

Praxisleitfaden

# Open Data nachhaltig umsetzen

Einstieg in das Datenmanagement für  
die Bereitstellung offener Verwaltungs-  
daten aus Nordrhein-Westfalen

Version 1.0 | 2022





## Liebe Leserinnen und Leser,

Ich freue mich, dass Sie sich zum übergreifenden Thema Datenmanagement informieren möchten. Insbesondere um Verwaltungsdaten nachhaltig als Open Data bereitzustellen, ist es wichtig, sich systematisch mit den Daten aus dem eigenen Fachbereich und mit dem Umgang mit Daten im Allgemeinen auseinanderzusetzen.

Um die Open Data-Aktivitäten dauerhaft voranzutreiben, ist es notwendig, die mit der Bereitstellung verbundenen Aufgaben effizient umzusetzen. Nur so können Aufwände gering gehalten und hochwertige Datensätze veröffentlicht werden.

Dies kann durch die Entwicklung und Etablierung eines Datenmanagements gelingen. Wie eine praktische Umsetzung in den Behörden des Landes erfolgen kann und welche relevanten Schritte erforderlich sind, wird in der vorliegenden Veröffentlichung dargestellt.

Ein systematischer Umgang mit den eigenen Daten ist dabei nicht nur im Hinblick auf die Veröffentlichung von Open Data nach den Vorgaben des §16a EGovG NRW und der Open Data-Verordnung relevant, sondern bildet auch die Grundlage dafür, neuen Herausforderungen der Digitalisierung zu begegnen und die Chancen einer datenbasierten Verwaltung wahrzunehmen.

Geschäftsprozesse können beispielsweise schneller an neue Anforderungen angepasst werden, wenn Wissen über die vorhandenen Daten, ihre Qualität, Prozesse der Datenverarbeitung sowie Potenziale und Grenzen einer Weiterverwendung vorhanden sind. Nur so ist es möglich, Arbeitsabläufe in einer agilen Verwaltung effizient umzugestalten.

Zusammen bringen wir mit einem besseren Datenmanagement die Bereitstellung offener Verwaltungsdaten voran und legen wichtige Bausteine für eine stärker datenbasierte Landesverwaltung. Herzlichen Dank für Ihr Engagement!



**Ihr Prof. Dr. Andreas Meyer-Falcke**  
Der Beauftragte der Landesregierung  
Nordrhein-Westfalen für  
Informationstechnik – CIO





# Inhaltsverzeichnis

<b>Vorwort</b>	<b>3</b>
<b>Abbildungsverzeichnis</b>	<b>6</b>
<b>Abkürzungsverzeichnis</b>	<b>7</b>
<b>1. Ziel und Aufbau des Leitfadens</b>	<b>8</b>
<b>2. Datenmanagement</b>	<b>11</b>
2.1 Definition Datenmanagement .....	12
2.2 Bedeutung von Datenmanagement für die Bereitstellung offener Daten der Behörden des Landes Nordrhein-Westfalen ...	12
<b>3. Bausteine eines Datenmanagements</b>	<b>14</b>
3.1 Geeignete Datensätze identifizieren .....	16
3.2 Vorhandene Rechte klären .....	18
3.3 Datenformate auswählen .....	19
3.4 Offene Nutzungsbestimmungen .....	22
3.5 Metadaten definieren .....	23
3.6 Bestehende Datenberichtspflichten .....	26
3.7 Automatisierungspotenziale .....	28
3.8 Datensätze und Metadaten veröffentlichen .....	31
3.9 Qualitäts- und Wissensmanagement .....	33
3.9.1 Qualitätsmanagement .....	33
3.9.2 Wissensmanagement .....	35
<b>4. Erweiterung des Datenmanagement-Zyklus</b>	<b>36</b>
<b>5. Quellenverzeichnis</b>	<b>40</b>
<b>6. Anhang</b>	<b>42</b>
Glossar .....	42
<b>Impressum</b>	<b>44</b>

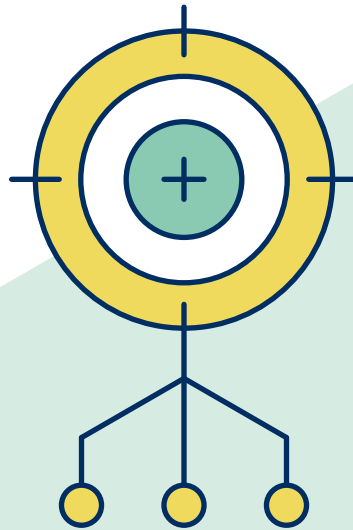
# Abbildungsverzeichnis

<b>Abbildung 1:</b>	Hauptfunktionen der offenen Daten .....	13
<b>Abbildung 2:</b>	Datenmanagement-Zyklus .....	15
<b>Abbildung 3:</b>	Daten- und Schnittstellenformate .....	19
<b>Abbildung 4:</b>	Beispiel zur Datenberichtspflicht des LANUV zur Luftqualität .....	27
<b>Abbildung 5:</b>	Ablauf des Ableitungsmechanismus der Metadaten am Beispiel der Stadt Köln .....	29
<b>Abbildung 6:</b>	Ablauf des Ableitungsmechanismus der Metadaten am Beispiel der Berichtspflichten des LANUV zur Luftqualität .....	30
<b>Abbildung 7:</b>	Prozess zur Veröffentlichung von Metadaten im Open.NRW-Portal .....	32
<b>Abbildung 8:</b>	Dashboard MQA-Komponenten des Open.NRW-Portals (Screenshot) .....	34
<b>Abbildung 9:</b>	Beispielhafter zu durchlaufender Datenmanagement-Zyklus, wenn keine Vorkenntnisse vorhanden sind .....	37
<b>Abbildung 10:</b>	Beispielhafte Erweiterung des Datenmanagement-Zyklus, wenn Vorkenntnisse vorhanden sind .....	39

## Abkürzungsverzeichnis

<b>API</b>	Application Programming Interface
<b>CSV</b>	Comma-separated Values
<b>CSW</b>	Catalogue Service for the Web
<b>DCAT-AP</b>	Data Catalogue Vocabulary-Application Profile for Data Portals in Europe
<b>DCAT-AP.de</b>	Data Catalogue Vocabulary-Application Profile for Data Portals in Deutschland
<b>EDP</b>	Europäisches Datenportal
<b>EDV</b>	Elektronische Datenverarbeitung
<b>EEA</b>	European Environment Agency
<b>EGovG</b>	E-Government-Gesetz (Deutschland)
<b>GDI-DE</b>	Geodateninfrastruktur Deutschland
<b>HTML</b>	Hypertext Markup Language
<b>INSPIRE</b>	Infrastructure for Spatial Information in Europe
<b>ISO</b>	Internationale Organisation für Normung
<b>JSON</b>	JavaScript Object Notation
<b>LANUV</b>	Landesamt für Natur, Umwelt und Verbraucherschutz
<b>MD</b>	Metadaten
<b>MQA</b>	Metadata Quality Assessment
<b>NRW</b>	Nordrhein-Westfalen
<b>OGC</b>	Open Geospatial Consortium
<b>PDF</b>	Portable Document Format
<b>RDF</b>	Resource Description Framework
<b>REST</b>	Representational State Transfer
<b>URL</b>	Uniform Resource Locator
<b>WFS</b>	Web Feature Service
<b>XML</b>	Extensible Markup Language

# 1. Ziel und Aufbau des Leitfadens





Akteurinnen und Akteure der öffentlichen Verwaltung erheben regelmäßig Daten oder lassen diese erheben. Das können beispielsweise Karten zur durchschnittlichen Lärmbelastung durch Straßenverkehr, Statistiken zur Bevölkerung oder Zahlen zu freien Platzkapazitäten in der Pflege sein. Das E-Government-Gesetz des Landes Nordrhein-Westfalen (EGovG NRW) verpflichtet mit § 16a die Behörden des Landes, solche Daten als Open Data bereitzustellen, sofern keine rechtlichen Hinderungsgründe wie etwa der Schutz personenbezogener Daten dem entgegenstehen.

Die Bereitstellung offener Verwaltungsdaten birgt zahlreiche Potenziale, beispielsweise eine Verbesserung der behördenübergreifenden Zusammenarbeit, die Stärkung des Vertrauens in Verwaltungshandeln, indem die Transparenz und Nachvollziehbarkeit erhöht wird, oder die Schaffung neuer, innovativer Anwendungen und Geschäftsmodelle.

Auch die datenbereitstellende Behörde kann von der Veröffentlichung von Open Data profitieren, bspw. dann, wenn die proaktive Veröffentlichung von Daten die Anzahl der Auskunftsanfragen verringert. Damit diese Potenziale zum Tragen kommen, ist es wichtig, dass der gesamte Prozess der Datenveröffentlichung so effizient wie möglich abläuft. Dies kann erreicht werden, indem ein entsprechendes Datenmanagement aufgebaut wird.

Der vorliegende Leitfaden soll die Behörden beim Aufbau eines Datenmanagements und bei der Bereitstellung der Verwaltungsdaten als Open Data unterstützen. Er beinhaltet die wesentlichen Schritte und erklärt wichtige Inhalte und Themen. Dabei werden die rechtlichen Rahmenbedingungen in Nordrhein-Westfalen, wie sie in §§ 16 und 16a EGovG NRW und in der Open Data-Verordnung <sup>[2]</sup> des Landes verankert sind, berücksichtigt. Der § 16 regelt hierbei, „wie“ Daten bereitgestellt werden sollen, und der § 16a legt fest, dass die Behör-

den des Landes dazu verpflichtet sind, bestimmte Daten als Open Data zu veröffentlichen. Diese und weitere rechtliche Grundlagen zur Bereitstellung offener Verwaltungsdaten sind in dem → **Open Data-Guide** <sup>[2]</sup> zusammengefasst und eingeordnet.

Der Leitfaden richtet sich an alle Mitarbeiterinnen und Mitarbeiter der Behörden des Landes Nordrhein-Westfalen, in deren Aufgabengebiet oder Zuständigkeitsbereich Daten erhoben oder verarbeitet werden oder die Daten durch Dritte erheben oder verarbeiten lassen. Keine Rolle spielt, mit welchem Vorwissen sie sich der Thematik „Offene Daten“ nähern, da das vorliegende Dokument als Blaupause für verschiedenste individuelle Rahmenbedingungen geeignet ist.<sup>1</sup> Bei Leserinnen und Lesern mit Vorwissen kann der Leitfaden dazu dienen, existierende Lücken oder Optimierungspotenziale zu identifizieren und anzugehen. Der Datenmanagement-Zyklus kann dann vielmehr als Bausteinlieferant angesehen werden, bei dem man sich an den notwendigen Stellen bedienen kann.

Einleitend werden die Ziele und der Aufbau des Leitfadens beschrieben und der Begriff „Datenmanagement“ definiert. Es wird erklärt, wie dieser im Kontext des Leitfadens und im Sinne der Veröffentlichung von offenen Daten der Behörden des Landes Nordrhein-Westfalen zu verstehen ist. Im Folgenden wird der sogenannte Datenmanagement-Zyklus vorgestellt, der die notwendigen Schritte auf dem Weg zu einer effizienten Veröffentlichung von offenen Daten beschreibt. Die Gliederung der nachfolgenden Kapitel orientiert sich an der Reihenfolge der Schritte, wie sie bei einem Datenmanagement-Zyklus vorzugsweise durchlaufen werden sollten: Dieser beginnt mit Handlungsempfehlungen zum Identifizieren von Datensätzen, die sich zur Veröffentlichung eignen, und der Klärung von damit verbundenen rechtlichen Fragestellungen. Danach wird die Auswahl geeigneter Daten- und Schnittstellenformate unter Berücksichtigung der Open Data-Verordnung beschrieben. Um Daten

---

<sup>1</sup> Bei der Entwicklung eines Datenmanagement-Konzepts kann es eine Herausforderung sein, die verschiedenen Wissensstände der Nutzerinnen und Nutzer zu berücksichtigen. Aufgrund der unterschiedlichen Wissensstände bzgl. des Umgangs mit (Meta-)Daten und deren Pflege können einzelne oder auch alle Bausteine des in diesem Dokument beschriebenen Datenmanagement-Zyklus bei einigen Akteurinnen und Akteuren schon vorhanden und bekannt sein.

als Open Data zu veröffentlichen, müssen zudem einige Regelungen beachtet werden. Im Kapitel über die offenen Nutzungsbestimmungen wird deshalb beschrieben, welche Bedingungen erfüllt werden sollten, damit ein Datensatz bereitgestellt werden darf und vielseitig weiterverwendet werden kann. Um die bereitgestellten Daten zudem auffindbar zu machen, spielen Metadaten im Veröffentlichungsprozess eine zentrale Bedeutung. Wie eine Metadatenbeschreibung aussehen kann, wird am Beispiel eines Datensatzes aufgezeigt. Bei der Veröffentlichung von Daten und dem Aufbau eines Datenmanagements ist es wichtig, bestehende Berichtspflichten zu beachten, über die eine Behörde verfügt. Das heißt bestimmte Informationen müssen zu einem bestimmten Zeitpunkt bereitgestellt bzw. „berichtet“ werden. Diese Berichtspflichten können z. B. bei Automatisierungsprozessen eine wichtige Grundlage bilden, wenn Schnittstellen geschaffen werden, um Daten direkt in das Open.NRW-Portal zu überführen. Dieser Prozess wird im Leitfaden anhand eines Beispiels aufgezeigt. Die Veröffentlichung der originären Daten und deren Metadaten erfolgt durch die datenbereitstellenden Behörden. Um Datensätze öffentlich zu machen, bedarf es einer Dateninfrastruktur (z. B. Fachportal, Homepage). Zudem müssen die Metadaten beschrieben und über das Open.NRW-Portal auffindbar gemacht werden. Es wird gezeigt, welche Kenntnisse benötigt werden und welche Daten und Metadaten wie und wo veröffentlicht werden können.



