



Rheinland-Pfalz

MINISTERIUM FÜR ARBEIT,
SOZIALES, TRANSFORMATION
UND DIGITALISIERUNG



HANDBUCH FÜR DIE BEREITSTELLUNG VON DATEN

OPEN DATA RHEINLAND-PFALZ

Inhalt

Einleitung	04
Bereitstellung von Open Data	06
Welchen Nutzen haben offene Daten?	07
Was sind die Open-Data-Prinzipien?	09
Hybride Bedrohungen – Open Data und Transparenz als Risiko oder Chance?	10
5-Sterne-Modell als Benchmark für offene Daten	12
5-Sterne offene Daten	12
Von Rohinformationen zu offenen Daten mit 5 Sternen	14
Open Data und Rechtsvorschriften	16
Transparenz- und Offene-Daten-Plattform des Landes Rheinland-Pfalz	23
Die Metadatenföderation GovData	26
Aufbau der Plattform	26
Angebundene Liefersysteme	28
Grundlagen zur Vorbereitung von Daten und Metadaten	29
Daten anonymisieren	30
Metadaten erstellen	32
Lizenzauswahl	33
Formatauswahl	34
Veröffentlichung von Metadaten auf open.rlp.de	35
Harvesting: Abholung von Metadaten	36
API: Anlieferung von Metadaten	38
Einstellsystem	38
Weitere Hinweise	39
Bildnachweise	40
Impressum	41

Einleitung

Die „Transparenz- und Offene-Daten-Plattform des Landes Rheinland-Pfalz“ (open.rlp.de) bietet als Metadatenplattform mit nutzerfreundlicher Suchfunktion einen zentralen Zugang zu Datensätzen und Dokumenten aus den rheinland-pfälzischen Verwaltungen, einschließlich der Kommunen.

Die Datensätze bestehen jeweils aus den Daten selbst und standardisierten Metadaten. Auf open.rlp.de werden die Metadaten durchsuchbar zur Verfügung gestellt. Die eigentlichen Daten verbleiben beim jeweiligen Bereitsteller und werden verlinkt. Dadurch, dass ausschließlich die Metadaten weitergegeben werden, ist eine kaskadierende und gleichzeitig schlanke Bereitstellung auf Länder-, Bundes- und europäischer Ebene gewährleistet. Denn open.rlp.de leitet die Metadaten der rheinland-pfälzischen offenen Daten an das Bundesportal weiter, das wiederum die gesammelten Daten aus Deutschland an das europäische Portal weiterleitet.

Offene Daten sind frei nutzbar. Sie dürfen vervielfältigt, verändert und in eigene Anwendungen eingebunden werden.

Open Data in Rheinland-Pfalz

Zur Erfüllung ihrer öffentlich-rechtlichen Aufgaben erheben die Behörden des Landes eine Fülle von unterschiedlichsten Daten, die nicht sensibel oder personenbezogen sind. Sie verfügen damit über einen enormen Datenpool, der auch über die Grenzen der jeweiligen Behörde hinaus nutzbar gemacht werden kann – sei es als Grundlage für Beteiligung, Wissen und Innovationen,

zur Ermöglichung neuer Geschäftsmodelle oder zur Förderung von Transparenz und Offenheit der Verwaltung. Um möglichst vielen Bürgerinnen und Bürgern, Unternehmen, der Wissenschaft, sozialen Einrichtungen und Initiativen sowie Einrichtungen der öffentlichen Verwaltung entgeltfrei Zugang zu diesen Daten zur freien Nutzung zu ermöglichen, sollte die Bereitstellung dieser Daten durch die Behörden des Landes als Open Data standardisiert erfolgen und signifikant erhöht werden.

Seit der Veröffentlichung des ersten Open-Government-Data-Portals in Rheinland-Pfalz am 5. März 2013 hat das Angebot an offenen Verwaltungsdaten (vor allem aus dem Bereich Geodaten) kontinuierlich zugenommen.

Auf der „Transparenz-Plattform“ (TPP), die seit dem 1. Januar 2016 in Betrieb ist, werden durch die Landesbehörden, die Kommunen, sowie weitere transparenzpflichtige Stellen diejenigen Informationsgegenstände veröffentlicht, für die eine Veröffentlichungspflicht bzw. eine freiwillige Veröffentlichungsmöglichkeit nach dem Landestransparenzgesetz vorgesehen sind.

Diese beiden vorhandenen Daten- und Informationsplattformen wurden inhaltlich weiter ausgebaut und zu einem zentralen, nutzungsfreundlichen Angebot zusammengeführt, das auch in das nationale Metadatenportal „GovData“ einspeist. Auch Kommunen sowie die sonstige mittelbare Landesverwaltung und weitere Akteure sollen im Kontext Open (Government) Data die Möglichkeit sowie Anreize erhalten, of-

fene Daten über das Landesportal auffindbar zu machen.

Wir entwickeln die Funktionen, die Oberfläche und die Datenbasis unseres Portals ständig weiter. Es lohnt sich, regelmäßig vorbeizuschauen!

An wen richtet sich dieses Handbuch?

Dieses Handbuch richtet sich an Personen aus der rheinland-pfälzischen Verwaltung, die sich für das Thema Open Data interessieren und offene Daten veröffentlichen möchten.

Wir freuen uns, wenn Sie das Handbuch nutzen, um Ihre Daten zu veröffentlichen und uns dabei unterstützen, unsere Plattform mit interessanten Daten zu füllen und so mehr offene Daten für unser Land nutzbar zu machen.

Was ist das Ziel des Handbuchs?

Nach der Lektüre dieses Handbuchs

- haben Sie ein besseres Verständnis von Open Data,
- kennen Sie die rechtlichen Grundlagen des Open-Data- und Transparenz-Prinzips in Rheinland-Pfalz,
- wissen Sie, was notwendig ist, um Ihre Offenen Daten auf open.rlp.de zu veröffentlichen,
- kennen Sie wichtige Begriffe rund um Open Data und können dazu im eigenen Verantwortungsbereich als Multiplikatorin oder Multiplikator einen wichtigen Beitrag leisten.

Wie ist dieses Handbuch aufgebaut?

In **Kapitel 1** werden grundlegende Informa-

tionen zum Begriff „Open Data“, den wirtschaftlichen und gesellschaftlichen Potenzialen, der praktischen Umsetzung von Open Data sowie der rechtlichen Verankerung des Open-Data-Prinzips bereitgestellt.

Kapitel 2 erläutert die zentrale technische Rolle, die open.rlp.de bei der Veröffentlichung von offenen Daten in Rheinland-Pfalz und darüber hinaus ausfüllt.

In **Kapitel 3** werden die zuvor dargestellten Grundlagen zur Veröffentlichung eines Datensatzes als Open Data um praktische Empfehlungen ergänzt.

Kapitel 4 stellt dar, wie eigene Daten und Metadatenkataloge auf open.rlp.de veröffentlicht werden können und welche Anforderungen mit einer Anbindung an die Plattform verbunden sind.

Kapitel 5 gibt einen kurzen Überblick über das Informationsangebot des Landes und führt Möglichkeiten auf, mit den Verantwortlichen der Plattform in Kontakt zu treten.



BEREITSTELLUNG
VON OPEN DATA

Im Rahmen des Projekts „Open Data“ ging open.rlp.de als Konsolidierung der bisherigen „Transparenzplattform“ (tpp.rlp.de) und des „Open-Government-Data-Portals“ (daten.rlp.de) im September 2024 online. Damit ist open.rlp.de die zentrale Plattform des Landes Rheinland-Pfalz für Transparenzinformationen und offene Daten.

Die neue Transparenz- und Offene-Daten-Plattform bietet Besucherinnen und Besuchern wichtige Informationen über die Verwaltung unseres Bundeslandes und ge-

währt Zugriff auf Daten, die kostenlos genutzt werden können. So ist open.rlp.de ein einheitlicher zentraler Zugang zu Verwaltungsdaten aus den rheinland-pfälzischen Landesbehörden sowie den rheinland-pfälzischen Kommunen.

Ziel ist es, diese Daten über einen zentralen Metadatenkatalog auffindbar zu machen. Im Sinne von Open Data wird außerdem angestrebt, die Verwendung offener Lizenzen zu fördern und das Angebot von maschinenlesbaren Rohdaten zu erhöhen.

Welchen Nutzen haben offene Daten?

Offene Daten sind der Treibstoff des digitalen Wandels. Sie nehmen eine Schlüsselfunktion für zentrale Aufgaben wie Forschung und Innovation, Transparenz und Bürgerbeteiligung oder Planung und Zusammenarbeit ein.

