatz/das Dokument/die App generell kostenfrei ist, ist hier einzutragen.
- **terms\_of\_use:** Die spezifischen Nutzungsbestimmungen des Datensatzes, des Dokuments oder der App können mit einem oder mehreren der folgenden drei Felder festgelegt werden.
  - **license\_id:** Lizenz aus einer festen Liste. Mögliche Werte sind nachstehender Tabelle zu entnehmen.
  - **license\_url:** Die URL, auf der die Lizenz des Datensatzes, Dokuments oder der App beschrieben und erklärt ist, ist in diesem Feld zu finden.
  - **other:** Freitext der die Nutzungsbestimmungen festlegt, falls diese keiner allgemein bekannten Lizenz entsprechen.

*Tabelle 2: Mögliche Werte für das Feld Lizenz*

license_id	Lizenzname
apache	Apache Software License
bsd-license	New and Simplified BSD Licenses
cc-by	Creative Commons Namensnennung
cc-by-sa	Creative Commons Weitergabe unter gleichen Bedingungen
cc-nc	Creative Commons Nicht-Kommerziell
cc-zero	Creative Commons Zero
gfdl	GNU Free Documentation License
gpl-3.0	GNU General Public License version 3.0 (GPLv3)
mit-license	MIT-Lizenz
mozilla	Mozilla Public License 1.0 (MPL)
mozilla1.1	Mozilla Public License 1.1 (MPL)
notspecified	keine Angabe
odc-odbl	OKD Compliant::Open Data Commons Open Database License (ODbL)

Abschließend wird definiert, ob das Metadatum zu einem Datensatz, einem Dokument bzw. einer Applikation gehört, aus welchem Sektor diese stammen und welche Datensätze ggfs. genutzt werden:

- **content\_type:** Dies ist ein technisches Feld, welches definiert ob sich der Metadateneintrag auf einen Datensatz, ein Dokument oder eine App bezieht.
- **used\_datasets:** Die verwendeten Datensätze bzw. Dokumente von Dokumenten und Apps werden mithilfe einer Liste mit URLs auf Metadaten erfasst.
- **sector:** Dieses Feld wird nur für Dokumente und Apps genutzt und zeigt an, ob eine App oder ein Dokument aus dem öffentlichen, dem privaten oder einem anderen Bereich kommt.

Der eigentliche Zugriff auf einen Datensatz, auf ein Dokument bzw. auf eine Applikation erfolgt über sogenannte Ressourcen, wobei zu jedem Datensatz, jedem Dokument bzw. jeder Applikation mehrere Ressourcen definiert sein können:

- **url:** Verweis auf die eigentliche Datendatei, das Dokument oder die Anwendung.
- **format:** Hier ist der Typ der Ressource anzugeben (für eine Datensatz bspw. csv, xls, etc.).
- **description:** Diese Beschreibung soll erläutern, welche Rolle die Ressource für den Datensatzes, das Dokument bzw. die App spielt, z. B. ob es sich dabei um eine bestimmte Zeitscheibe oder Schlüsselliste handelt.
- **language:** Sprache als ISO 639-1-Code, in der die Ressource verfasst ist.
- **hash:** Bei statischen Dateien die SHA2-Prüfsumme bzw. bei dynamischen Daten ein Zertifikat zur Überprüfung der Integrität

## 4. Ausblick

Die OGPD-Plattform wird dann zu einem gut nutzbaren Informationsportal werden, wenn die in der OGPD-Plattform bereitgestellten Metadaten einem einheitlichen und von allen AkteurInnen getragenen Metadaten-Schema folgen. Daher wird derzeit für den OGPD-Prototyp ein Metadaten-Schema auf Basis der Ausarbeitung in der Studie "Open Government Data Deutschland" entwickelt. Dieses ist entlang nationaler und internationaler Entwicklungen kontinuierlich zu aktualisieren. Zudem wird angestrebt, dass sich die OGPD-Verantwortlichen auf EU-Ebene engagieren, um ein zum deutschen Metadaten-Schema kompatibles europäisches Open-Government-Data-Metadaten-Schema mitzugestalten.

Ein Teil dieses Engagements sollte zudem darauf abzielen, ein öffentliches Linked Data-Vokabular zu erstellen und zu pflegen, so dass das Metadaten-Schema semantisch repräsentiert und die Metadaten einfacher weiterverarbeitet werden können. Ein solches Vokabular kann ebenso helfen, ein Metadaten-Schema als gemeinsamen Standard über Länder hinweg zu etablieren.