### Qualitätshinweis:

Qualitäts- und Wissensmanagement spielen in allen Schritten des Datenmanagement-Zyklus eine wichtige Rolle und stellen eine Querschnittsaufgabe dar. In vielen Kapiteln zeigt der Leitfaden diesbezüglich Hinweise auf. ■

**Wichtig ist: Das vorgestellte Datenmanagement sollte nicht einmalig, sondern im Sinne eines Datenmanagement-Zyklus nach seiner ersten Implementierung regelmäßig durchlaufen werden. Dabei kann und sollte der Zyklus je nach Reifegrad der jeweiligen Behörde bei der Umsetzung eines Open Data-Datenmanagements angepasst werden. Der Leitfaden zeigt abschließend beispielhaft Erweiterungen und Anpassungen des Datenmanagement-Zyklus auf.**

Das Dokument fügt sich in den von der Beratungsstelle Open Data beim Beauftragten der Landesregierung Nordrhein-Westfalen für Informationstechnik, CIO, zur Verfügung gestellten **→Open Data-Guide** <sup>[2]</sup> ein und verweist ggf. auf diesen. An den entsprechenden Stellen des vorliegenden Dokuments wird, auch um Doppelungen zu vermeiden, mit Links oder Literaturquellen (gesammelt am Ende im Quellenverzeichnis) auf die entsprechenden Fundstellen verwiesen, um bei Bedarf und/oder Interesse vertieft weiterlesen zu können.

In Verbindung mit den in diesem Dokument gegebenen Handlungsempfehlungen, verwendeten Beispielen und verlinkten Quellen bietet das Dokument ein umfangreiches Unterstützungsangebot, das den Leserinnen und Lesern dabei hilft, ihre Daten möglichst effizient und breit zu veröffentlichen. Gleichzeitig soll das Bewusstsein für den Umgang mit offenen Daten gefördert werden. In jedem Ressort steht zudem eine Open Data-Ansprechpartnerin bzw. ein Open Data-Ansprechpartner zur Verfügung, die bzw. der rund um Open Data beraten kann und die Aktivitäten im jeweiligen Ressort koordiniert.

## 2. Datenmanagement



In jedem modernen Berufsumfeld fallen tagtäglich riesige Mengen von Datensätzen an. Dank elektronischer Datenverarbeitung (EDV) ist das Erfassen und Organisieren der Informationen heute kein Problem mehr.

Durch die zunehmende Vernetzung und die steigende Datenflut entstehen neue Anforderungen an die Handhabung der Datenmengen. Im besten Fall sollen Daten also nicht mehrfach erfasst werden und außerdem jederzeit auffindbar sein und mehrere Personen müssen gleichzeitig auf denselben Datenbestand zugreifen können. Außerdem sollen sie sicher sein, z. B. vor Verlust durch Hardware-Defekte. Nicht zuletzt sind rechtliche

Aspekte zu beachten, d. h. Aufbewahrungsfristen, Aktualitätsanforderungen und Berichtspflichten sowie die Aufgabe der öffentlichen Verwaltung ihr Verwaltungshandeln transparent zu machen und ihre Daten der Öffentlichkeit zur Verfügung zu stellen.

Die Komplexität dieses Themenbereichs hat zur Entwicklung einer neuen Disziplin in der Informationstechnik geführt, das Datenmanagement. Im Sinne des Datenmanagements werden bestimmte Anforderungen an den Umgang mit digitalen Daten gestellt. Der Begriff beschreibt eher einen Prozess als einzelne Maßnahmen.

## 2.1 Definition Datenmanagement

Datenmanagement lässt sich als eine Daueraufgabe definieren, die beschreibt, wie Daten geordnet und strukturiert verwaltet werden können – ganz unabhängig vom jeweiligen Kontext. Bestandteile eines Datenmanagements sind Daten und Metadaten, die die Daten beschreiben und es ermöglichen, Daten zu ermitteln, in Verzeichnisse aufzunehmen und zu nutzen. Mit einem guten Datenmanagement können die Voraussetzungen für eine Optimierung vieler nachgelagerter Prozesse gebildet werden,

ebenso wie eine bessere Verfügbarkeit von Daten und der Wissensaufbau zum Umgang mit Daten. Um das zu erreichen, ist ein kontinuierlicher Prozess erforderlich. Diese Kontinuität schließt aber nicht aus, dass schrittweise Anpassungen vorgenommen werden können, um Teilschritte des Datenmanagements bei Bedarf zu ergänzen oder auszubauen.

## 2.2 Bedeutung von Datenmanagement für die Bereitstellung offener Daten der Behörden des Landes Nordrhein-Westfalen

Aufgrund der Bedeutung eines Datenmanagements wird für Open Data in § 6 der Open Data-Verordnung vorgeschrieben, ein Datenmanagement in den Behörden des Landes dauerhaft zu etablieren, um die effiziente Bereitstellung und Aktualisierung nach § 16a des E-Government-Gesetzes Nordrhein-Westfalen sicherzustellen.

Das Veröffentlichen von Daten und das damit verbundene Datenmanagement sollte dabei nicht als reiner Selbstzweck betrachtet werden, um den gesetzlichen Anforderungen des § 16a EGOvG NRW nachzukommen – denn v. a. durch die Nachnutzung der Daten durch unterschiedliche verwaltungsinterne und -externe Akteurinnen und

Akteure kann der Wert der Daten mitunter deutlich gesteigert werden. Die Förderung der Nutzung offener Daten ist somit zentral für deren „Wertschöpfung“.

Damit offene Daten für die Verwaltung selbst Nutzen entfalten und als Analyse- und Steuerungsinstrument effizient eingesetzt werden können, müssen dabei zwei zentrale Faktoren berücksichtigt werden: die jeweilige Fachlichkeit der Behörde, inkl. der fachlichen Zuständigkeiten, sowie die Erfahrung der Behörde in Bezug auf den Umgang mit Daten.

## Offene Daten erfüllen in ihrer Nachnutzung mehrere Hauptfunktionen:



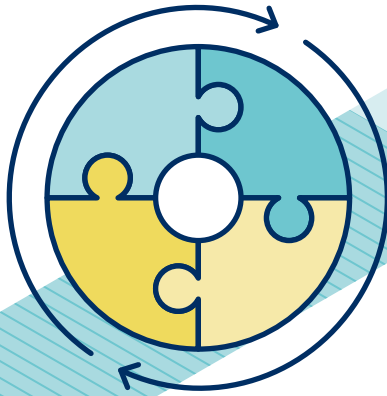
Abbildung 1: Hauptfunktionen der offenen Daten

Schlussendlich geht es darum, offene Daten einerseits in den einzelnen Behörden intern umfassend verfügbar zu machen, andererseits die konkrete Bereitstellung und Arbeit mit den Daten (Daten-

management) zu ermöglichen und verschiedenen Akteurinnen und Akteuren gezielt auch extern zugänglich zu machen.

# 3.

## Bausteine eines Datenmanagements



Mit dem Datenmanagement wird der Prozessablauf dargestellt, welchen es für die Erstellung und Veröffentlichung von Datensätzen und den dazugehörigen Metadaten zu etablieren gilt. Im Sinne eines Datenmanagement-Zyklus sollte dieser nach seiner ersten Implementierung regelmäßig durchlaufen werden. Das Vorgehensmodell basiert auf den Erfahrungen aus ähnlichen Projekten im nationalen und internationalen Umfeld (z. B. EDP, LANUV

Nordrhein-Westfalen, Urbane Datenräume) und ist auf die rechtlichen Rahmenbedingungen in Nordrhein-Westfalen zugeschnitten worden. Es dient als Blaupause. Dabei kann es zudem Ausgangspunkt für die Weiterentwicklung bzw. Analyse und Bewertung eines individuell auf den Bedarf einer Behörde des Landes Nordrhein-Westfalen zugeschnittenen Vorgehensmodells sein (vgl. → **Abschnitt 4** – Erweiterung des Datenmanagement-Zyklus).

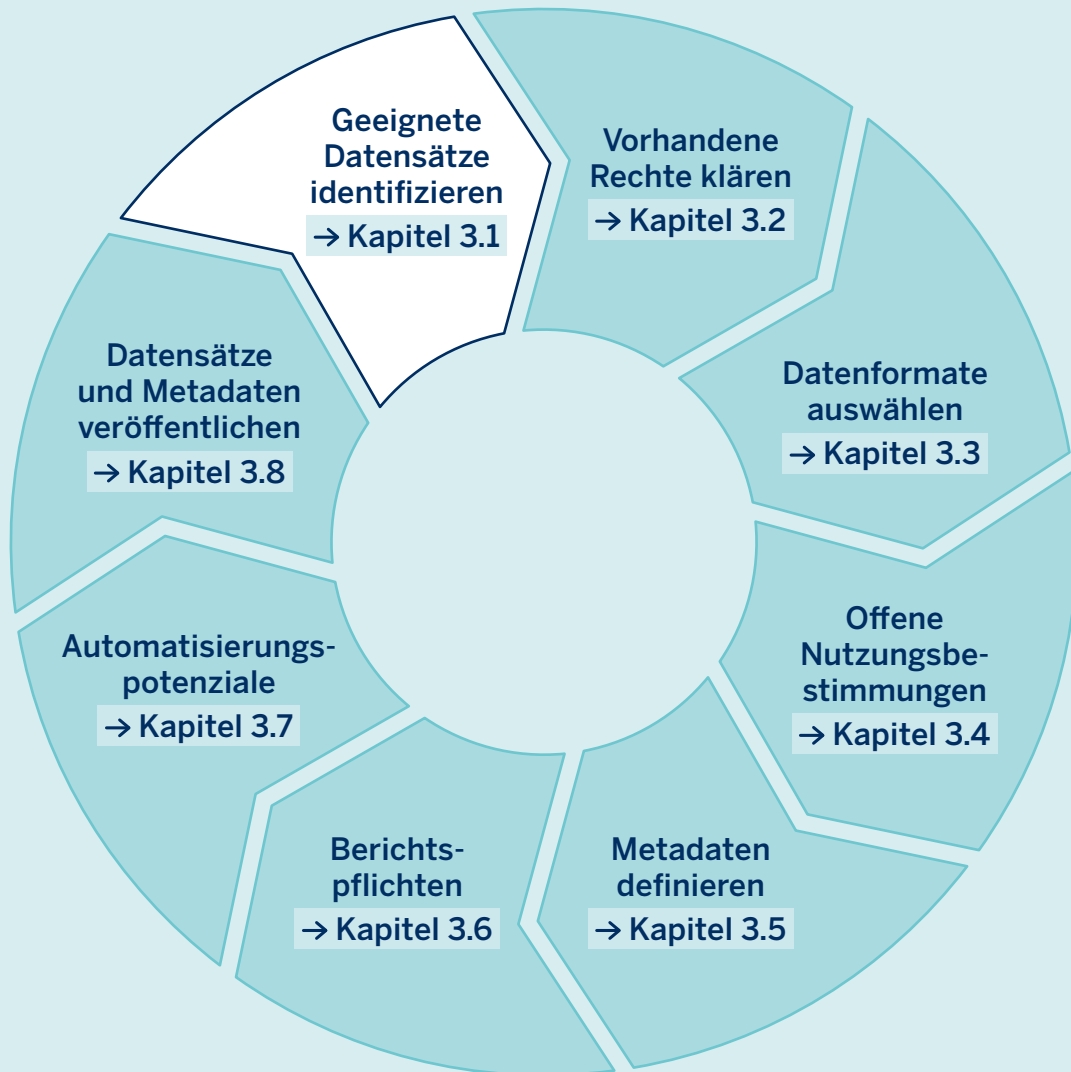


Abbildung 2: Datenmanagement-Zyklus

Die in → **Abbildung 2** aufgeführten Bausteine bieten den Behörden des Landes Nordrhein-Westfalen eine Orientierung, welche Prozessschritte zur Veröffentlichung von Daten zu etablieren

sind. In den nachfolgenden Kapiteln werden die einzelnen Bausteine und deren jeweilige Funktion im Datenmanagement-Zyklus erläutert. Geeignete Datensätze identifizieren

## 3.1 Geeignete Datensätze identifizieren



In diesem Kapitel werden Methoden vorgestellt, die helfen, für Open Data geeignete Datensätze zu ermitteln:

- Suche bereits veröffentlichter Daten (Fremdscreening)
- Hausinterne Abfrage (Selbstscreening)
- Durchführung von Dateninterviews ■

Zunächst stellt sich die Frage, welche Datensätze überhaupt innerhalb einer Behörde vorliegen und für eine Veröffentlichung im Sinne von Open Data geeignet sind. Im Folgenden werden verschiedene Methoden vorgestellt, die zur Beantwortung dieser Frage genutzt werden können. Wichtig ist in diesem Zusammenhang, dass die Wahl der Methoden keinen ausschließenden Charakter hat, sondern verschiedene Methoden auch ergänzend genutzt werden können. Weitere Maßnahmen, die diesen Prozess sinnvollerweise flankierend begleiten können, werden im **→ Kapitel 3.9.2 – Wissensmanagement** – behandelt.