Die Bereitstellung von offenen Daten fördert die Transparenz und Nachvollziehbarkeit des Regierungshandelns. Damit einhergehend ermöglichen offene Daten die Teilhabe an der (digitalen) Gesellschaft und stärken das Vertrauen in die staatlichen Institutionen. Nicht zuletzt schaffen sie eine faktenbasierte Grundlage für die aktive Beteiligung am gesellschaftlichen Diskurs und unterstützen die Entwicklung von innovativen Lösungen gesellschaftlicher Herausforderungen, wie etwa dem demografischen Wandel.

Offene Daten können daher auch genutzt werden, um Desinformationen und Falsch-

meldungen im politisch-gesellschaftlichen Raum proaktiv zu bekämpfen, indem Sachverhalte verlässlich und fundiert anhand von Primärquellen dargestellt werden. Dies stärkt den (zivil-)gesellschaftlichen Zusammenhalt und die Demokratie.

Ihr Beitrag zur Gewinnung neuer wissenschaftlicher Erkenntnisse und der Entwicklung neuer Geschäftsmodelle macht offene Daten zu exzellenten Treibern von Forschung und Innovation. Im Bereich der Umweltforschung können offene Daten beispielsweise dazu beitragen, den Einfluss menschlicher Aktivitäten auf die Umwelt besser zu verstehen. Szenarioanalysen mit realen Daten ermöglichen die Entwicklung nachhaltiger Umweltschutzmaßnahmen sowie die Simulation ihrer Auswirkungen. Die Analyse von Verkehrsdaten und die Entwicklung datenbasierter Verkehrsüberwachungsanwendungen verbessern die Verkehrssteuerung und reduzieren Staus.

Offene Daten schaffen neue Geschäftsmodelle. Sie steigern die Effizienz von Unternehmen und tragen zur Entwicklung besserer Produkte bei. Aktuelle Mobilitätsdaten ermöglichen beispielsweise Logistikunternehmen eine bessere Routenplanung und aktuelle Klimadaten können für die Optimierung der Ernte herangezogen werden.

Gerade für Start-ups und kleine spezialisierte Unternehmen bieten offene Daten ein großes Potenzial. Einer [Studie der EU-Kommission](#) aus dem Jahr 2020 zufolge wird der Marktwert offener Daten im Jahr 2025 199,5 Mrd. Euro betragen. Offene Daten sind also auch Treiber des wirtschaftlichen Wachstums.

Nicht zuletzt für die Verwaltung selbst bietet die Veröffentlichung ihrer Informationen in Form von offenen Daten ebenfalls einen deutlichen Mehrwert. Je größer die Anzahl offen verfügbarer Daten, desto einfacher ist es, innerbehördliche Wissenssilos

aufzubrechen und redundante Arbeiten zu vermeiden. Das Verwaltungshandeln wird außerdem effektiver gestaltet, wenn zusätzliche Informationen aus unterschiedlichen Quellen berücksichtigt werden. Dies geschieht bei kontextualisierten offenen Daten, sogenannten „Linked Open Data“.

Über open.rlp.de können auch Bürgerinnen und Bürger in ihrem Alltag direkt von dem Angebot offener Verwaltungsdaten profitieren. Sie können zum Beispiel auf Informationen wie Bebauungspläne von Wohn-, Industrie- und Gewerbegebieten oder veröffentlichte Dokumente der Gemeinden zu einer Vielzahl von Themen zugreifen. Wir erweitern den Bestand an Informationen auf open.rlp.de stetig, indem wir immer neue Datenquellen an open.rlp.de anbinden – mit dem Ziel, Rheinland-Pfalz als Standort für Unternehmen sowie Privatpersonen attraktiver zu machen.



Was sind die Open-Data-Prinzipien?

Mitarbeiterinnen und Mitarbeiter in der öffentlich-rechtlichen Verwaltung erheben und generieren tagtäglich eine Fülle von Informationen zur Erfüllung ihrer dienstlichen Aufgaben. Diese Informationen stellen eine wertvolle Ressource dar und können, wenn sie nicht personenbezogen oder aus anderen Gründen schützenswert sind, in der Regel frei zugänglich und ohne Einschränkungen zur freien Verwendung und Verbreitung veröffentlicht werden.

Man spricht in diesem Kontext von „Open Data“ bzw. „Open Government Data“ (offene Verwaltungsdaten). Die zehn Open-Data-Prinzipien definieren dabei Kriterien, die den Grad der Verfügbarkeit und Nachnutzbarkeit messbar machen.

Diese zehn Prinzipien sollen alle offenen Daten erfüllen:

1. Vollständigkeit

Die Daten zu bestimmten Themen werden so vollständig wie möglich veröffentlicht. Die Daten sind mit Metadaten versehen, die sie standardisiert beschreiben.

2. Primärquelle

Die Daten werden mit dem höchstmöglichen Feinheitsgrad gesammelt und veröffentlicht und nicht in aggregierten oder modifizierten Formaten.

3. Aktualität der Daten

Die Daten sollen veröffentlicht werden, sobald sie erhoben und zusammengestellt wurden.

4. Zugänglichkeit

Die Informationen sollen ohne physische und technische Hürden barrierefrei zugänglich sein. Programmierschnittstellen (APIs), die das Herunterladen sämtlicher Daten als Gesamtpaket ermöglichen, sowie eine einfache Suche und leichte Auffindbarkeit der Daten unterstützen die Zugänglichkeit.

5. Maschinenlesbarkeit

Es werden Datenformate verwendet, die von einem Computer weiterverarbeitet werden können. Während z.B. eine PDF-Datei für die Nutzerinnen und Nutzer gut lesbar ist, bleibt sie für eine automatische Weiterverwendung schwer zugänglich. Als gut maschinenlesbar gelten u.a. die folgenden Formate: .txt, .csv, .json, .xml, .rss.

6. Diskriminierungsfreiheit

Nach diesem Prinzip hat jede Person zu jedem Zeitpunkt Zugriff auf die Daten, ohne sich dabei identifizieren oder eine Rechtfertigung für ihr Handeln erbringen zu müssen.

7. Offene Standards

Die gespeicherten Daten müssen so zugänglich und lesbar sein, dass keine Notwendigkeit zur Anschaffung einer bestimmten Software und damit verbundene Kosten besteht.

8. Lizenzierung

Maximale Offenheit beinhaltet, dass öffentliche Informationen eindeutig als Werk der Verwaltung ausgewiesen und für die Allgemeinheit ohne Restriktionen oder Auferlegung von Nutzungsbedingungen, die sich als Barrieren auswirken können, zugänglich sind.

9. Dauerhaftigkeit

Die einmal online gestellten Informationen bleiben mit einer angemessenen Dokumentation der Versionen online und sind über lange Zeit hinweg verfügbar.

10. Nutzungskosten

Selbst die Auferlegung von symbolischen Nutzungsgebühren stellt ein Hindernis dar. Die Erhebung von Gebühren kann die Nutzung von Daten und damit mögliches Wirtschaftswachstum, Steuereinnahmen oder die Schaffung eines sozialen Mehrwerts verhindern.

Hybride Bedrohungen – Open Data und Transparenz als Risiko oder Chance?

In modernen Konfliktszenarien haben sich gezielte Falschinformationen und -meldungen in den Medien sowie Beeinflussung des Meinungsklimas und Wahlbeeinflussung in liberalen Demokratien und offenen Gesellschaften als ernste hybride Bedrohungen erwiesen. Als positives Potenzial offener Daten wurde in den vorangegangenen Abschnitten aufgezeigt, dass sie ein Mittel sein können, um Falschmeldungen durch eine Offenlegung der Primärquellen zu entlarven. Bürgerinnen und Bürger haben durch offene Daten die Möglichkeit, selbst Einblick in die Prozesse und das Handeln von Regierung und Verwaltung zu nehmen.

Die Verfügbarkeit und erleichterte Auffindbarkeit von Informationen stellt jedoch im Umkehrschluss ebenfalls eine Erleichterung für staatliche und nichtstaatliche Akteure mit Angriffsabsichten dar, da das Informationsangebot auch für Desinformation und Falschmeldungen instrumentalisiert werden kann. Die Verfügbarkeit offener Schnittstellen ermöglicht zudem systematische Analysen zur Identifizierung von Schwachstellen und Angriffsvektoren.

Durch die Schaffung und den Ausbau eines Open-Data-Ökosystems in Rheinland-Pfalz und die damit verbundene gemeinsame Nutzung technischer und organisatorischer Infrastrukturen wird also möglicherweise eine größere Angriffsfläche geboten.

Den genannten Risiken sollten wir jedoch nicht mit Einschränkungen der Verfügbarkeit offener Daten begegnen, um uns ihrer Vorteile nicht zu berauben. Vielmehr müssen wir diesen Risiken mit konkreten

Anforderungen in Bezug auf IT-Sicherheit und organisatorischen Maßnahmen begegnen. Im Rahmen einer Data-Governance muss beispielsweise sichergestellt werden, dass Richtlinien und sichere Methoden implementiert werden, um vor der Aggregation und Veröffentlichung von Daten eine Anonymisierung vorzunehmen.