## Referenzen

Klessmann et al (2012) – Jens Klessmann, Philipp Denker, Ina Schieferdecker, Sönke E. Schulz: "Open Government Data Deutschland: Eine Studie zu Open Government in Deutschland im Auftrag des Bundesministerium des Innern",  
[http://www.bmi.bund.de/SharedDocs/Downloads/DE/Themen/OED\\_Verwaltung/ModerneVerwaltung/opengovernment.pdf](http://www.bmi.bund.de/SharedDocs/Downloads/DE/Themen/OED_Verwaltung/ModerneVerwaltung/opengovernment.pdf), zuletzt besucht am 11.09.2012



- Both et al (2012) – Wolfgang Both, Ina Schieferdecker : "Berliner Open Data-Strategie: Organisatorische, rechtliche und technische Aspekte ofener Daten in Berlin. Konzept, Pilotsystem und Handlungsempfehlungen",  
[http://www.berlin.de/projektzukunft/fileadmin/user\\_upload/pdf/sonstiges/Berliner\\_Open\\_Data-Strategie\\_2012.pdf](http://www.berlin.de/projektzukunft/fileadmin/user_upload/pdf/sonstiges/Berliner_Open_Data-Strategie_2012.pdf), zuletzt besucht am 11.09.2012
- INSPIRE (2012) – INSPIRE: Infrastructure for Spatial Information in the European Community,  
<http://inspire.jrc.ec.europa.eu/>, zuletzt besucht am 11.09.2012
- ISO (2012) - Übersetzung des ISO 19115, <http://www.gdi-de.org/thema2009/uebersetzungiso>, zuletzt besucht am 11.09.2012
- InGriDCATalog (2012) – Koordinierungsstelle PortalU:  
<http://www.kst.portalu.de/ingrid/ingriDCATalog.html>, zuletzt besucht am 11.09.2012
- GeoNetwork (2012) – Open-Source-Geo-Metadatenkatalog: <http://geonetwork-opensource.org>, zuletzt besucht am 11.09.2012
- CKAN (2012) – CKAN: Comprehensive Knowledge Archive Network: <http://ckan.org/>, zuletzt besucht am 11.09.2012
- CKAN Metadaten (2012) – CKAN Metadaten-Schema: <http://docs.ckan.org/en/ckan-1.7.1/domain-model.html>, zuletzt besucht am 11.09.2012
- Dublin Core (2012) – Dublin Core Metadaten: <http://dublincore.org/>, zuletzt besucht am 11.09.2012
- CKAN INSPIRE (2012) CKAN INSPIRE Extension: <https://github.com/okfn/ckanext-inspire>, zuletzt besucht am 11.09.2012
- JSON Schema (2012): JavaScript Object Notation: <http://json-schema.org/>, zuletzt besucht am 11.09.2012
- JSON Schema Spec (2012): JSON Draft Specification: <http://tools.ietf.org/html/draft-zyp-json-schema-03>, zuletzt besucht am 11.09.2012
- REST (2000): Roy Thomas Fielding: Architectural Styles and the Design of Network-based Software Architectures, PhD Thesis, Chapter 5, University of California, Irvine.
- DataGovUK (2012) – Open Data Portal von Großbritannien: <http://data.gov.uk/>, zuletzt besucht am 11.09.2012
- PortalU (2012) – Umweltportal Deutschland: <http://www.portalu.de/>, zuletzt besucht am 11.09.2012
- GeoPortal (2012) – Geoportal Deutschland: <http://www.geoportal.de/>, zuletzt besucht am 11.09.2012
- Destatis (2012) – Statistisches Bundesamt Deutschland: <https://www.destatis.de/>, zuletzt besucht am 11.09.2012

## **Ansprechpartnerin**

*Prof. Dr.-Ing. Ina Schieferdecker*

Ina Schieferdecker leitet das Kompetenzzentrums für Modellieren und Testen (MOTION) am Fraunhofer FOKUS, Berlin und ist Professorin an der Freien Universität Berlin. Sie beschäftigt sich mit Methoden und Werkzeugen für kosteneffiziente qualitativ-hochwertige Systeme, leitet das Innovationszentrum „IKT für smarte Städte“ bei FOKUS und koordiniert technisch die Open Data-Plattformen in Berlin und für Deutschland.