Eine pragmatische Methode, geeignete Datensätze zu identifizieren, besteht darin, organisationsweit nach Daten zu suchen, die bereits veröffentlicht, aber noch nicht „offen“ im Sinne von Open Data sind. Dies ist beispielsweise der Fall, wenn die Daten noch nicht in einem offenen Format gemäß der Open Data-Verordnung vorliegen oder noch nicht auf dem Open.NRW-Portal bereitgestellt werden.

Bei diesen bereits veröffentlichten Daten dürften datenschutzrechtliche Fragestellungen grundsätzlich bereits geklärt sein, sodass ein geringerer Aufwand zu erwarten ist, wenn man diese Daten auch als Open Data bereitstellen will.

Für die Erfassung von Datensätzen einer Organisation, die bereits online auf Webseiten verfügbar sind, können Suchmaschinen oder Webcrawler<sup>2</sup> verwendet werden. Hierbei bietet es sich an, gezielt nach „nicht offenen“ Datenformaten zu suchen, bspw. nach PDF-Dokumenten.

Eine weitere Methodik zur Erhebung der insgesamt vorhandenen Datensätze in einer Organisation ist die organisationsweite Abfrage. Hierbei melden die Fachabteilungen ihre vorhandenen Datensätze z. B. mithilfe eines kurzen Fragebogens, sodass im Ergebnis eine organisationsweite Erfassung aller vorliegenden Datensätze in Form eines Datenkatalogs entsteht.

---

<sup>2</sup> Mithilfe von Webcrawlern und Harvestern können Domains nach bestimmten Dateitypen durchsucht werden. Die Ergebnisse werden in Form einer strukturierten Liste bereitgestellt.



Ergänzend können zur Identifizierung von geeigneten Datensätzen auch sogenannte „Dateninterviews“ geführt werden. Bei dieser Methode werden Einzel- oder Gruppeninterviews mit den Datenverantwortlichen aus den Fachorganisationseinheiten geführt. Bei diesem Vorgehen kann vor allem auf die aus Sicht der Datenverantwortlichen besonders für eine Veröffentlichung geeigneten Daten eingegangen werden. Darunter fallen bspw. Daten, bei denen eine hohe Nachfrage besteht. Diese Nachfrage kann sowohl bei internen als auch bei externen Stakeholderinnen und Stakeholdern bestehen. Eine hohe Nachfrage durch interne Stakeholderinnen und Stakeholdern kann sich bspw. durch häufige Fragen nach bestimmten Daten durch Kollegen und Kolleginnen aus anderen Abteilungen oder Referaten äußern. Gehen häufig Anfragen bzgl. Daten ein, die ihren Ursprung nicht in der eigenen Organisation haben, deutet dies auf eine hohe Nachfrage von Externen hin.

Hilfreich für die erfolgreiche Durchführung von Dateninterviews sind z. B. ein Leitfaden, anhand dessen die Interviews strukturiert werden, und ein Datenkatalog, der bereits eine grobe Übersicht über die vorhandenen Daten liefert. Die Methode der Dateninterviews ist ressourcenintensiv. Daher sollte sie nur in Ergänzung, aber nicht als grundlegende Erhebungsmethode eingesetzt werden.

Finden in der Behörde weitere Digitalisierungsvorhaben statt, wie bspw. Geschäftsprozessoptimierungen oder Umsetzungsprojekte im Rahmen des Onlinezugangsgesetzes, sollte bei der Durchführung dieser Vorhaben frühzeitig die Erhebung und Identifikation von Daten berücksichtigt werden (vgl. auch § 16a Abs. 8 EGovG NRW).

## 3.2 Vorhandene Rechte klären



In diesem Kapitel wird erläutert, dass einer Prüfung der Datensätze hinsichtlich möglicher rechtlicher Hinderungsgründe nichts entgegensteht. Dazu gibt es folgende unterstützende Dokumente:

- Open Data-Checkliste
- Praxisleitfaden für Datensouveränität ■

Zeitgleich zur Erhebung der vorhandenen Datenbasis oder im Anschluss folgt die Prüfung, ob die Datensätze unter die Veröffentlichungspflicht nach § 16a EGovG NRW fallen und ob sie veröffentlicht werden dürfen, weil keine sogenannten Hinderungsgründe nach § 16a Abs. 3 EGovG NRW wie bspw. datenschutzrechtliche Gründe vorliegen. Hier kann auf die → **Checkliste** zur Prüfung der Verpflichtungen aus § 16a EGovG NRW <sup>[3]</sup> zurückgegriffen werden.

Ein Bestandteil dieser → **Checkliste** ist bspw. die Prüfung, ob Urheberrechte einer Veröffentlichung entgegenstehen. Hat die Datenbereitstellerin bzw. der Datenbereitsteller die Urheberrechte selbst inne, so kann sie/er diese Daten ohne Einschränkungen bzgl. der Nutzungsrechte veröffentlichen.

Anders verhält es sich, wenn die Urheberrechte bei anderen Behörden oder auch bei Dritten, wie einem Unternehmen, liegen. Ist dies der Fall, dürfen die Daten nur mit Erlaubnis der Person mit Urheberschaft veröffentlicht werden. Dazu muss der datenveröffentlichenden Stelle ein vertragliches Nutzungsrecht für die zu veröffentlichenden Daten eingeräumt werden. Aus diesem Grund empfiehlt es sich, bei zukünftigen Vergabeverfahren die Veröffentlichung von Daten a priori als Anforderung festzuhalten (vgl. § 16a Abs. 8 EGovG NRW). Hierzu wird in Kürze eine Hilfestellung in Form eines

→ **Praxisleitfadens für Datensouveränität** zur Verfügung gestellt.



### Qualitätshinweis:

Ein wichtiges Merkmal von Open Data ist, dass die Daten keinen Personenbezug aufweisen. Sollte ein Datensatz einen Personenbezug aufweisen, ist dies aber nicht automatisch ein Hinderungsgrund für die Veröffentlichung, denn der Datensatz kann u. U. mit geeigneten Anonymisierungstechniken dennoch veröffentlicht werden. ■

Wurde festgestellt, dass ein Datensatz nach § 16a EGovG NRW zu veröffentlichen ist, muss der Datensatz in den nächsten Schritten in eine für die Veröffentlichung geeignete Form, beispielsweise hinsichtlich der Lizenz und des Dateiformats, gebracht werden. Die Open Data-Verordnung enthält entsprechende Vorgaben, die zu beachten sind. Manche Datensätze werden diese Vorgaben bereits erfüllen, während bei anderen Anpassungen vorgenommen werden müssen. Die nachfolgenden Kapitel zeigen auf, wie die Veröffentlichung unter Berücksichtigung der Vorgaben aus der Open Data-Verordnung umgesetzt werden kann, und erläutern dies.

## 3.3 Datenformate auswählen



In diesem Kapitel wird die Wahl der passenden Datenformate für die Bereitstellung der Datensätze behandelt:

- Es wird das Bestreben nach einer Bereitstellung von möglichst offenen Datei- oder Schnittstellenformaten beschrieben
- Es wird auf Anlage 1 der Open Data-Verordnung und die darin enthaltene Liste der zu verwendenden Formate verwiesen
- Zur Prüfung der maximalen Offenheit wird das 5-Sterne-Modell vorgestellt ■

Die Daten sollen maschinenlesbar und in offenen Datei- oder Schnittstellenformaten und vorrangig über offene Schnittstellen bereitgestellt werden. Die Open Data-Verordnung gibt mit Anlage 1 hier für geeignete Formate vor.



Abbildung 3: Daten- und Schnittstellenformate

Bei der letztendlichen Auswahl aus den Datenformaten sollte darüber hinaus einem an Nutzerinnen und Nutzern orientierten Ansatz gefolgt werden, mit dem Ziel, die Daten so offen wie möglich anzubieten.

Bezüglich einer ebenfalls auf Nutzerinnen und Nutzer ausgerichtete Veröffentlichung stellen sich die folgenden Fragen:

1. Wer sind die Adressaten der Daten?
2. Wie möchte die Zielgruppe die Daten nutzen und in welchem Kontext könnten die Daten eingesetzt und bearbeitet werden? (Siehe auch → **Kapitel 3.4** – Offene Nutzungsbestimmungen.)
3. In welcher Form benötigt die Zielgruppe die Daten (z. B. Schnittstelle, Format, Aktualität)?
4. Wie offen müssten die Daten sein und wie sollten sie aufbereitet sein, um die Qualität und Weiternutzung zu unterstützen?

Je nachdem wie diese Fragen beantwortet werden, können Schlüsse für die geeignete Bereitstellung der Daten gezogen werden. Bei Daten kann es die Weiterverarbeitung massiv erleichtern, wenn diese nicht nur lesbar, bspw. in einem PDF, sondern auch in einer editierbaren Form, z. B. in einer Excel-Datei, vorliegen (siehe Tabelle 1: 2 bzw. 3 Sterne). Bei Daten, die von einer hohen Aktualität geprägt sind (z. B. Hochwasserdaten oder Stellenmarktdaten), sollte das Datenformat so gewählt werden, dass dieser Notwendigkeit nach häufigen oder regelmäßigen Aktualisierungen bestmöglich nachgekommen werden kann. Dies ist bei Bereitstellung der Daten mittels einer Schnittstelle der Fall.

Da die verschiedenen möglichen Datenformate aus der Verordnung eine unterschiedliche Offenheit haben, sollte zusätzlich das 5-Sterne-Modell <sup>[4]</sup> herangezogen werden, um die max. mögliche Offenheit zu erreichen. Dieses Modell kommt auch im Open.NRW-Portal zum Einsatz. Im 5-Sterne-Modell werden die Datenformate anhand einer Offenheits-Skala bezüglich der Kosten und des Nutzens bewertet. Die Festlegung der Skala im Open.NRW-Portal ist folgendermaßen definiert:



Skala	Beschreibung
<b>&gt; Verfügbar</b> ★☆☆☆☆	Hierbei handelt es sich um einfache Daten, die im Web zur Verfügung gestellt werden. Nutzerinnen und Nutzer können die Daten anschauen, abspeichern, ausdrucken und teilen, aber nicht verändern (z. B. JPEG, PNG).
<b>&gt; Strukturiert</b> ★★☆☆☆	Hierbei handelt es sich um strukturierte Daten, die im Web zur Verfügung gestellt werden. Die Daten kann man direkt mit proprietärer Software bearbeiten, um sie zusammenzufassen, zu visualisieren, Berechnungen durchzuführen etc. (z. B. XLSX). Die Daten können außerdem in ein anderes (strukturiertes) Format exportiert werden.
<b>&gt; Nicht proprietär</b> ★★★☆☆	Hierbei handelt es sich nicht um die Verwendung von proprietären Daten, sondern um offene weitverbreitete Datenformate (z. B. Open-Document-Formate wie CSV).
<b>&gt; Referenzierbar</b> ★★★★☆	Hierbei handelt es sich um die Verwendung von eindeutigen, referenzierbaren Verlinkungen zu den Daten (z. B. RDF).
<b>&gt; Verknüpft</b> ★★★★★	Hierbei handelt es sich um Verknüpfungen der referenzierbaren Daten mit anderen Daten aus anderen Kontexten.