Es obliegt den veröffentlichenden Stellen, zu bewerten, ob es sich bei Daten um sicherheitsrelevante Informationen handelt. Nur so kann eine Veröffentlichung von personenbezogenen oder anderweitig schützenswerten Daten ausgeschlossen werden. Bei der Auswahl der Lizenzierung und den Nutzungsbedingungen sollte außerdem darauf geachtet werden, dass sie Haftungsausschlüsse und Hinweise auf eine mögliche Fehlinterpretation der Daten enthalten. Kapitel 3 dieses Handbuchs liefert weitere Hinweise zur praktischen Umsetzung dieser Themen.

Bei der Planung, Konzeption und Umsetzung von open.rlp.de wurden die Themenfelder Sicherheit und Datenschutz von Beginn an mitbedacht, sodass nach dem Grundsatz „Security und Privacy by Design“ gearbeitet wurde. Die Infrastruktur der Plattform ist in einer sicheren und resilienten Umgebung integriert. Zusätzlich wurden Maßnahmen ergriffen, um eine Robustheit gegen Angriffe (z.B. DDoS) und erhöhte Zugriffszahlen zu gewährleisten. Letzteres wird auch durch ein erhöhtes Aufkommen verschiedener Bots bedingt.

Die „CKAN-API“ von open.rlp.de, die einen automatisierten Abruf der Daten ermög-

licht, wurde auf die wesentlichen Funktionen eingeschränkt, um einen Abfluss von schützenswerten Informationen zu verhindern. Zusätzlich wurden Rate-Limits für den Abruf und die Anlieferung der Daten über die Harvester implementiert, um die Systemlast besser kontrollieren zu können.

Bei Bedarf werden Unregelmäßigkeiten im Datenbestand durch die Plattform-Verantwortlichen an die Liefersysteme zurückgemeldet. Zusätzlich existieren verschiedene Infomaterialien, über die sich die Liefersysteme zum Thema der Datenbereitstellung informieren können.

Beim Angebot von open.rlp.de handelt es sich lediglich um einen Metadatenkatalog. Die eigentlichen Transparenzinformationen

und offenen Daten liegen weiterhin bei den veröffentlichenden Stellen oder Fachportalen. Dadurch besteht im Falle eines Ausfalls der Plattform weiterhin eine hohe Verfügbarkeit der ursprünglichen dezentralen Systeme und der Abruf der Informationen ist weiterhin gewährleistet.

Mit Blick auf moderne hybride Bedrohungen ist ein sicherheitsbewusster und verantwortungsvoller Umgang mit offenen Daten notwendig. Durch die Umsetzung konkreter Maßnahmen, Technologien und eines angemessenen Risikomanagements wird die Verwaltung befähigt, den Bürgerinnen und Bürgern Transparenz zu bieten, und dabei gleichzeitig die eigene Sicherheit im Informationsraum zu gewährleisten.



5-Sterne-Modell als Benchmark für offene Daten

Das „5-Sterne-Modell“ von Tim Berners-Lee beschreibt das Open-Data-Prinzip aus der Perspektive der technischen Nutzbarkeit und Interoperabilität. Das Modell definiert fünf Qualitätsstufen für die Bereitstellung offener Daten und ermöglicht dabei eine Bewertung des Nutzens der bereitgestellten Daten.



5-Sterne offene Daten



Schaubild: „5-Sterne-Modell“ nach Sir Berners-Lee

Stufe 1: Die Veröffentlichung von Daten erfolgt unter einer offenen Lizenz. Das Format, das verwendet wird, ist dabei unerheblich.

Stufe 2: Zusätzlich wird die Anforderung eines strukturierten Formats erhoben, bspw. in Form einer Excel-Tabelle anstelle eines Bildes.

Stufe 3: Die Veröffentlichung der Daten muss zusätzlich in einem offenen, nicht proprietären Format erfolgen, z.B. als CSV anstelle einer Excel-Tabelle.

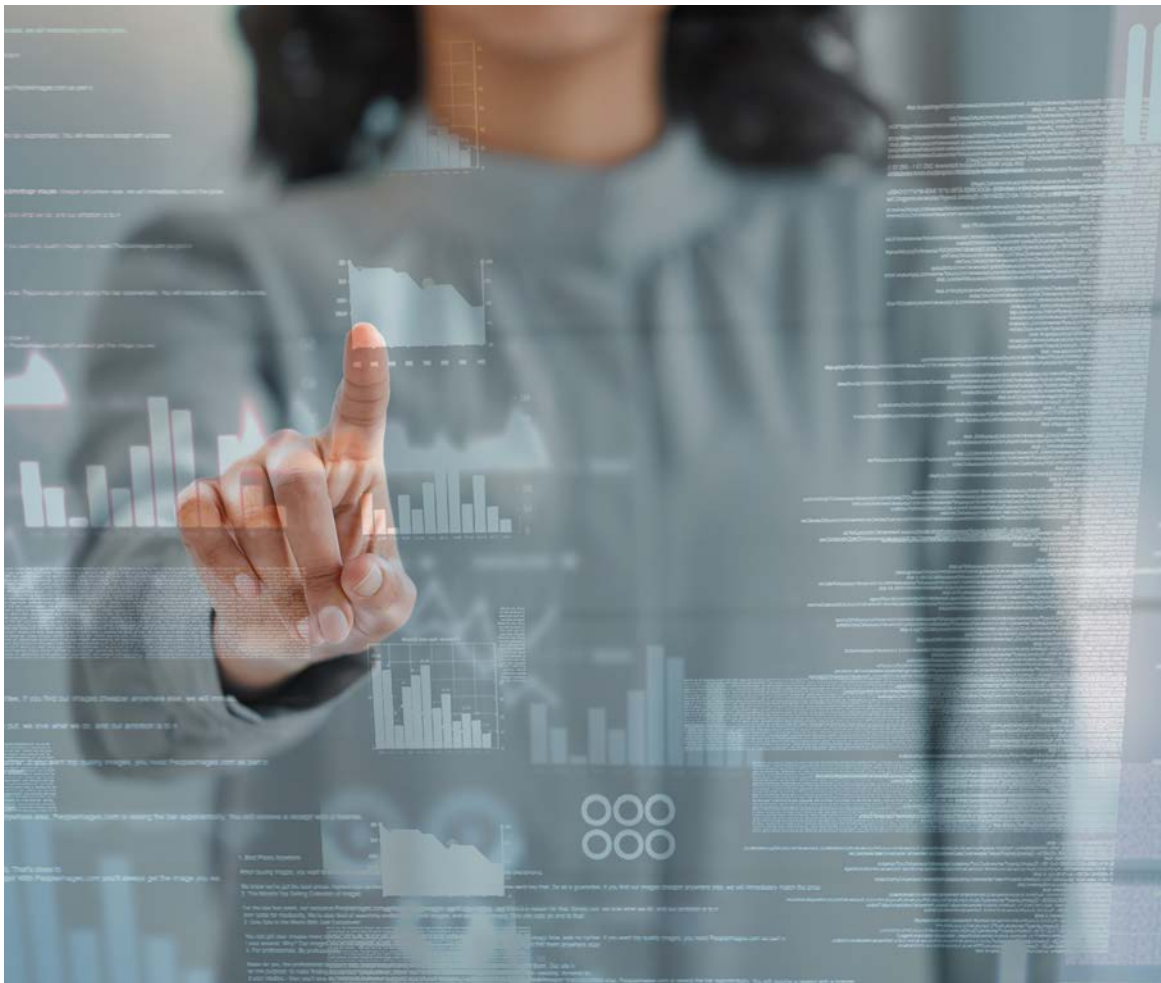
Stufe 4: Die Daten sollen im Web mit einem „Unified Resource Identifier“ (URI) veröffentlicht werden, um eine persistente Referenzierung mittels Verlinkung zu ermöglichen.

Stufe 5: Die Daten werden mit anderen Daten verlinkt, um die Daten in einen Kontext zu stellen.

Was ist ein Unified Resource Identifier (URI)?

Bei einem „Uniform Resource Identifier“ (URI) handelt es sich um eine eindeutige Zeichenfolge, die der Identifikation einer abstrakten oder physischen Ressource dient. Das bedeutet, dass eine URI nicht zwangsläufig aufgelöst werden kann, d.h. zu einer tatsächlich erreichbaren oder existierenden Ressource führen muss. Die Syntax für URIs ist im [RFC 3986](#)-Standard definiert.