Tabelle 1: Open.NRW-Offenheits-Skala, abgeleitet aus dem 5-Sterne-Modell nach Tim Berners-Lee

Dies bedeutet im Umkehrschluss nicht, dass weniger gängige Datenformate nicht veröffentlicht werden können (beispielsweise Excel-Dateien als xls- und CSV-Datei). Vielmehr ist es wünschenswert, dass diese Datensätze dann zusätzlich und nicht ausschließlich im Portal bereitgestellt werden. Aufgrund der Vielzahl von Nutzerinnen und

Nutzern und ihren unterschiedlichen Interessen kann es im Ergebnis auch sinnvoll sein, die Daten in mehreren Formaten zu veröffentlichen. Neben den vorgegebenen Formaten der Open Data-Verordnung können Datensätze also aus verschiedenen, weiteren Formaten bestehen und zusätzlich in diesen bereitgestellt werden.

## 3.4 Offene Nutzungsbestimmungen



In diesem Kapitel geht es um die Festlegung der Nutzungs- und Lizenzbestimmung für die spätere Weiternutzung der bereitgestellten Open Data-Datensätze, welche in den Lizenzbestimmungen und Nutzungsrechten der Open Data-Verordnung (§ 5 Abs. 1) und in Anlage 2 der Verordnung beschrieben sind. ■

Bei der Veröffentlichung von Daten legen Nutzungsbestimmungen fest, unter welchen Bedingungen ein Datensatz verwendet werden darf. Die Wahl der Nutzungsbestimmungen für die zur Bereitstellung vorgesehenen Daten ist ein wichtiger Punkt, der nicht nur in Hinblick auf Rechte Dritter und spezielle gesetzliche Vorgaben in Nordrhein-Westfalen, sondern auch vor dem Hintergrund des Anwendungskomforts berücksichtigt werden muss.

In der Open Data-Verordnung wird mit § 5 Abs. 1 festgelegt, dass die bereitgestellten Daten einschließlich zugehöriger Metadaten mit einem Nutzungsrecht (Lizenz) zu versehen sind. Die Lizenz ist dabei so zu wählen, dass die bereitgestellten Daten frei und uneingeschränkt genutzt werden können. Die Anlage 2 der Verordnung nennt die für die Bereitstellung von Daten nach § 16a EGovG NRW zu verwendenden Lizenzen:

**Folgende Lizenzen sind bevorzugt zu nutzen, da sie eine einfache Weiterverwendung der Daten ermöglichen**

→ **Datenlizenz Deutschland – Zero – Version 2.0**

→ **CC0 1.0 Universell (CC0 1.0) Public Domain Dedication**

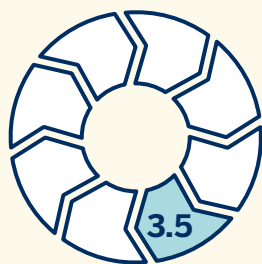
**Folgende Lizenzen sind zu nutzen, wenn eine Namensnennung erfolgen muss**

→ **Datenlizenz Deutschland – Namensnennung – Version 2.0**

→ **Creative Commons Namensnennung 4.0 International (CC BY 4.0)**

In Ausnahmefällen kann nach § 5 Abs. 2 aus bestimmten Gründen auch eine andere Lizenz gewählt werden. Diese muss jedoch den vorgegebenen Lizenzbedingungen aus Absatz 1 und Absatz 2 entsprechen. <sup>[1]</sup>

## 3.5 Metadaten definieren



In diesem Kapitel wird die Handhabung von Metadaten im Open Data-Kontext behandelt:

- Über Metadaten werden die Datensätze öffentlich such- und auffindbar gemacht
- Das Open.NRW-Portal ist die zentrale Metadatenplattform für die Veröffentlichung von Metadaten
- Als Grundlage für die Veröffentlichung der Metadaten wird das Metadatenschema DCAT-AP.de verwendet ■

Im Veröffentlichungsprozess von Datensätzen sind Metadaten von zentraler Bedeutung, denn durch sie werden die Daten der Behörden erst öffentlich auffindbar gemacht. Metadaten sind Informationen, die Daten beschreiben und es ermöglichen, Daten zu ermitteln, in Verzeichnisse aufzunehmen und zu nutzen.<sup>[5]</sup> Die Verwendung von Metadaten ist vor allem für Onlineplattformen und Suchmaschinen von großer Bedeutung. Durch sie ist es möglich, offene Daten zu suchen und leichter auffindbar zu machen.

Für die Veröffentlichung von Metadaten offener Verwaltungsdaten wird in Deutschland der vom IT-Planungsrat festgelegte formale Austauschstandard für offene allgemeine Verwaltungsdaten – DCAT-AP.de –<sup>[5]</sup> verwendet. Bei DCAT-AP.de handelt es sich um das deutsche Profil des europäischen Metadatenstandards DCAT-AP. Mit DCAT-AP wird auch die Kompatibilität mit dem EU-Standard sichergestellt.<sup>[6]</sup>

In der Spezifikation von DCAT-AP.de<sup>[6]</sup> werden alle wichtigen Informationen beschrieben, die man benötigt, um die Metadaten in das Open.NRW-Portal zu überführen. Die Inhalte beziehen sich auf unterschiedliche Klassen, die als verpflichtend, empfohlen oder optional gekennzeichnet sind. Die wichtigsten Klassen im DCAT-AP.de-Modell sind Dataset, Catalog und Distribution.

In der Klasse **Dataset** werden die Datensätze beschrieben, welche in dem Open.NRW-Portal bereitgestellt werden und an anderer Stelle, z. B. der Homepage der Behörde, veröffentlicht sind. Der **Catalog** beinhaltet die Metadaten zu den Datensätzen und die **Distribution** beschreibt, wie auf die Datensätze zugegriffen wird bzw. wie diese bereitgestellt werden können. Für ein besseres Verständnis, wie das Metadatenmodell im Open.NRW-Portal verankert ist, wird dieses nachfolgend anhand eines Beispiels aus dem Open.NRW-Portal beschrieben:

## Klasse: Catalog

➤ Titel	Open.NRW – das Open-Government-Portal in Nordrhein-Westfalen
➤ Beschreibung	Transparenz und Teilhabe sind zentrale Bausteine des Open Governments in Nordrhein-Westfalen. Das Open.NRW-Portal ist Informations- und zugleich Open Data-Plattform der Landesregierung Nordrhein-Westfalen. Als Nutzenden der Plattform steht Ihnen sowohl ein stetig wachsender Bestand an offenen Verwaltungsdaten zur Verfügung als auch zahlreiche Beispiele und Informationen zu laufenden und vergangenen Projekten. Außerdem erfahren Sie, auf welche Weise Sie sich aktiv einbringen können.
➤ URL	<a href="https://open.nrw/">https://open.nrw/</a>

## Klasse: Dataset

➤ Titel	Sitzungstermine der politischen Gremiensitzungen Münster
➤ Beschreibung	Dieser Datensatz enthält den Sitzungskalender der politischen Gremiensitzungen Münster für das Jahr 2022. Im Rahmen der Open Data-Initiative der Stadtverwaltung Münster erhalten Sie auf dieser Seite den Sitzungskalender in maschinenlesbaren Formaten, die z. B. in Tabellenkalkulations- und Kalenderprogrammen verarbeitet werden können. Weitere Informationen sowie eine Liste der Gremientermine für den Internetbrowser erhalten Sie unter folgender URL des Ratsinformationssystems Münster: <a href="https://www.stadtmuenster.de/sessionnet/sessionnetbi/info.php">https://www.stadtmuenster.de/sessionnet/sessionnetbi/info.php</a>
➤ Veröffentlichende Stelle	Leider wurden keine Kontaktdaten hinterlegt
➤ Veröffentlicht über	Stadt Münster
➤ Erstellt am	08.03.2022
➤ Aktualisiert am	10.08.2022
➤ Aktualisierungsfrequenz	Jährlich
➤ Kategorie	Regierung und öffentlicher Sektor
➤ Schlagwörter	Politik und Wahlen
➤ Offenheits-Skala	Nicht proprietär
➤ Links	<a href="https://opendata.stadt-muenster.de/">https://opendata.stadt-muenster.de/</a> <a href="https://www.stadt-muenster.de/ratsservice/politik-und-verwaltung">https://www.stadt-muenster.de/ratsservice/politik-und-verwaltung</a> <a href="https://open.nrw/dcatap/dataset/sitzungstermine-der-politischen-gremiensitzungen-munster-ms.rdf">https://open.nrw/dcatap/dataset/sitzungstermine-der-politischen-gremiensitzungen-munster-ms.rdf</a>



## Klasse: Distribution

➤ URL	<a href="https://opendata.stadt-muenster.de/sites/default/files/ratsinformation_termine.zip">https://opendata.stadt-muenster.de/sites/default/files/ratsinformation_termine.zip</a>
➤ Beschreibung	In diesem ZIP-Archiv enthalten ist eine iCalendar-Datei mit den Sitzungsterminen der politischen Gremien in Münster 2022.
➤ Dokumentation	<a href="https://opendata.stadt-muenster.de/dataset/sitzungstermine-der-politischen-gremiensitzungen-m%C3%BCnster/resource/dee53d2f-108a-433f-bfe0">https://opendata.stadt-muenster.de/dataset/sitzungstermine-der-politischen-gremiensitzungen-m%C3%BCnster/resource/dee53d2f-108a-433f-bfe0</a>
➤ Lizenz	Datenlizenz Deutschland Namensnennung 1.0

Tabelle 2: Beispiel, basierend auf den wichtigsten Klassen des DCAT-AP.de-Metadatenmodells

Das Open.NRW-Portal ist das zentrale Metadatenportal für Verwaltungsdaten aus Nordrhein-Westfalen. Daten, die nach dem E-Government-Gesetz Nordrhein-Westfalen als Open Data zu veröffentlichen sind, sind samt Metadaten über Open.NRW zugänglich zu machen. Es kommen – abhängig von einigen Faktoren – verschiedene Wege der Veröffentlichung in Betracht.

**1. Festlegung des geeigneten Portals für die initiale Veröffentlichung:** Grundsätzlich müssen die Metadaten aber nur einmal veröffentlicht und bereitgestellt werden, in Abhängigkeit von ihrer Thematik und einem etwaigen Raumbezug entweder im GEOportal.NRW oder im Open.NRW-Portal. Da zwischen den beiden Portalen ein automatisierter Austausch von Metadaten erfolgt, werden auf dem GEOportal.NRW bereits veröffentlichte offene Daten ebenso auf Open.NRW veröffentlicht. Für die Personen, die die Daten bereitstellen, ist nichts weiter zu tun. Genauere Informationen hierzu sind im [→ Leitfaden zur Metadatenerfassung](#) <sup>[7]</sup> und im [→ Open Data-Guide](#) <sup>[2][3]</sup> auf Open.NRW zu finden.

**2. Manuelle Veröffentlichung:** Ist dies nicht der Fall und die Daten werden auch an anderer Stelle nicht veröffentlicht, können sie über das Open.NRW-Portal über entsprechende Metadaten zur Verfügung gestellt werden. Zur Erfassung der Metadaten stellt das Portal eine kostenfreie Softwarelösung zur Verfügung. Es muss eine Registrierung <sup>[8]</sup> erfolgen. **Achtung:** Metadaten zu Daten mit einem geografischen Raumbezug sowie INSPIRE-relevante oder GDI-DE-pflichtige Daten müssen im GEOportal.NRW erfasst werden.<sup>3</sup> Durch diesen

starken Raumbezug unterscheidet sich die Erfassung von Metadaten im GEOportal.NRW zu einem gewissen Grad von der Erfassung im Open.NRW-Portal.

**3. Automatisierte Veröffentlichung:** Wurden die eigenen Daten bereits online veröffentlicht, z. B. auf der eigenen Homepage oder einem Portal, oder besteht eine umfassende interne Datenbank, müssen die Daten zudem über das Open.NRW-Portal öffentlich auffindbar gemacht werden. Hierzu ist vorgesehen, dass eine Schnittstelle mit DCAT-AP.de angeboten wird. Genauere Informationen finden Sie in [→ Kapitel 3.7 – Automatisierungspotenziale](#).



### Qualitätshinweis:

Über Metadaten werden Daten überhaupt erst auffindbar. Die Auffindbarkeit der Daten leidet jedoch, wenn die Metadaten eine geringe Qualität aufweisen. Daher sollten stets, wo möglich, kontrollierte Vokabulare verwendet werden. Titel und Beschreibungen des Datensatzes sollten so verständlich wie möglich sein und um zeitliche und geografische Angaben ergänzt werden. ■

<sup>3</sup> Das GEOportal.NRW ist die zentrale Vermittlungsstelle zwischen Nutzerinnen bzw. Nutzern und Bereitstellerinnen bzw. Bereitstellern von Geodaten in Nordrhein-Westfalen. Weiterhin dient es als Informationsplattform hinsichtlich des aktuellen Stands der INSPIRE-Umsetzung und des Aufbaus der Geodateninfrastruktur in Nordrhein-Westfalen (GDI-NW).

## 3.6

# Bestehende Datenberichtspflichten



In diesem Kapitel werden Bezüge zwischen Datenberichtspflichten als wiederkehrenden Verwaltungsabläufen und der Open Data-Veröffentlichungspolitik aufgezeigt. ■

Ein erfolgreiches Open Data-Datenmanagement in den Behörden des Landes Nordrhein-Westfalen sollte nicht losgelöst von den zum Teil bereits seit mehr als einem Jahrzehnt existierenden elektronischen Datenmanagement-Prozessen etabliert werden, wie sie unter anderem im Kontext von Datenberichtspflichten bestehen (z. B. INSPIRE oder Berichtspflichten der Bundesbehörden gegenüber der EEA).

Bei Berichtspflichten handelt es sich um Verwaltungsabläufe, welche wiederkehrend zu einem vorgegebenen Zeitpunkt durchgeführt und an eine verantwortliche Stelle übersendet werden müssen. § 12 Abs. 1 S. 1 EGovG NRW legt fest, dass ab Dezember 2025 alle diese Verwaltungsabläufe auf elektronischem Weg durchzuführen sind. Dabei müssen die elektronische Akte, die elektronische Laufmappe und das ersetzende Scannen eingeführt werden. Die Verwaltungsabläufe werden dadurch wesentlich vereinfacht und schaffen neue Möglichkeiten für die Behörden beim Umgang mit bestehenden Datenberichtspflichten.<sup>[9]</sup>

Die Feststellung bestehender Datenberichtspflichten ist deshalb ein essenzieller Schritt und die Grundlage für Automatisierungspotenziale, Effektivitäts- und Effizienzsteigerungen (wie zum

Beispiel bei der Veröffentlichungspflicht nach der INSPIRE-Richtlinie oder bei Geodaten, die bereits als Open Data maschinenlesbar, mit genormten Metadaten beschrieben sowie als Metadaten und Daten über APIs verfügbar sind). Auch in (PDF-) Dokumenten können dabei Daten enthalten sein, die als offene Daten einen Mehrwert für andere Akteurinnen und Akteure bedeuten würden und daher extrahiert werden müssen.

Auch kann es im Kontext der Datenberichtspflichten darum gehen, „doppelte“ Arbeit zu vermeiden, indem bereits existierende Kanäle und Prozesse mit- oder nachgenutzt werden.

Um einen Überblick darüber zu bekommen, wie ein Workflow zu Datenberichtspflichten aussehen kann, wird in der nachfolgenden Abbildung 4 ein solcher am Beispiel der Datenberichtspflicht des Landesamts für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen (LANUV)<sup>[10]</sup> bzgl. der Luftqualitätskontrolle anhand der EU-Luftqualitätsrichtlinie 2008/50/EG<sup>[10]</sup> aufgezeigt.



### Beurteilung der Luftqualität

- Durchführung von Messungen der Luftqualität
- Beurteilung der Messungen in Bezug auf gesundheitliche Schadstoffe (z. B. Kohlenmonoxid)



### Kontrolle der Luftqualität

- Kontrolle der Messungen anhand von vorgegebenen Grenzwerten
- Erforderliche Maßnahmen bei Überschreitung der Grenzwerte



### Pläne

- Erstellung von Luftqualitätsverbesserungsplänen für die Gebiete mit überschrittenen Grenzwerten
- Erstellung der Pläne muss innerhalb von zwei Jahren nach Feststellung erfolgen



### Berichtspflicht

- Bereitstellung der Informationen zur Luftqualität in Form von Jahresberichten
- Informationen müssen kostenlos über leicht zugängliche Medien zur Verfügung gestellt werden

Abbildung 4: Beispiel zur Datenberichtspflicht des LANUV zur Luftqualität

Das Beispiel zur Kontrolle der Luftqualität soll deutlich machen, dass in jeder behördlichen Einrichtung Datenberichtspflichten vorhanden sind und die entsprechenden Daten nach § 16a EGovG NRW auch als Open Data zur Verfügung gestellt werden müssen. Dabei ist es möglich, dass an ver-

schiedenen Stellen einzelne oder mehrere Verwaltungsabläufe in automatisierte Prozesse überführt und somit die Arbeiten erleichtert und beschleunigt werden. Konkrete Automatisierungspotenziale müssen dabei von Fall zu Fall identifiziert und implementiert werden.

## 3.7 Automatisierungspotenziale



In diesem Kapitel geht es um die Identifizierung von wiederkehrenden Prozessschritten im Datenmanagement, mit dem Ziel, diese zu automatisieren und dadurch die Arbeitsprozesse der Behörden zu erleichtern und zu beschleunigen. ■

Eine große Anzahl von Datensätzen oder auch hochdynamische Daten wie Echtzeitdaten in einzelnen Behörden sind jeweils Indikatoren dafür, dass bei der wiederkehrenden Aufgabe der Veröffentlichung und Pflege von offenen Daten der Automatisierung verschiedener Arbeitsschritte eine entscheidende Bedeutung zukommt. In der Regel sollten bereits existierende Informations- und Kommunikationswege und damit verbundene Prozesse nachgenutzt oder nutzbar gemacht werden.

Es besteht dabei die Möglichkeit, verschiedene Prozessschritte im Datenmanagement zu automatisieren. Dadurch können die Arbeitsprozesse der Behörden erleichtert werden. Es ist daher von großer Wichtigkeit, dass jede Behörde wiederkehrende Arbeitsprozesse identifiziert und diese mit den Ansprechpersonen des Open.NRW-Portals bezüglich der notwendigen Ressourcen, Standards sowie vorhandenen Schnittstellen und deren

Möglichkeiten abstimmt. Die Open Data-Ansprechpersonen des jeweiligen Ressorts können hierbei unterstützen.

Automatisierungspotenziale gibt es unter anderem bei:

1. der automatisierten Ableitung der Metadaten aus einem externen Metadaten-System,
2. der automatisierten Ableitung der Metadaten aus den Daten,
3. der automatisierten Aktualisierung der Metadaten im Open.NRW-Portal,
4. der automatisierten Bereitstellung der Metadaten im Metadatenkatalog des Open.NRW-Portals.

Die automatisierte Ableitung der Metadaten aus externen Metadaten-Systemen wird am Beispiel eines bereits existierenden Prozesses der Stadt Köln illustriert:

Die Stadt Köln verfügt über eine für die Erfassung und Pflege der Metadaten verantwortliche Person. Diese Metadaten werden in einem eigenen Metadaten-System der Stadt Köln gespeichert, dem stadt-eigenen Geoportal. Damit die Metadaten in das Metadaten-System des Open.NRW-Portals automatisiert überführt werden können, wird ein Ableitungsmechanismus (Harvester) benötigt. Mit diesem Ableitungsmechanismus wird die Schnittstelle des Metadaten-Systems der Stadt Köln ausfindig gemacht und die Metadaten werden in

das Open.NRW-Metadaten-System überführt. Hierbei müssen die Metadaten der Stadt Köln im DCAT-AP.de-Metadaten-Schema vorliegen, um ins Open.NRW-Portal übernommen werden zu können.

Der Ableitungsprozess wird durch Administratorinnen und Administratoren des Open.NRW-Portals initialisiert und ausgeführt, die Stadt Köln muss lediglich das eigene Metadaten-System (Geoportal) pflegen und – falls nicht vorhanden – eine DCAT-AP.de-konforme Schnittstelle anbieten. In der nachfolgenden → **Abbildung 5** ist der beschriebene Ableitungsmechanismus der Metadaten der Stadt Köln beispielhaft dargestellt.

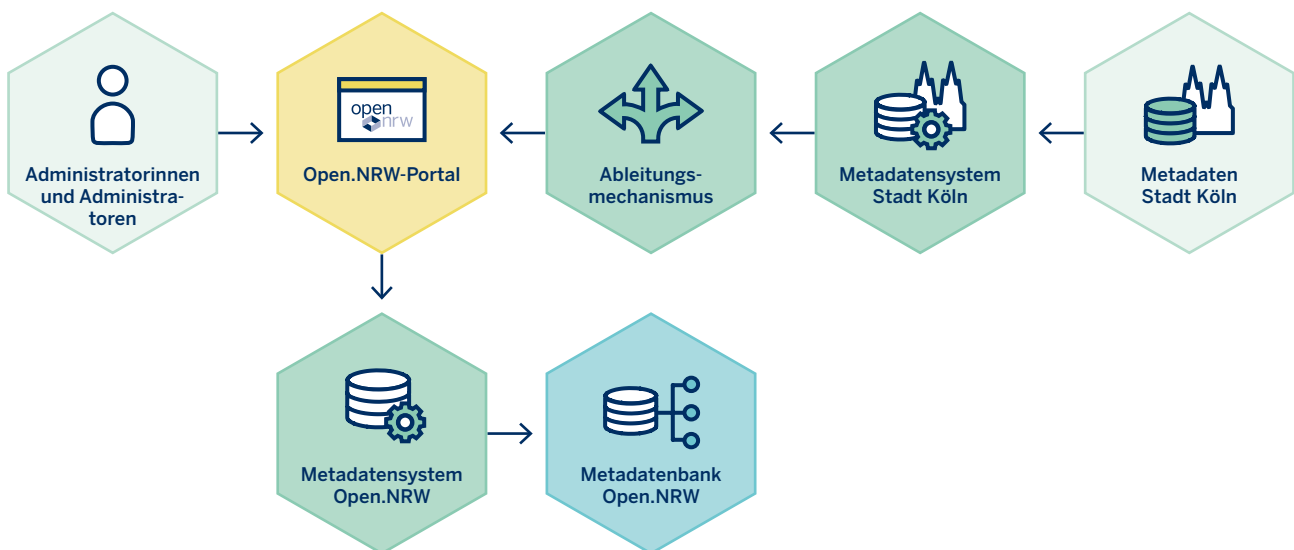


Abbildung 5: Ablauf des Ableitungsmechanismus der Metadaten am Beispiel der Stadt Köln

Der oben beispielhaft beschriebene Lösungsansatz steht bereits im Open.NRW-Portal [12] zur Verfügung und ist in der Lage, die Metadaten aus

anderen Portalen abzuleiten. Hierzu wird als Vorgabe das standardisierte DCAT-AP.de-Metadatenmodell verwendet.

Im folgenden Beispiel wird gezeigt, wie die automatisierte Ableitung der Metadaten aus den Daten erfolgen kann. Dafür wird das Beispiel zur Daten-

berichtspflicht des LANUV in Bezug auf Luftqualität (→ Kapitel 3.6) nachfolgend abgebildet (→ siehe Abbildung 6).

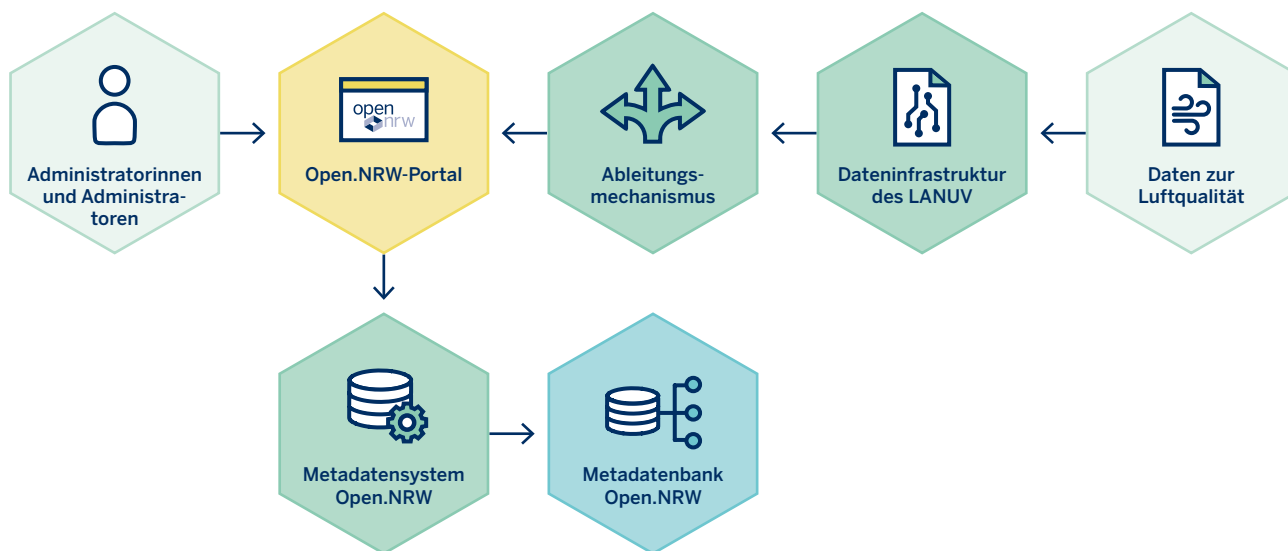


Abbildung 6: Ablauf des Ableitungsmechanismus der Metadaten am Beispiel der Berichtspflichten des LANUV zur Luftqualität