Im Alltag haben wir täglich mit URIs zu tun. Beispielsweise besitzt jede Internetseite eine eindeutige URI in Form einer URL („Uniform Resource Locator“). Bei URLs handelt es sich also um eine Teilmenge von URI.



Von Rohinformationen zu offenen Daten mit 5 Sternen

Das 5-Sterne-Modell kann als Leitfaden für die Veröffentlichung von Informationen als offene Daten dienen. Grundsätzlich führt die Veröffentlichung von Informationen unter einer offenen Lizenz zur digitalen Verfügbarkeit und damit zum ersten Stern. Das Dateiformat ist dabei unerheblich.

Werden die Informationen zusätzlich in einem strukturierten Dateiformat zur Verfügung gestellt, sind die Anforderungen für den zweiten Stern erfüllt. Strukturierte Formate sind wertvoll, weil sie die compu-

tergestützte Analyse und Visualisierung ermöglichen.

Maschinenlesbarkeit impliziert, dass Computer ohne menschliches Eingreifen Daten verarbeiten können. Darüber hinaus können strukturierte Daten von einem Format in ein anderes, ebenfalls strukturiertes Format übertragen werden und tragen so zur Interoperabilität der Daten bei. Beispiele für maschinenlesbare Formate sind .xml, .csv und .json. Formate wie .xlsx oder .pdf können unter bestimmten Bedingungen maschinenlesbar sein.

Exkurs: Wann ist ein PDF strukturiert und maschinenlesbar?

Strukturierte Daten und Maschinenlesbarkeit sind eng miteinander verknüpft, weil maschinenlesbare Daten nur dann effizient verarbeitet werden können, wenn sie in einem strukturierten Format vorliegen. Das geläufige PDF ist hierbei ein gutes Beispiel, um die Unterschiede zu verdeutlichen:

PDFs sind primär für die Lesbarkeit durch Menschen konzipiert und erfüllen nicht zwangsläufig die Kriterien eines strukturierten, maschinenlesbaren Datenformats. PDF-Dokumente können jedoch, wenn sie als „tagged PDF“ generiert sind, eine Inhaltsstruktur enthalten und entsprechend strukturiert sein. Häufig ist diese aber entweder gar nicht vorhanden oder nicht standardisiert genug, um eine uneingeschränkte maschinelle Verarbeitung zu garantieren. Damit erfüllen Informationen, die als PDF veröffentlicht werden, häufig lediglich die Kriterien für den ersten Stern des 5-Sterne-Modells.

Ein Nachteil einiger strukturierter Formate ist, dass für ihre Verarbeitung kommerzielle, d.h. proprietäre, Software benötigt wird. Das schränkt den Kreis der Nutzerinnen und Nutzer ein und widerspricht dem Open-Data-Prinzip. Aus diesem Grund knüpft das

5-Sterne-Modell den dritten Stern an die Nutzung von offenen Formaten. Offene Formate haben den Vorteil, dass sie keine proprietäre Software für die Verarbeitung voraussetzen. Typische offene Dateiformate sind .csv, .xml und .json.

i Mehr Informationen zu Datenformaten im Hinblick auf die Anforderungen des 5-Sterne-Modells finden Sie in der Handreichung „Datenformate“ auf der Transparenz- und Offene-Daten-Plattform open.rlp.de.

Die ersten drei Sterne des 5-Sterne-Modells fokussieren sich auf das Format der Daten. Ab dem vierten Stern werden Unified Resource Identifier (URIs) als Bezeichner für Dinge verwendet. Die Daten sind dabei nicht mehr lediglich über das Web verfügbar, sondern werden ein Teil des Webs. Die Verwendung von URIs ermöglicht eine ein-

deutige Referenzierung einzelner Entitäten innerhalb eines Datensatzes und die Verlinkung darauf durch andere von außen. Die Bereitstellung der Daten unter einer URI ist aufwändiger und erfordert gegebenenfalls ein anderes, geeigneteres Format, wie beispielsweise das „Resource Description Framework“ (RDF).

Exkurs: Resource Description Framework (RDF)

RDF ist ein vom „World Wide Web Consortium“ (W3C) entwickeltes Standardmodell für den Datenaustausch im Internet. Die Struktur von RDF setzt sich aus Subjekt, Prädikat und Objekt zusammen, die als sogenanntes „Triple“ logische Aussagen über Ressourcen bilden. Folgende Aussage verdeutlicht die Darstellung verbal:

Das Gutenberg-Museum (Subjekt) liegt in (Prädikat) Mainz (Objekt).

Dadurch, dass mehrere Aussagen zu einer Ressource getroffen werden können und das Objekt eines Triples das Subjekt eines anderen Triples bilden kann, entsteht eine Verknüpfungsstruktur, wie folgende Aussage verdeutlicht:

Mainz (Subjekt) ist die Landeshauptstadt von (Prädikat) Rheinland-Pfalz (Objekt).

Eine Repräsentation von Subjekt, Prädikat und Objekt erfolgt in Form von URIs, wobei es sich bei dem Objekt auch um ein Literal handeln kann. Literale kommen zum Einsatz, wenn man keine Ressource, sondern einen einfachen Wert (Zahlen- oder Buchstabenfolgen aus UNICODE-Zeichen) hinterlegen möchte. Geläufige Beispiele hierfür sind Datumsangaben oder Freitext. So kann der Titel von <https://open.rlp.de> als Literal „open.rlp“ definiert werden:

<<https://open.rlp.de>> <<http://purl.org/dc/terms/title>> „open.rlp“

Literale können über eine URI mit einem Datentyp versehen werden. Eine Datumsangabe gemäß ISO 8601 beispielsweise kann wie folgt angegeben werden:

*<<https://open.rlp.de>><<http://purl.org/dc/terms/modified>>
„2025-03-20T01:27:38.472915“^^<http://www.w3.org/2001/XMLSchema#dateTime>*

Zur Darstellung von RDF existieren verschiedene Darstellungsformen, wie beispielsweise „N-Triples“ (im Beispiel verwendet), „RDF/XML“ und „Turtle“.

❶ Mehr Informationen zu RDF finden Sie in der Handreichung „Linked Open Data“ auf der Transparenz- und Offene-Daten-Plattform open.rlp.de.

Erfüllen die Daten die Voraussetzungen der ersten vier Sterne, ist die Basis für den fünften Stern gelegt: „Linked Open Data“. Die Idee hinter Linked Open Data ist, Informationen aus verschiedenen Quellen zu vernetzen und so deren Wert zu steigern.

Wertsteigerung bedeutet in diesem Fall, dass eine bessere Datenintegration ermöglicht, Interoperabilität gefördert und

der Zugriff auf verteilte Wissensbestände erleichtert werden. Der fünfte Stern erfordert daher die Verknüpfung von Daten mit anderen Informationen und Daten. Durch die Verlinkung wird die Sichtbarkeit und Auffindbarkeit aller Daten erhöht und die Informationen gewinnen durch die Einbettung in einen Kontext an Bedeutung und Verständlichkeit.

Generative Künstliche Intelligenz und offene Daten

Das Aufkommen generativer Künstlicher Intelligenz (KI) birgt neues Potential für offene Daten. Offene Daten können als Trainingsdaten verwendet werden und damit für eine breitere informative und sprachliche Basis für große Sprachmodelle sorgen. Um die Hürden für die Nutzung dieser Daten möglichst gering zu halten, ist es zwingend erforderlich, Daten strukturiert, maschinenlesbar und offen bereitzustellen.

Open Data und Rechtsvorschriften

Das Potenzial offener Daten wird auf unterschiedlichen politischen Ebenen erkannt und schlägt sich in entsprechenden Richtlinien der Europäischen Union, in Bundesgesetzen sowie in der rheinland-pfälzischen Landesgesetzgebung nieder.

Stand Sommer 2025 wird in Rheinland-Pfalz ein Open-Data-Gesetz erarbeitet. In diesem sollen Landesbehörden mittels einer Soll-Vorschrift und im Rahmen ihrer Ressourcen dazu verpflichtet werden offene Daten bereitzustellen. Für Kommunen ist keine Pflicht, aber ein Angebot zur Unter-

stützung bei der Veröffentlichung vorgesehen.

Open-Data-Richtlinie und HVD-Durchführungsverordnung

Auf europäischer Ebene hat die „Richtlinie (EU) 2019/1024 über offene Daten und die Weiterverwendung von Informationen des öffentlichen Sektors“ (sog. „Open-Data-Richtlinie“) die Förderung der breiten Nutzung öffentlicher und öffentlich finanzierter Daten zur Unterstützung digitaler Innovationen zum Ziel. Dabei steht insbesondere der Bereich der künstlichen Intelligenz im

Fokus. Die Richtlinie definiert Mindestanforderungen für die Nachnutzung von öffentlichen Informationen und hat das Ziel, Open Data zu fördern und hebt die Prinzipien „Open-by-Design“ und „Open-by-Default“ hervor.

Entsprechend schreibt die Richtlinie die Veröffentlichungen von Informationen als of-

fene Daten in maschinenlesbaren, strukturierten und offenen Formaten vor. Darüber hinaus sollten die Dateiformate sowie die zugehörigen Metadaten möglichst offenen Standards entsprechen. Die Verwendung von möglichst offenen Standardlizenzen wird angeregt. Die Open-Data-Richtlinie bildet die Basis für die „Durchführungsverordnung (EU) 2023/138“.

Open-by-Design und Open-by-Default

Die Prinzipien Open-by-Design und Open-by-Default fördern die Bereitstellung offener Daten, unterscheiden sich jedoch in ihrer Herangehensweise. Open-by-Design bedeutet, dass Offenheit bereits bei der Erhebung, Speicherung und Bereitstellung von Daten berücksichtigt wird. Dies umfasst die Nutzung offener Formate, standardisierter Schnittstellen (APIs) und maschinenlesbarer Strukturen, um die Nachnutzung der Daten zu erleichtern.

Zudem soll sichergestellt werden, dass rechtliche Rahmenbedingungen wie offene Lizenzen von Beginn an berücksichtigt werden.

Open-by-Default zielt darauf ab, dass öffentliche Daten grundsätzlich unter offenen Lizenzen veröffentlicht werden, es sei denn, es gibt triftige Gründe für eine Einschränkung: etwa Datenschutz, Sicherheitsrisiken oder rechtliche Vorgaben. Während Open-by-Design sicherstellt, dass Offenheit strukturell verankert ist, setzt Open-by-Default darauf, Offenheit als Standard anzunehmen.

Die „Durchführungsverordnung (EU) 2023/138“ ist auch als „High Value Dataset-Verordnung“ (HVD-Verordnung) bekannt und verpflichtet öffentliche Stellen der EU-Mitgliedstaaten, bestimmte Datensätze des öffentlichen Sektors, sog. hochwertige Datensätze (engl. high value datasets) kostenfrei, in einem maschinenlesbaren Format und unter offenen Lizenzen bereitzu-

stellen. Diese Datensätze werden aufgrund ihres hohen sozioökonomischen Potenzials als besonders wertvoll für Gesellschaft, Umwelt und Wirtschaft eingestuft. Hochwertige Datensätze sind Datensätze, die unter eine der folgenden Kategorien fallen: Geodaten, Erdbeobachtung und Umwelt, Meteorologie, Statistik, Unternehmen und Mobilität.

Merkmale hochwertiger Datensätze

Themenbereiche

- A. Sozial
- B. Klimawandel und Umwelt
- C. Wirtschaft
- D. Innovation und KI
- E. Öffentliche Dienste und Verwaltung
- F. Nachnutzung

Verbindliche technische Anforderungen

- 1. Creative-Commons-BY-4.0 oder gleichwertige Lizenz
- 2. Kostenfrei
- 3. API und Massendownload
- 4. Maschinenlesbares Format
- 5. Umfassende Metadaten
- 6. Öffentlich verfügbare Dokumentation

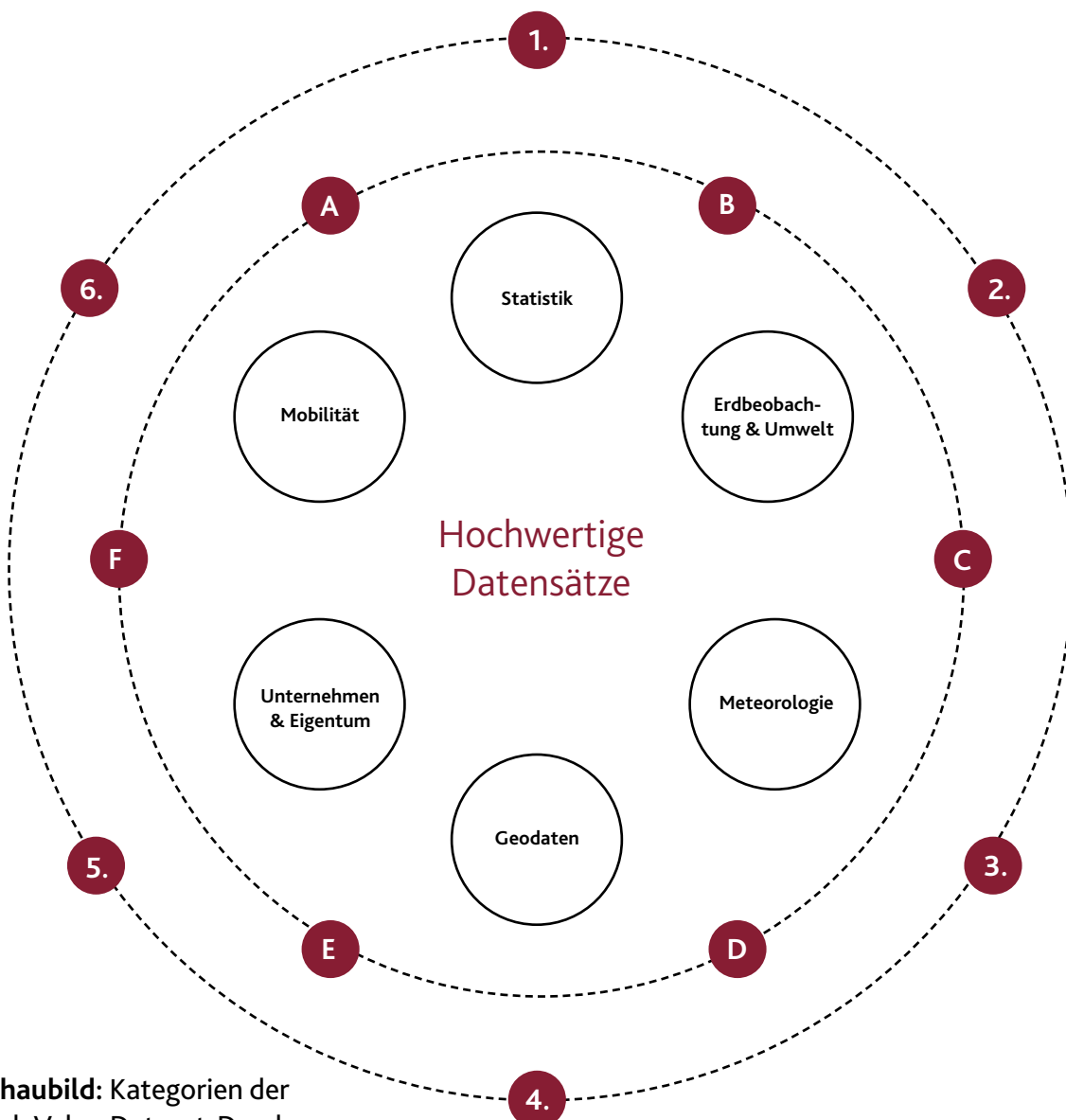


Schaubild: Kategorien der High Value Dataset-Durchführungsverantwortung

Die Daten müssen als HVD-Datensätze in den Metadaten gekennzeichnet sein und als Massendownload über eine Anwendungs-Programmier-Schnittstelle (engl. „Application Programming Interface“, API) abrufbar sein. Die Veröffentlichung hochwertiger Datensätze obliegt den jeweiligen Verwal-

tungsstellen, die im Besitz der Daten sind. Dies umfasst unter anderem Ministerien, deren nachgeordnete Bereiche sowie Kommunen und kommunale Einrichtungen. Die Europäische Kommission hat bereits angekündigt, dass diese Kategorien in Zukunft erweitert werden können.

Metadaten

Metadaten sind strukturierte Daten, die die eigentlichen (offenen) Daten beschreiben, d.h. Daten, die Informationen über Daten enthalten. Sie helfen dabei, Daten zu organisieren, zu identifizieren, zu beschreiben oder leichter auffindbar zu machen. Metadaten können genauso wie die eigentlichen Daten strukturiert und in einem standardisierten Format gespeichert werden. Ein Beispiel für einen in der Europäischen Union weit verbreiteten Standard für Metadaten ist [DCAT-AP](#) („Data Catalog Vocabulary Application Profile“).