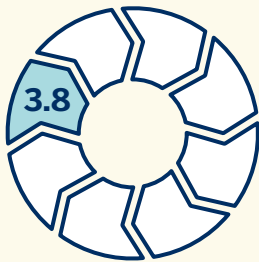
Die Administratorinnen und Administratoren des Open.NRW-Portals initialisieren und führen den erforderlichen Ableitungsmechanismus aus. Dabei muss die Dateninfrastruktur des LANUV über eine Schnittstelle zugänglich sein. Der Ableitungsmechanismus enthält die wesentlichen Informationen zur Ableitung der Metadaten direkt aus den Daten zur Luftqualität. Es werden die einzelnen Informationen aus den Daten über einen Befehl abgerufen. Dabei ist es wichtig, dass die Metadaten des Datensatzes zur Luftqualität dem Metadatenmodell des Open.NRW-Portals entsprechen, damit sie ohne zusätzliche Anpassungen im Ableitungsmechanismus abgeleitet werden können. Wenn die Metadaten nicht im gleichen Metadatenmodell vorliegen, werden sie während des Ableitungsmechanismus umgewandelt. In diesem Prozess können nur die

Metadaten zur Luftqualität abgeleitet werden, da der Ableitungsmechanismus an die spezifischen Eigenschaften der Daten angepasst ist. Es können nur die Metadaten abgeleitet werden, welche in den Daten enthalten sind, individuelle Titel oder Beschreibungstexte müssten im Nachgang im Open.NRW-Portal eingepflegt werden. Über eine Komponente zur Überprüfung der Qualität können die Metadaten im Open.NRW-Portal auf ihre Richtigkeit analysiert werden.

Voraussetzung für die Nutzung der Automatisierung ist, dass die datenhaltenden Stellen über Kenntnisse in der Erstellung und Pflege von Metadaten verfügen, die bei Bedarf durch Handlungsanleitungen aufgebaut oder gefördert werden sollten.

## 3.8

# Datensätze und Metadaten veröffentlichen



In diesem Kapitel geht es um die tatsächliche Veröffentlichung der Daten:

- > Veröffentlichung der Datensätze über eine Dateninfrastruktur
- > Beschreibung der technischen Möglichkeiten zur Veröffentlichung von Metadaten im Open.NRW-Portal des Landes Nordrhein-Westfalen ■

Die Veröffentlichung der Datensätze selbst erfolgt durch die Datenbereiterinnen bzw. Datenbereiter und benötigt eine Dateninfrastruktur. Bei Datensätzen handelt es sich um die originären Daten (z. B. CSV), welche im Open.NRW-Portal durch Metadaten (z. B. Titel, Beschreibung etc.) beschrieben werden und auffindbar sein sollen. Die Daten-Infrastruktur dient dazu, dass die originären Daten online zur Verfügung stehen und abrufbar sind. In der Regel wird eine eigene Dateninfrastruktur der Behörde benötigt. Behörden, welche über eine eigene Dateninfrastruktur verfügen, können ihre Datensätze über Verlinkungen (Zugangs-URL) in den Metadaten zu der jeweiligen Quelle des Datensatzes im Open.NRW-Portal bereitstellen. Für Behörden, welche noch nicht über eine eigene Dateninfrastruktur verfügen und bei denen aufgrund der Beschaffenheit und/oder des Umfangs es auch nicht verhältnismäßig erscheint, diese aufzubauen, wird es in Zukunft die Möglichkeit geben, Datensätze in angemessenem Umfang direkt im Open.NRW-Portal zu veröffentlichen.



### Qualitätshinweis:

Die Daten und deren dazugehörige Metadaten sollte man nach der initialen Veröffentlichung im Open.NRW-Portal regelmäßig auf ihre Aktualität überprüfen, um unvollständige oder fehlerhafte (Meta-) Datensätze zu vermeiden. Es ist daher empfehlenswert, dass jede Behörde festlegt, wer für die Erfassung und Pflege der Daten und deren Metadaten verantwortlich ist. Die Zuständigkeiten und Rollen der Fachpersonen können in den Behörden unterschiedlich ausgestaltet sein. In den Geschäftsbereichen der Landesbehörden gibt es Open Data-Ansprechpartnerinnen und -Ansprechpartner, die hierbei weiterhelfen und Auskunft geben können. ■

Die nachfolgende Abbildung 7 zeigt beispielhaft die Veröffentlichung von Metadaten, welche im Open.NRW-Portal bereitgestellt werden können.

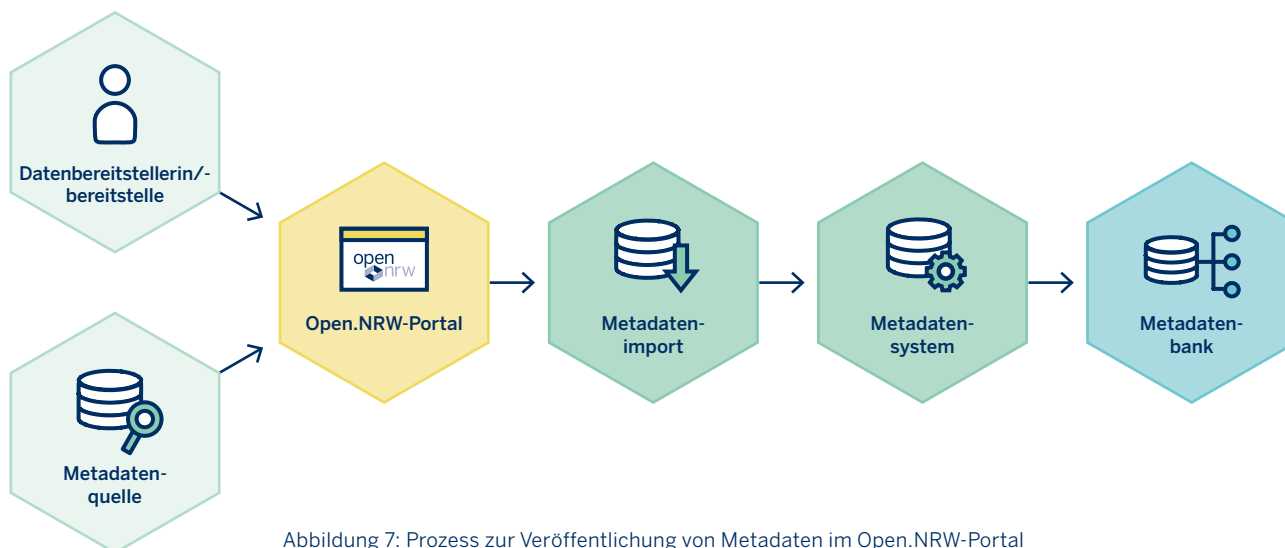


Abbildung 7: Prozess zur Veröffentlichung von Metadaten im Open.NRW-Portal

Die Metadatenquelle kann von dem automatisierten Ableitungsmechanismus des Open.NRW-Portals angefragt werden und dann lassen sich die Informationen extrahieren. Diese werden anschließend über den Metadatenimport in das Metadatensystem der Metadatenbank des Open.NRW-Portals gespeichert. Dieser Prozess kann auch über Datenbereitstellerinnen und -bereitsteller erfolgen, welche im Open.NRW-Portal die Metadaten händisch in das Metadatensystem überführen.



### Qualitätshinweis:

Wenn Daten auf einer eigenen Infrastruktur veröffentlicht und über Verlinkungen im Open.NRW-Portal bereitgestellt werden, muss die Zugänglichkeit der Daten gewährleistet sein. Die gesetzten Verlinkungen müssen langlebig sein. Es gilt, insbesondere sogenannte Broken Links zu vermeiden, also Links, die auf ein nicht mehr erreichbares Ziel verweisen. ■



## 3.9 Qualitäts- und Wissensmanagement

### 3.9.1 Qualitätsmanagement

In diesem Kapitel geht es um die Qualitätseigenschaften der Datensätze und Metadaten:

- Diese Qualität muss vor der Veröffentlichung sichergestellt werden
- Dazu gibt es im Open.NRW-Portal eine Metadata-Quality-Assessment-Komponente (MQA), über die die Qualitätseigenschaften von Metadaten überprüft werden können. ■

Damit als Open Data veröffentlichte Daten von Dritten mehrwertbringend verwendet werden können, müssen die Daten von guter Qualität sein. Zu Einschränkungen bei der Weiterverwendung der Daten kommt es beispielsweise, wenn die Daten unvollständig, veraltet oder nicht korrekt sind, wenn Titel und Beschreibungen fehlen oder unvollständig sind und die Daten in proprietären und nicht maschinenlesbaren Formaten veröffentlicht werden. Ziel muss es daher sein, die Daten vor der Veröffentlichung in einen qualitativ hochwertigen Zustand zu bringen. Das Qualitätsmanagement kann dabei als Querschnittsthema betrachtet werden, welches in verschiedenen Phasen des Datenmanagement-Prozesses mitgedacht werden muss. Zahlreiche Qualitätshinweise finden Sie zum Beispiel in den Kapiteln zu den folgenden Themen:

- **Datensätze identifizieren**
- **Metadaten definieren**
- **Datensätze veröffentlichen**

Der Leitfaden zu qualitativ hochwertigen Daten und Metadaten <sup>[13]</sup> gibt darüber hinaus zahlreiche weitere Handlungsempfehlungen für die Aufwertung von konkreten Datenformaten, wie bspw. CSV/XLSX, XML, JSON, GeoJSON, RDF, REST und WFS. Die dort enthaltenen Qualitätsmerkmale sollten

nicht erst bei einer anstehenden Veröffentlichung von Daten umgesetzt werden, sondern vielmehr bereits bei der Erstellung eines Datensatzes Berücksichtigung finden. So lässt sich der Aufwand, der vor einer Veröffentlichung für eine manuelle Aufbereitung der Daten anfallen würde, auf ein Minimum reduzieren oder sogar gänzlich vermeiden.

Nach der Veröffentlichung der Daten sollte eine kontinuierliche Qualitätssicherung erfolgen. Hierzu gehört beispielsweise eine regelmäßige Überprüfung der Metadaten auf Aktualität und Korrektheit. Webstatistiken und Zugriffszahlen der Portale, auf denen Daten veröffentlicht werden, kann man darüber hinaus nutzen, um herauszufinden wie häufig bestimmte Datensätze heruntergeladen werden. Das Open.NRW-Portal veröffentlicht regelmäßig eine Webstatistik mit bestimmten Kennzahlen. Solche Informationen können hilfreich dabei sein, die Nachfrage nach Datensätzen besser einzuschätzen und entsprechend das Datenangebot zu justieren. Für besonders nachgefragte Datensätze könnte bspw. ein kürzeres Aktualisierungsintervall gewählt werden, während für weniger nachgefragte Datensätze der Aufwand bei der Vorbereitung und Veröffentlichung minimiert wird.

Das Open.NRW-Portal bietet den Datenberei-  
 tstellerinnen und -bereitstellern die Möglich-  
 keit, mittels der sogenannten Metadata-Quality-  
 Assessment-Komponente (MQA) die Qualität der  
 eigenen Metadaten zu analysieren und nachzu-  
 vollziehen, inwieweit diese im Open.NRW-Portal  
 genutzt und verbessert werden können. Open.  
 NRW bietet durch MQA die Möglichkeit, dass jede  
 Datenberei- tstellerin und jeder Datenberei-  
 tsteller (in MQA als Katalog bezeichnet) ihre bzw. seine

spezifischen Metadatenauswertungen, ein sog.  
 Dashboard, sieht. Das Katalog-Dashboard beant-  
 wortet Fragen zur allgemeinen Qualität wie Auf-  
 findbarkeit, Zugänglichkeit, Interoperabilität oder  
 Wiederverwendbarkeit und zu speziellen Metada-  
 tenelementen. Hervorzuheben sind Auswertun-  
 gen zur Maschinenlesbarkeit oder Lizenzangabe,  
 welche für offene Daten elementar sind. [14] Die  
 nachfolgende → **Abbildung 8** zeigt das Dashboard.