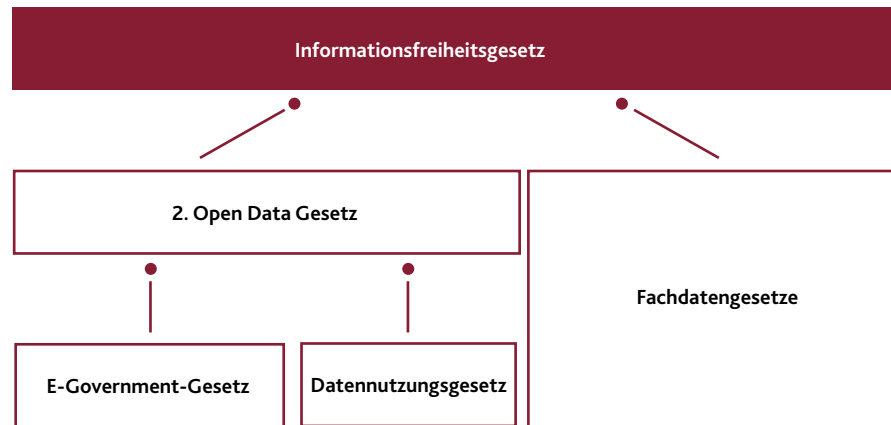
Zweites Open-Data-Gesetz und Fachdatengesetze

Die spezifischen Regelungen auf nationaler Ebene basieren auf der Umsetzung der „EU-Richtlinie 2019/1024“ sowie der „Durchführungsverordnung 2023/138“. Entsprechend ist es auch ihr Ziel, den Zugang zu öffentlichen Informationen zu fördern und die Nachnutzung von Daten durch Bürgerinnen und Bürger sowie Unternehmen zu erleichtern. Neben dem „Informationsfreiheitsgesetz“ (IFG) gelten insbesondere das „Zweite Open-Data-Gesetz“ (2. ODG) sowie diverse Fachdatengesetze und Vorgaben. Das IFG gewährt bereits seit 2006 Bürgerinnen und Bürgern das Recht gegenüber Behörden, Einrichtungen und Organen des

Bundes, Einsicht in amtliche Dokumente zu nehmen. Dies kann jedoch nur erfolgen, sofern dem keine schutzwürdigen Interessen entgegenstehen. Es stellt chronologisch die zentrale Säule und den Ausgangspunkt des in weiteren Gesetzen verankerten Open-Data-Prinzips in der nationalen Gesetzgebung dar.

Das IFG schafft die Basis, um einerseits den freien Zugang zu amtlichen Informationen sicherzustellen und andererseits den Grundstein für eine effektive und moderne Nutzung der Daten im Sinne von Transparenz und Innovation zu legen.

Schaubild: Nationale Gesetze mit Open-Data-Bezug



Während das IFG damit allgemein den Zugang zu Informationen fördert, definiert das 2. ODG spezifische Anforderungen an die Veröffentlichung von Daten. Es bildet den zentralen Baustein der nationalen Umsetzung der EU-Richtlinie 2019/1024. Geltende Vorgaben des Gesetzes zur Förderung der elektronischen Verwaltung, des „E-Government-Gesetzes“ (EGovG), werden darin angepasst und es führt das „Datennutzungsgesetz“ (DNG) ein.

Das EGovG verpflichtet Bundesbehörden und – abhängig von den ihnen delegierten Aufgaben – auch andere öffentliche Stellen zur Veröffentlichung von Informationen in standardisierten, maschinenlesbaren und nicht-proprietären Formaten. Die Daten sind dabei unter möglichst wenig restriktiven Lizenzen ohne Gebühren bereitzustellen. Zusätzlich soll eine Veröffentlichung bzw. Bereitstellung der Daten mitsamt den Metadaten auf dem nationalen Metadatenportal „GovData“ erfolgen.

Das DNG hat einen breiteren Geltungsbe-

reich. Es definiert die Art und Weise der Veröffentlichung, für alle öffentlichen Stellen, die Daten von öffentlichem Interesse besitzen, für die eine Veröffentlichungspflicht besteht. Öffentliche Stellen sind jene, die eine hoheitliche Aufgabe wahrnehmen. Der Datenbegriff bezieht sich auf vorhandene Aufzeichnungen und ist breit gefasst. Diese sollen nach dem DNG maschinenlesbar, in einem nicht-proprietärem Format und nach Möglichkeit unter offenen Lizenzen veröffentlicht werden. Auch das DNG verpflichtet die Kennzeichnung von Daten mittels Metadaten und die Veröffentlichung auf GovData. Zusätzlich fordert das Gesetz unter anderem die Nutzung standardisierter Schnittstellen (z. B. APIs) und Metadatenformate bei der Veröffentlichung, um die einfache Integration und Weiterverarbeitung der Daten in unterschiedlichen Systemen zu ermöglichen.

Neben diesen zentralen Regelungen existieren diverse Fachgesetze, welche weitere rechtliche Aspekte bezüglich der Veröffentlichung von Daten und Informationen

beinhalten. Einige Beispiele umfassen das „Umweltinformationsgesetz“ (UIG), das „Gesetz zur Verbesserung der gesundheitsbezogenen Verbraucherinformation – Verbraucherinformationsgesetz“ (VIG), das „Gesundheitsdatennutzungsgesetz“ (GDNG), das „Gesetz über den Deutschen Wetterdienst“ (DWDG), das „Gesetz über die Statistik der See- und Binnenschifffahrt, des Güterkraftverkehrs, des Luftverkehrs sowie des Schienenverkehrs und des gewerblichen Straßen-Personenverkehrs

– Verkehrsstatistikgesetz“ (VerkStatG), das „Gesetz über die Erhebung von streckenbezogenen Gebühren für die Benutzung von Bundesautobahnen und Bundesstraßen – Bundesfernstraßenmautgesetz“ (BFStrMG), das „Gesetz über die Nutzung und Sicherung von Archivgut des Bundes – Bundesarchivgesetz“ (BArchG) sowie die „Bundeshaushaltsordnung“ (BHO). Zusätzlich sind ein Mobilitätsdatengesetz sowie ein Forschungsdatengesetz in Planung.



Landestransparenzgesetz Rheinland-Pfalz

Das „Landestransparenzgesetz Rheinland-Pfalz“ (LTranspG) ist seit 2016 in Kraft und verfolgt das übergeordnete Ziel, die Transparenz und Offenheit der öffentlichen Verwaltung zu erhöhen und die Möglichkeit der Kontrolle staatlichen Handelns durch die Bürgerinnen und Bürger zu verbessern. Das Gesetz schafft die rechtliche Grundlage dafür, dass amtliche bzw. behördliche Informationen sowie Umweltinformationen über eine gemeinsame rechtliche Regelung veröffentlicht werden, unabhängig von der Art ihrer Speicherung.

Das LTranspG beinhaltet landesspezifische Regelungen zur Veröffentlichung von Informationen. Die Informationen sollen in offenen, plattformunabhängigen, maschi-

nenlesbaren Formaten nach anerkannten offenen Standards zur Verfügung gestellt werden. Bereitgestellte Informationen sollen darüber hinaus mit Metadaten in einem offenen Standard versehen werden. Der Zugang und die Nutzung der Transparenzplattform sollen grundsätzlich kostenfrei, anonym und ohne Einschränkungen möglich sein.

Die Digitalstrategie des Landes Rheinland-Pfalz sowie die Open-Data-Strategie sehen zusätzlich die Einführung eines rheinland-pfälzischen Open-Data-Gesetzes vor, das sich Stand Herbst 2025 in der Vorbereitung befindet und, als zweite Säule neben dem LTranspG, das Open-Data-Prinzip auf rechtlicher Ebene komplettieren soll.

Die Transparenz- und Offene-Daten-Plattform: open.rlp.de

Ein zentrales Element des LTranspG ist die Schaffung einer Transparenz-Plattform, auf der amtliche Informationen und Umweltinformationen veröffentlicht werden und somit Bürgerinnen und Bürgern kostenfrei zur Verfügung gestellt werden. Zusätzlich verpflichtet das Gesetz die Landesbehörden, nicht schützenswerte Informationen proaktiv auf der Plattform bereitzustellen. Beispiele dafür sind Ministerratsbeschlüsse, Gutachten, öffentliche Pläne, Geodaten oder Daten von Unternehmen, an denen das Land beteiligt ist.

Dem ist das Land Rheinland-Pfalz mit der Einführung der Landestransparenz-Plattform im Jahr 2016 nachgekommen, die nun in die [Transparenz- und Offene-Daten-Plattform](https://open.rlp.de) open.rlp.de überführt wurde. Hier werden nach dem Landestransparenzgesetz veröffentlichungspflichtige Informationen sowie offene Daten veröffentlicht. Von open.rlp.de aus werden die Metadaten der offenen Daten aus Rheinland-Pfalz an das Datenportal für Deutschland, govdata.de, geleitet.



TRANSPARENZ- UND OFFENE- DATEN-PLATTFORM DES LANDES RHEINLAND-PFALZ

Die Transparenz- und Offene-Daten-Plattform des Landes Rheinland-Pfalz, open.rlp.de, macht die Landes- und Kommunalverwaltung für alle zugänglich, indem die Bestände an offenen Daten sowie Transparenzinformationen in Form eines umfassenden Metadatenkatalogs auf der Plattform veröffentlicht und durchsuchbar gemacht werden.

Dabei werden die eigentlichen Daten durch Verlinkungen repräsentiert, sodass sie in Fachportalen bzw. bei den veröffentlichenden Stellen verbleiben, wo sie gespeichert und gepflegt werden. Ein Beispiel für ein solches Fachportal ist das „GeoPortal“ des Landes Rheinland-Pfalz, <https://www.geoportal.rlp.de>.

Fachportale sammeln in der Regel einen Bestand an offenen Daten zu einem bestimmten Thema und bieten dazu einen eigenen Metadatenkatalog zum Abruf über eine Schnittstelle an. open.rlp.de fragt diese Schnittstellen in regelmäßigen Abständen ab und führt die einzelnen Kataloge zu einem zentralen Metadatenkatalog für Rheinland-Pfalz zusammen. Dieser Katalog wird dann über einen RDF-Endpunkt bereitgestellt, der unter anderem einen Abruf durch das Datenportal des Bundes, „GovData“, ermöglicht.

Auf Bundesebene führt GovData die einzelnen Kataloge der Länder zusammen und leitet diesen Metadatenkatalog zum europäischen Portal, „data.europa.eu“, aus.



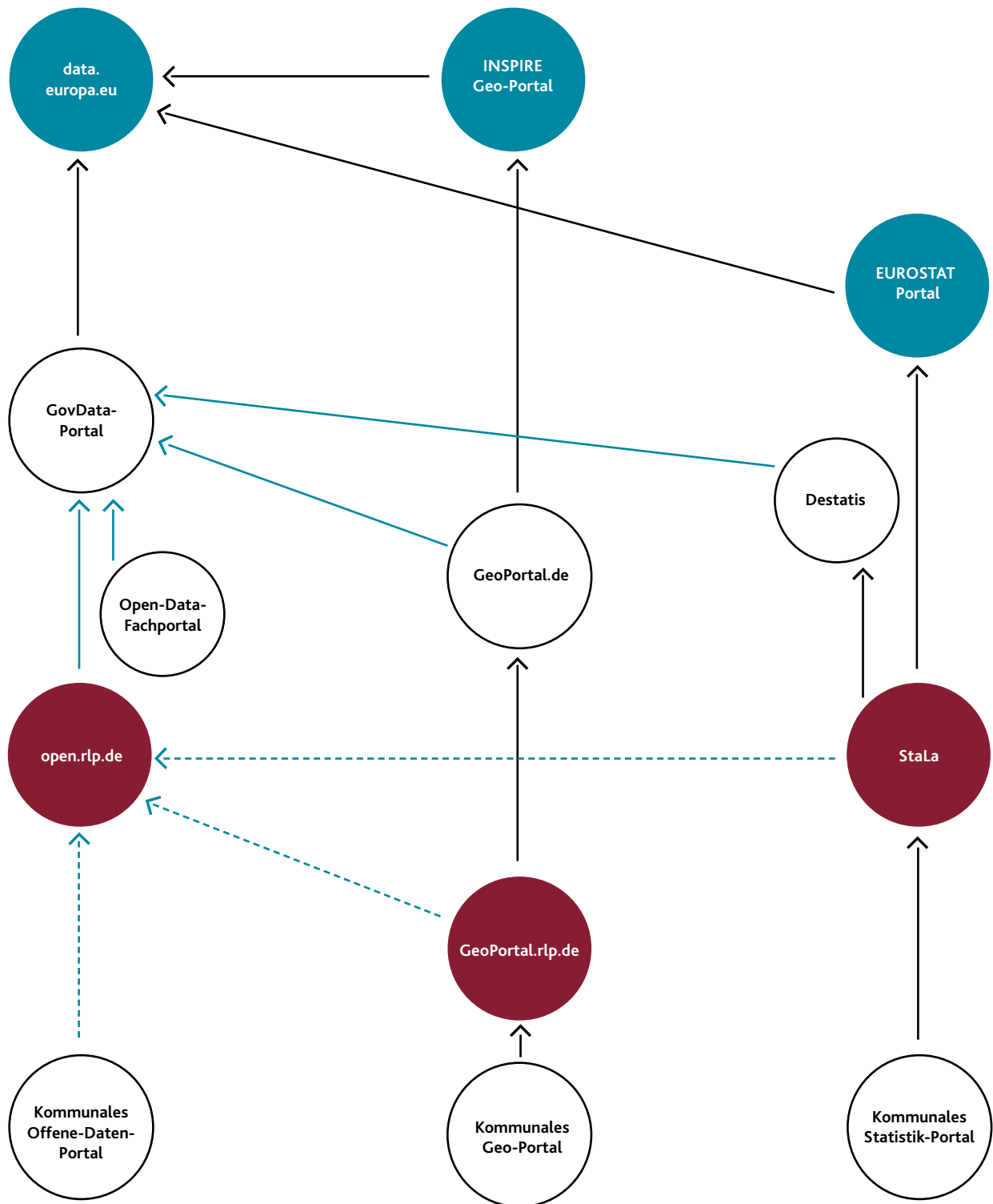


Abbildung 1: Standardisierung in der Metadatenförderung GovData.

DCAT-AP.de
verpflichtend

DCAT-AP.de
empfohlen

Die Metadatenföderation GovData

Um dem Standardisierungsbedarf, der durch Datenfluss zwischen den Portalen entsteht, gerecht zu werden und die EU-weite Interoperabilität zwischen den Datenkatalogen zu gewährleisten, kommen beim Austausch zwischen den Portalen der europäische Metadatenstandard „DCAT-AP“ und dessen nationale Ausprägungen zum Einsatz. Diese Standards basieren auf dem „Resource Description Framework“ (RDF) und werden im Regelfall als RDF/XML serialisiert.

Die deutsche Profilierung des Standards heißt „DCAT-AP.de“. Er wurde auf Beschluss des IT-Planungsrats veröffentlicht und seine Nutzung ist im GovData-Verbund verpflichtend. DCAT-AP.de ergänzt DCAT-AP um zusätzliche Einschränkungen und Erweiterungen und bildet damit die Spezifika der deutschen Verwaltung, wie zum Beispiel

Gemeinde- und Regionalschlüssel, mittels kontrollierter Vokabulare als Metadaten ab.

Kontrollierte Vokabulare sind Sammlungen an standardisierten Begriffen. Sie dienen der Vermeidung von Mehrdeutigkeit und erhöhen die Genauigkeit der Beschreibungen sowie den Grad der Standardisierung. All dies dient der Optimierung der Metadatenqualität.

Entsprechend basiert der Metadatenkatalog auf open.rlp.de ebenfalls auf dem Standard DCAT-AP.de. Auf der Ebene unterhalb der Landesportale ist die Befolgung von DCAT-AP.de nicht verbindlich – sie wird lediglich empfohlen. Eine Verwendung des Standards durch die Fachportale erleichtert jedoch die Anbindung an open.rlp.de und ist somit erwünscht.

Aufbau der Plattform

Die Plattform bietet Möglichkeiten zum Abruf, zur Speicherung und zur Bereitstellung von Metadaten. Nutzer können den Metadatenbestand entweder über die graphische Oberfläche, die API oder über den RDF-Endpunkt abrufen und durchsuchen. Darüber hinaus bietet open.rlp.de ein umfassendes Angebot an redaktionellen Inhalten mit Informationen zu den Themen Transparenz und offene Daten.

Architektonisch baut open.rlp.de maßgeblich auf zwei Komponenten auf. Ein TYPO3-System steht als Frontend bereit. Es dient der Anzeige von Suchergebnissen und wird

von den Verantwortlichen der Plattform genutzt, um redaktionelle Inhalte zu pflegen. Über die CKAN-API findet der Datentransfer mit dem CKAN statt, das als Backend dient.

CKAN steht für „Comprehensive Knowledge Archive Network“. Dabei handelt es sich um ein Open-Source-Datenmanagementsystem, das weltweit als Basis für Offene-Daten-Portale genutzt wird. Es bietet Funktionen zur Speicherung, Verwaltung und Bereitstellung von Daten und den zugehörigen Metadaten. Veröffentlichte Datenbestände können über eine Facettensuche

auf Basis des gesamten Metadatenbestands durchsucht werden. Dabei setzt CKAN auf die Suchplattform Solr. Zur dauerhaften Speicherung und Verwaltung der Daten kommt eine PostgreSQL-Datenbank zum Einsatz. Darüber hinaus verfügt CKAN standardmäßig über eine Schnittstelle, die den maschinellen Abruf der Daten als JSON erlaubt. Über Erweiterungen werden Harvester zum Abruf der Metadaten bei den Lie-

fersystemen und eine Unterstützung von DCAT-AP.de integriert. So kann gewährleistet werden, dass der aktuelle Datenbestand in Form eines Metadatenkatalogs auf open.rlp.de in regelmäßigen Abständen aus den Datenbeständen der einzelnen Liefersysteme zusammengeführt und der eigene Katalog über einen RDF-Endpunkt ausgeleitet werden kann.