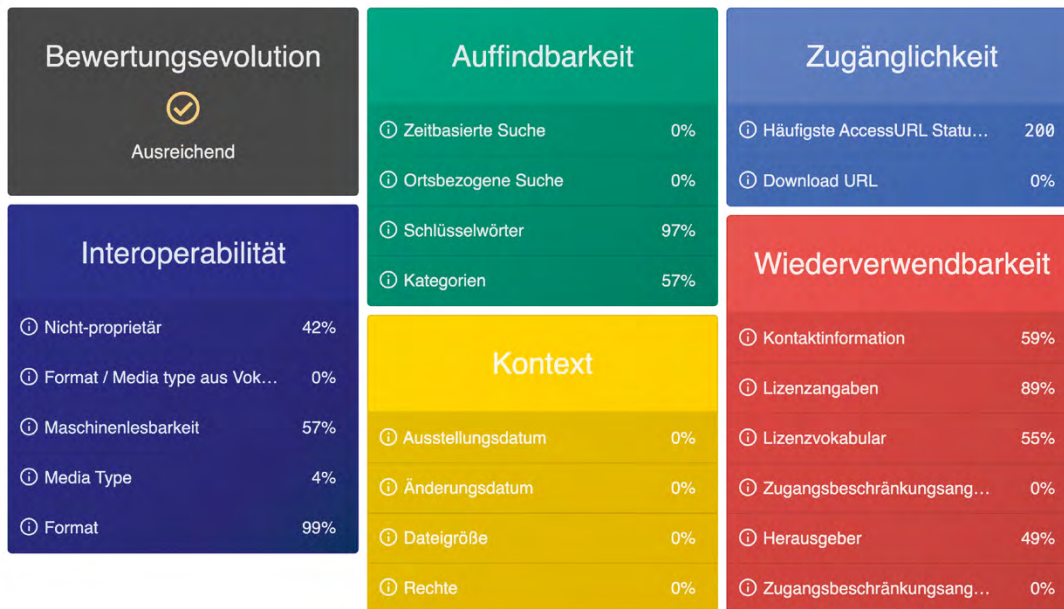


Abbildung 8: Dashboard MQA-Komponenten des Open.NRW-Portals (Screenshot)

Wichtig ist auch das Feedback von den Daten-  
 nutzerinnen und -nutzern, da sie für die Weiter-  
 verwendung der Daten ggf. andere Qualitäts-  
 anforderungen haben als die Datenberei- telerin-  
 nen und -berei- teler bei deren Veröffentlichung.  
 Daher ist es empfehlenswert, den Austausch mit

den Datennutzerinnen und -nutzern zu suchen und  
 Feedback zu ermöglichen. Dies kann beispielswei-  
 se durch die Angabe einer Kontaktadresse in den  
 Metadaten erfolgen oder durch einen gemein-  
 samen Austausch im Rahmen von Datenwork-  
 shops und Meetups.

## 3.9.2 Wissensmanagement

In diesem Kapitel geht es um die Einführung eines strategischen Wissensmanagements innerhalb der Behörden:

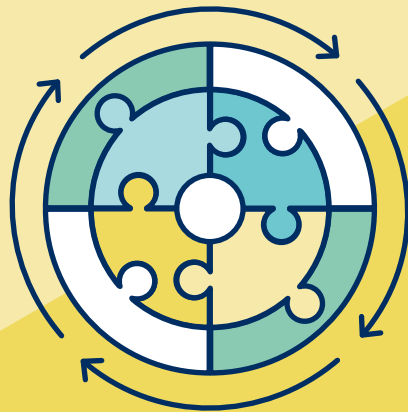
- Aufbau eines Datenmanagements
- Sensibilisierung der Mitarbeitenden ■

Wie die vorangegangenen Kapitel zeigen, gibt es diverse Bausteine, die die Bereitstellung offener Daten ermöglichen und ggf. vereinfachen können. Wie dieser Leitfaden verdeutlicht, ist das Thema Datenmanagement durchaus komplex. Umso wichtiger ist es, innerhalb der Behörde frühzeitig ein strategisches Wissensmanagement einzuführen, um diese Thematik aufzubauen und die Mitarbeitenden entsprechend zu sensibilisieren. Hierbei bietet sich als erster Schritt an, bereits in der Phase der Datenidentifizierung (→ vgl. **Abschnitt 3.1**) die Mitarbeitenden über das Thema Open Data und Datenmanagement zu informieren, beispielsweise über das hauseigene Intranet. Ein Angebot mit Informationsmaterialien zu Open Data seitens der

Beratungsstelle Open Data wird stetig erweitert und ist unter anderem in Form des → **Open Data-Guides** <sup>[2]</sup> auf NRW connect extern und auf dem Open.NRW-Portal zu finden.

Neben der Bereitstellung von Informationen gibt es weitere Formate, wie bspw. die Durchführung von Workshops oder Best-Practice-Dialogen, die etabliert werden können, um das notwendige Wissen rund um Open Data und Datenmanagement bei den Mitarbeitenden aufzubauen und zu erweitern. Bei Unterstützungsbedarf und für die ressortweite Koordinierung stehen die Open Data-Ansprechpartnerinnen und Ansprechpartner der jeweiligen Ressorts zur Verfügung.

# 4. Erweiterung des Datenmanagement- Zyklus



In → Kapitel 3 wurde der → **Datenmanagement-Zyklus** vorgestellt. In diesem sind alle Prozessschritte abgebildet, welche man benötigt, um Nutzerinnen und Nutzern des Open.NRW-Portals die Metadaten zu veröffentlichten Datensätzen zur Verfügung zu stellen.

Für eine Erweiterung des Datenmanagement-Zyklus sollten verschiedene mögliche Startpunkte berücksichtigt werden. Aufgrund des unterschiedlichen Wissensstands beim Umgang mit Metadaten und deren Pflege können einzelne oder auch alle Bausteine des Datenmanagement-Zyklus bei einigen Akteurinnen und Akteuren schon vorhanden sein.

Bei Akteurinnen und Akteuren ohne Vorwissen sollte der Datenmanagement-Zyklus einmal komplett durchlaufen werden. Im Rahmen einer Konzeption sollte bei den verschiedenen Schritten dann überlegt werden, wie eine erweiterte Lösung aussehen könnte. Dieser Leitfaden bietet Handlungsempfehlungen oder zumindest Orientierung, welche Personen an welcher Stelle involviert sein sollten, kann aber ein maßgeschneidertes Datenmanagement-Konzept nicht ersetzen. Bei Akteurinnen und Akteuren, welche nicht über Erfahrung bzgl. der Bereitstellung von Datensätzen in Portalen verfügen, sollten z. B. zunächst grundsätzliche Zuständigkeiten geklärt und notwendige Informationen, bezogen auf Metadatenerfassung, -pflege und -bereitstellung, eingeholt werden (→ siehe **Abbildung 9**).

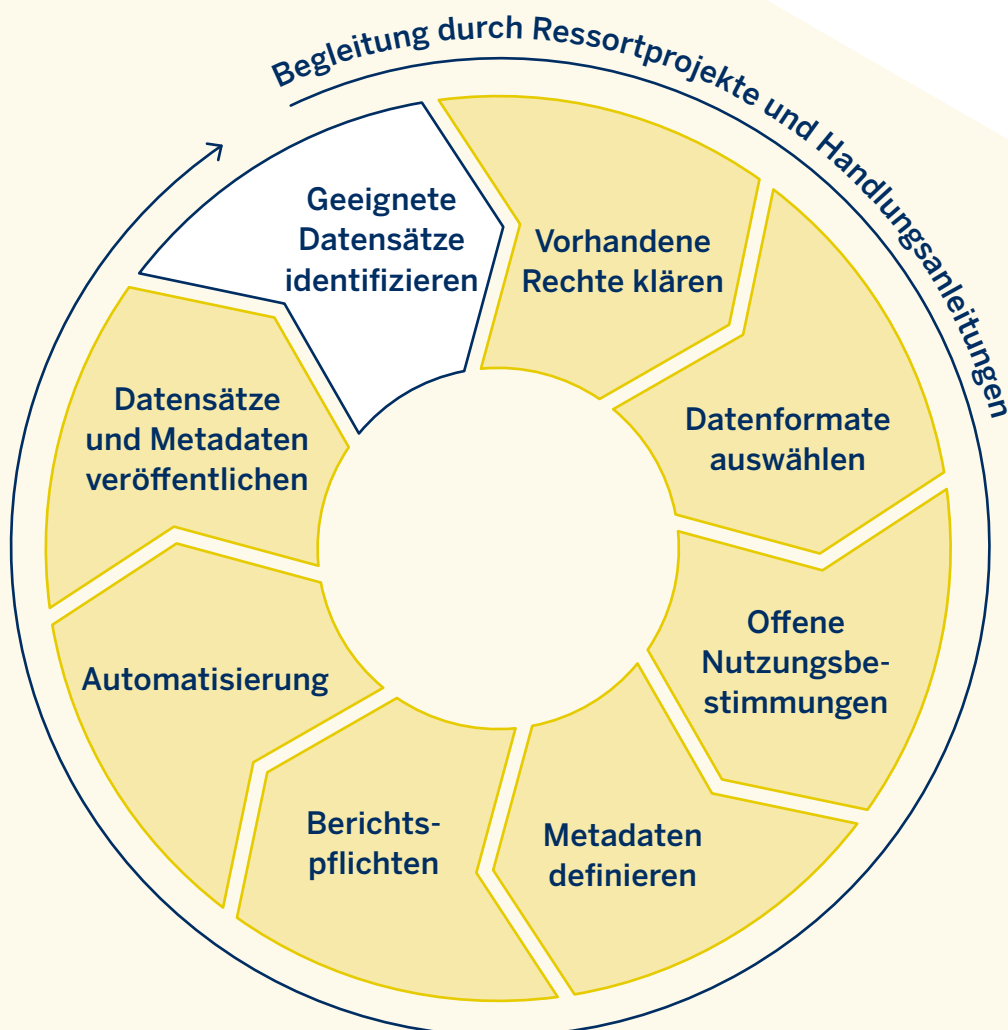


Abbildung 9: Beispielhafter zu durchlaufender Datenmanagement-Zyklus, wenn keine Vorkenntnisse vorhanden sind

Für Akteurinnen und Akteure mit ausreichendem Kenntnisstand, sollte man den Datenmanagement-Zyklus nicht zwangsläufig neu aufsetzen, sondern einzelne Prozessschritte nach Bedarf erweitern oder anpassen. Hierzu wurde bspw. die Identifizierung von Automatisierungspotenzialen (**→ Kapitel 3.7**) genannt, welche als erweiternde Komponente in den Datenmanagement-Zyklus mit integriert werden kann. Dadurch haben Akteurinnen und Akteure die Möglichkeit, wiederkehrende Prozessschritte im Zyklus zu automatisieren.

Die automatisierten Prozesse können bei gleicher Aufgabe auch von mehreren unterschiedlichen Behörden nachgenutzt werden und schaffen dadurch einen nachhaltigen Umgang mit der Metadatenerfassung und -pflege im Open.NRW-Portal.

Eine andere mögliche Erweiterung des Datenmanagement-Zyklus ist eine Bewertung der identifizierten Datensätze in Bezug auf die Erfüllung der Vorgaben aus der Open Data-Verordnung, bevor sie durch Metadaten ins Open.NRW-Portal integriert werden. Bei diesen Daten kann es sich um proprietäre Datenformate handeln, welche von den Behörden als eigenständiges Format genutzt, aber im Open.NRW-Portal nicht unterstützt werden. Hier wäre es sinnvoll, eine Datenaufbereitung (z. B. Transformation des proprietären Datenformats in ein nicht proprietäres) durchzuführen, um das Datenformat in eines der vorgeschriebenen und unterstützten Formate zu überführen. Mit dieser Erweiterung im Datenmanagement-Zyklus können im Vorfeld nicht unterstützte Datenformate im Open.NRW-Portal vermieden werden. In der **→ Abbildung 10** ist beispielhaft der erweiterte Datenmanagement-Zyklus abgebildet. Dabei sind die erweiterten Inhalte im Datenmanagement-Zyklus grün dargestellt.



Abbildung 10: Beispielhafte Erweiterung des Datenmanagement-Zyklus, wenn Vorkenntnisse vorhanden sind

In beiden Fällen der beispielhaften Erweiterung des Datenmanagement-Zyklus sollte innerhalb jeder Behörde eine Überprüfung auf den Wissensstand bzgl. der Erfassung und Pflege von Metadaten durchgeführt werden. Dadurch kann man heraus-

finden, welche Prozessschritte für die Metadaterfassung und -pflege im Open.NRW-Portal noch nicht vorhanden sind, und diese individuell im Datenmanagement-Zyklus erweitern.

# 5.

## Quellenverzeichnis

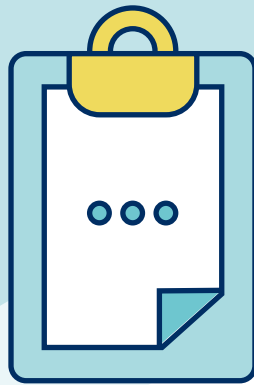


- [1] Nordrhein-Westfalen, IM, „Gesetz- und Verordnungsblatt (GV. Nordrhein-Westfalen.)“, 2022. Online verfügbar: [https://recht.nrw.de/lmi/owa/br\\_vbl\\_detail\\_text?anw\\_nr=6&vd\\_id=20121&ver=8&val=20121&sg=0&menu=0&vd\\_back=N](https://recht.nrw.de/lmi/owa/br_vbl_detail_text?anw_nr=6&vd_id=20121&ver=8&val=20121&sg=0&menu=0&vd_back=N). (Zugriff am 20.01.2022.)
- [2] Open.NRW, „Open Data-Guide Grundlagen“, 2022. Online verfügbar: [https://open.nrw/system/files/media/document/file/open\\_data\\_guide\\_grundlagen.pdf](https://open.nrw/system/files/media/document/file/open_data_guide_grundlagen.pdf).



- [3] Open.NRW, „Checkliste zur Prüfung der Verpflichtung aus § 16a EGovG NRW“, 2022. Online verfügbar: [https://open.nrw/system/files/media/document/file/open\\_data\\_checkliste.pdf](https://open.nrw/system/files/media/document/file/open_data_checkliste.pdf).
- [4] Findeisen, Matthias; Kim, James G.; Hausenblas, Michael, „5-Sterne-Modell nach Tim Berners-Lee“, 31.08.2015. Online verfügbar: <https://5stardata.info/de/>. (Zugriff am 22.12.2021.)
- [5] GovData, Geschäfts- und Koordinierungsstelle, „DCAT-AP.de“. Online verfügbar: <https://www.dcat-ap.de/>. (Zugriff am 23.12.2021.)
- [6] GovData, „DCAT-AP.de Spezifikation Version 1.1“, 08.10.2020. Online verfügbar: <https://www.dcat-ap.de/def/dcatde/1.1/spec/specification.pdf>. (Zugriff am 23.12.2021.)
- [7] Ministerium für Heimat, Kommunales, Bau und Digitalisierung des Landes Nordrhein-Westfalen, „Leitfaden zur Metadatenerfassung“, 2022. Online verfügbar: [https://open.nrw/system/files/media/document/file/leitfaden\\_zur\\_metadatenerfassung.pdf](https://open.nrw/system/files/media/document/file/leitfaden_zur_metadatenerfassung.pdf).
- [8] Ministerium für Heimat, Kommunales, Bau und Digitalisierung des Landes Nordrhein-Westfalen, „Leitfaden zur Registrierung“, 2022. Online verfügbar: [https://open.nrw/system/files/media/document/file/leitfaden\\_zur\\_registrierung.pdf](https://open.nrw/system/files/media/document/file/leitfaden_zur_registrierung.pdf).
- [9] Ministerium für Heimat, Kommunales, Bau und Digitalisierung des Landes Nordrhein-Westfalen, „E-Government-Strategie und E-Government-Gesetz“, 2022. Online verfügbar: <https://www.mhkbd.nrw/themen/digitalisierung/die-e-government-strategie-und-das-e-government-gesetz-nrw>. (Zugriff am 10.02.2022.)
- [10] „RICHTLINIE 2008/50/EG DES EUROPÄISCHEN PARLAMENTS UND DES RATES vom 21. Mai 2008 über Luftqualität und saubere Luft für Europa“, 29.05.2018. Online verfügbar: [https://igsvtu.lanuv.nrw.de/vtu/doc.app?P\\_VTU\\_SYSID=002-31&DATEI=6/dokus/60034.pdf](https://igsvtu.lanuv.nrw.de/vtu/doc.app?P_VTU_SYSID=002-31&DATEI=6/dokus/60034.pdf). (Zugriff am 21.02.2022.)
- [11] Nordrhein-Westfalen, LANUV, „Bericht über die Luftqualität im Jahr 2020“, 02.03.2021. Online verfügbar: [https://www.lanuv.nrw.de/fileadmin/lanuv/luft/immissionen/ber\\_trend/Bericht\\_ueber\\_die\\_Luftqualitaet\\_im\\_Jahr\\_2020.pdf](https://www.lanuv.nrw.de/fileadmin/lanuv/luft/immissionen/ber_trend/Bericht_ueber_die_Luftqualitaet_im_Jahr_2020.pdf). (Zugriff am 21.02.2022.)
- [12] Ministerium für Heimat, Kommunales, Bau und Digitalisierung des Landes Nordrhein-Westfalen, „Open.NRW-Portal“, 2022. Online verfügbar: <https://open.nrw/>. (Zugriff am 18.02.2022.)
- [13] Das Fraunhofer-Institut für Offene Kommunikationssysteme FOKUS, Berlin, „Leitfaden für qualitativ hochwertige Daten und Metadaten“, 2019. Online verfügbar: <https://www.nqdm-projekt.de/de/downloads/leitfaden>. (Zugriff am 19.01.2022.)
- [14] Ministerium für Wirtschaft, Innovation, Digitalisierung und Energie des Landes Nordrhein-Westfalen, „MQA2 – Metadata Quality Assessment Methodologie“, 2021. Online verfügbar: <https://mqa2-open.nrw.de/mqa/methodology?locale=de>. (Zugriff am 05.05.2022.)

# 6. Anhang



## Glossar

### **API**

Eine API (englisch für Application Programming Interface) ist eine Programmierschnittstelle. Sie wird von einem Softwaresystem bereitgestellt und ermöglicht es anderen Programmen, mit diesem Softwaresystem zu kommunizieren. APIs werden oftmals von Datenherausgeberinnen und Datenherausgebern bereitgestellt und ermöglichen es Programmen oder Apps, die Daten direkt über das Web zu lesen. Vorteilhaft an der Bereitstellung der Daten per API ist, dass nicht der gesamte Datensatz heruntergeladen werden muss, da die API lediglich die benötigten Daten bereitstellt.

### **CSV**

CSV (Comma-separated Values) ist ein Standardformat für strukturierte Daten. Aufgrund der Einfachheit, Offenheit und Maschinenlesbarkeit wird CSV häufig für die Veröffentlichung von offenen Daten verwendet.

### **Daten**

Daten beschreiben originale Erzeugnisse, welche in Form von Messungen, Beobachtungen oder statistischen Auswertungen meist selbst erhoben werden.

## Datenbank

In einer Datenbank werden die originalen Daten sowie deren Metadaten strukturiert zur elektronischen Datenverwaltung gespeichert.

## Datenformate

Datenformate geben an, wie Daten gespeichert werden können. Sie werden kenntlich gemacht durch die Dateiendung (z. B. .xml).

## Datennutzerinnen und -nutzer

Als Datennutzerinnen und -nutzer werden diejenigen natürlichen oder juristischen Personen bezeichnet, welche die von Datenbereiterinnen und -bereiter zur Verfügung gestellten Daten gemäß den vorgesehenen Nutzungsbestimmungen für ihre eigenen Zwecke verwenden.

## Datensatz

Bei einem Datensatz handelt es sich um eine Sammlung von Daten. Die Daten können vom gleichen oder unterschiedlichen Typ sein.

## DCAT-AP.de

DCAT-AP.de ist der deutsche Metadatenstandard für offene Verwaltungsdaten. Als gemeinsames deutsches Metadatenmodell zum Austausch von offenen Verwaltungsdaten wird DCAT-AP.de von GovData spezifiziert und ist seit Ende 2017 bei GovData im Einsatz. DCAT-AP.de ist als eine Ableitung vollständig kompatibel mit dem europäischen Standard DCAT-AP.

## GovData

GovData ist das nationale Datenportal für Deutschland und bietet eine zentrale Anlaufstelle für Daten aus allen Verwaltungsebenen. Das Portal kann unter folgender URL aufgerufen werden: <https://www.govdata.de/>.

## JSON

JSON ist ein leistungsfähiges Format, welches gut geeignet ist für den Datenaustausch zwischen verschiedenen Anwendungen. Es kann komplexe Datenstrukturen beschreiben, ist sowohl für Menschen als auch für Maschinen gut lesbar und unabhängig von Plattform und Programmiersprache.

## Lizenz

Eine Lizenz beschreibt Nutzungsbestimmungen für die Verwendung eines Werkes. Über eine Lizenz räumt die Urheberin bzw. der Urheber des Werkes der Lizenznehmerin bzw. dem Lizenznehmer gewisse Nutzungsrechte zu bestimmten Nutzungsbedingungen ein. Standardlizenzverträge, wie z. B. Creative Commons oder Datenlizenz Deutschland 2.0, erleichtern der Urheberin bzw. dem Urheber die Festlegung der Nutzungsbestimmungen.

## Metadaten

Die Metadaten sind beschreibende Informationen zu den originalen Daten. Über diese können die Daten im Internet zugänglich gemacht werden.

## Metadatenimport

Beim Metadatenimport werden die Metadaten aus den Quelldaten abgerufen und in das Datenbanksystem des Open.NRW-Metadatensystems kopiert. Dabei werden die Metadaten unverändert in das Datenbanksystem kopiert.

## Metadatensystem

Das Metadatensystem speichert die Metadaten der originalen Daten im System des Open.NRW-Portals. Die Speicherung erfolgt dabei in einem Datenbanksystem.

## RDF

Das Resource Description Framework ist ein Modell zur Speicherung von Daten und Metadaten. Darin werden miteinander verknüpfte Daten in Form von Tripeln gespeichert.

## Spezifikation (DCAT-AP.de)

In der Spezifikation zum DCAT-AP.de-Metadatenmodell werden die semantischen Regelungen beschrieben, welche im Open.NRW-Portal als Basis für die Metadatenerfassung verwendet werden.

## XML

XML (Extensible Markup Language) ist ein Format zur hierarchischen Strukturierung von Daten. XML wurde entwickelt, um plattformübergreifend Daten austauschen zu können.

# Impressum

## Herausgeber

Ministerium für Heimat, Kommunales,  
Bau und Digitalisierung  
des Landes Nordrhein-Westfalen  
Referat „Reden, Publikationen“

Jürgensplatz 1  
40219 Düsseldorf  
Telefon: 0211 8618-50  
info@mhkbd.nrw.de  
www.mhkbd.nrw

## Kontakt

Beauftragter der Landesregierung  
Nordrhein-Westfalen  
für Informationstechnik (CIO)

Beratungsstelle Open Data  
open-data@open.nrw.de

Das Projekt wurde im Rahmen  
von Open.NRW umgesetzt.  
www.open.nrw

## Konzept, Inhalt und Redaktion

con terra GmbH  
www.conterra.de

Fraunhofer FOKUS  
www.fokus.fraunhofer.de

## Design

made in  
www.madein.io



### **Stand**

Dezember 2022

### **Lizenz**

Creative Commons Namensnennung 4.0  
International (CC BY 4.0)

### **© 2022/MHKBD D-503**

Die Publikation steht zum Download bereit  
unter: [www.mhkbd.nrw/broschueren](http://www.mhkbd.nrw/broschueren)

Diese Druckschrift wird im Rahmen der Öffentlichkeitsarbeit der Landesregierung Nordrhein-Westfalen herausgegeben. Sie darf weder von Parteien noch von Wahlwerberinnen und -werbern oder Wahlhelferinnen und -helfern während eines Wahlkampfes zum Zwecke der Wahlwerbung verwendet werden. Dies gilt für Landtags-, Bundestags- und Kommunalwahlen sowie auch für die Wahl der Mitglieder des Europäischen Parlaments.

Missbräuchlich ist insbesondere die Verteilung auf Wahlveranstaltungen, an Informationsständen der Parteien sowie das Einlegen, Aufdrucken oder Aufkleben parteipolitischer Informationen oder Werbemittel. Untersagt ist gleichfalls die Weitergabe an Dritte zum Zwecke der Wahlwerbung.

Eine Verwendung dieser Druckschrift durch Parteien oder sie unterstützende Organisationen ausschließlich zur Unterrichtung ihrer eigenen Mitglieder bleibt hiervon unberührt. Unabhängig davon, wann, auf welchem Wege und in welcher Anzahl diese Schrift der Empfängerin oder dem Empfänger zugegangen ist, darf sie auch ohne zeitlichen Bezug zu einer bevorstehenden Wahl nicht in einer Weise verwendet werden, die als Parteinahme der Landesregierung zu Gunsten einzelner politischer Gruppen verstanden werden könnte.

Ministerium für Heimat, Kommunales,  
Bau und Digitalisierung  
des Landes Nordrhein-Westfalen



## Ministerium für Heimat, Kommunales, Bau und Digitalisierung des Landes Nordrhein-Westfalen

Jürgensplatz 1  
40219 Düsseldorf  
Telefon: 0211 8618-50  
Telefax: 0211 8618-54444

info@mhkbd.nrw.de  
www.mhkbd.nrw

-  **MHKBD\_NRW**
-  **MHKBD.NRW**
-  **mhkbd\_nrw**
-  **MHKBD\_NRW**