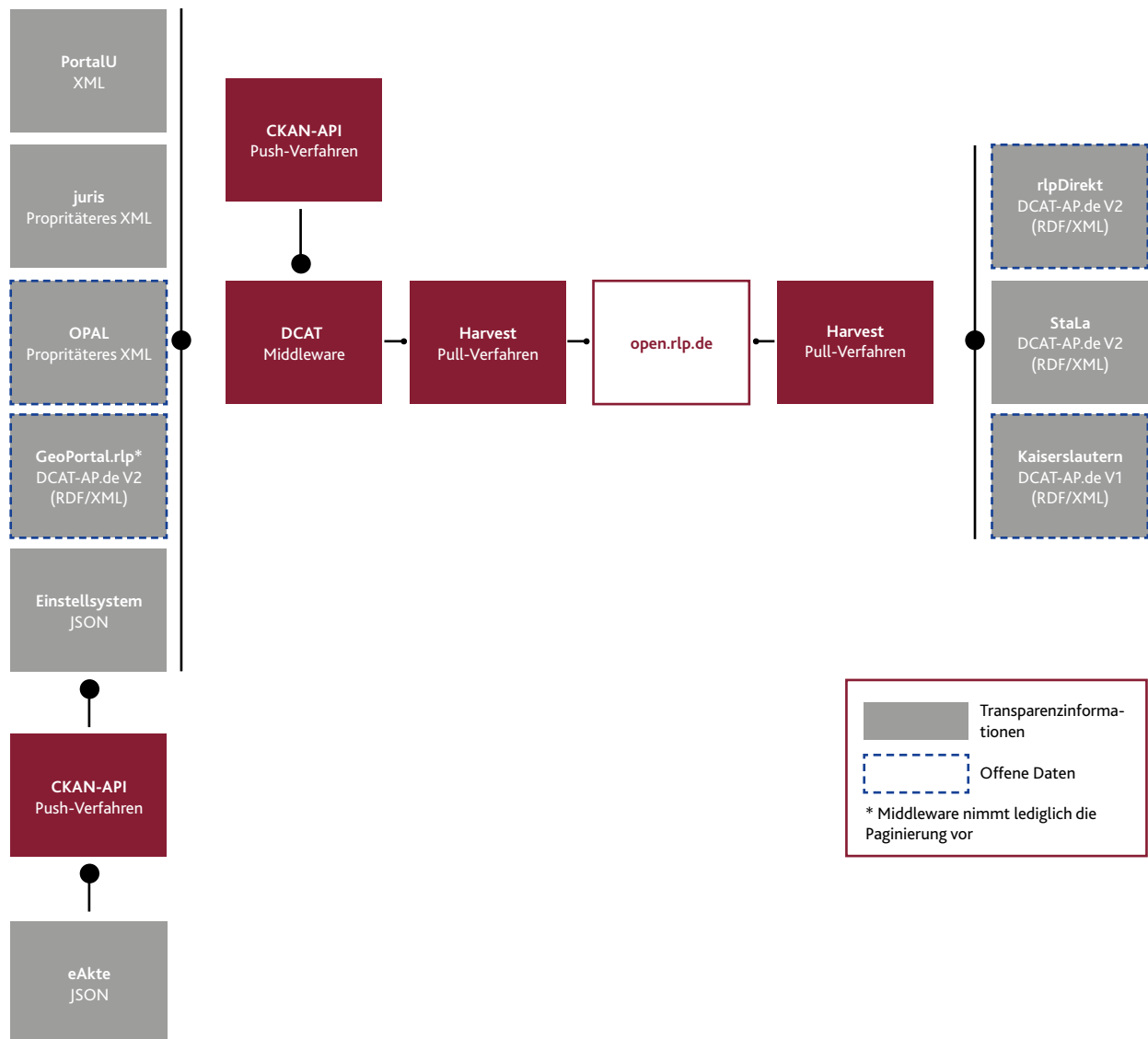
Angebundene Liefersysteme

Der Metadatenkatalog auf open.rlp.de setzt sich aus den Datenbeständen der verschiedenen Liefersysteme zusammen, die zeitgesteuert über Harvester durch das CKAN-System abgerufen werden.

Da eine Bereitstellung der Metadaten als DCAT-AP.de-konformes RDF/XML durch die

Liefersysteme nicht immer gewährleistet werden kann, wurde für den Abruf eine Zwischenschicht (Middleware) implementiert, die bei Bedarf eine Konsolidierung der einzelnen Datenbestände und die Überführung in den Standard vornimmt.

Abbildung 2: An open.rlp.de angebundene Liefersysteme und Modi der Anbindung.



Die Verantwortlichen bei open.rlp.de sind stets bestrebt, neue Datenbestände an die Plattform anzubinden. In Kapitel 4 finden Sie weitere Hinweise für eine Anbindung an open.rlp.de.



GRUNDLAGEN ZUR VORBEREITUNG VON DATEN UND METADATEN

Für die auf der Transparenz- und Offene-Daten-Plattform des Landes Rheinland-Pfalz zur Verfügung gestellten Daten und Dokumente sind die einzelnen Datenbereitsteller verantwortlich. Daher müssen die Daten vor der Veröffentlichung der Metadaten auf open.rlp.de aufbereitet werden, damit sie die gewünschten inhaltlichen und technischen Voraussetzungen erfüllen. Außerdem sollte sichergestellt werden, dass die Da-

ten nicht personenbezogen oder aus anderen Gründen schützenswert sind. Die dafür erforderlichen Schritte sind (unter Umständen) die Datenanonymisierung, das Erstellen der zugehörigen Metadaten sowie die Auswahl der Lizenzen und des Formates. Folgendes Schaubild skizziert die Schritte, die in den darauffolgenden Abschnitten erläutert werden.

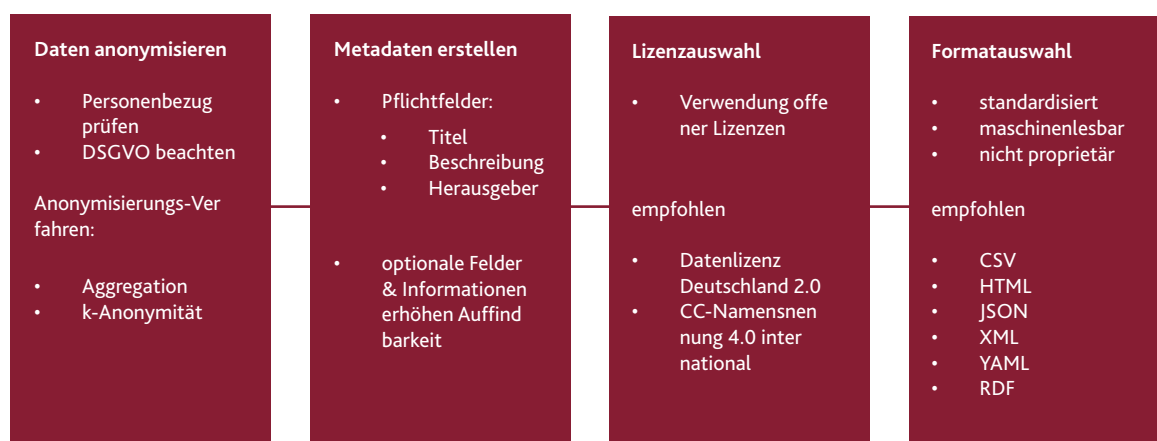


Schaubild: Vorbereitung von Daten und Metadaten

Daten anonymisieren

Informationen, die zur Veröffentlichung als offene Daten bereitgestellt werden, sollten dahingehend geprüft werden, ob sie personenbezogene Daten enthalten. Ist dies nicht der Fall, können die Daten verwendet werden.

Falls die Daten Personenbezüge enthalten, kann entschieden werden, dass eine Veröf-

fentlichung dennoch sinnvoll ist. Dies gilt vor allem dann, wenn die Daten einen besonderen gesellschaftlichen Nutzen haben. Damit ihre Veröffentlichung nicht gegen geltende datenschutzrechtliche Bestimmungen verstößt, müssen diese Informationen anonymisiert werden.

Personenbezug in Daten

Die „Europäische Datenschutzgrundverordnung“ (DSGVO) legt die rechtlichen Rahmenbedingungen für den Datenschutz fest und wird durch Bundes- und Landesgesetze (BDSG, LDSG RLP) ergänzt. In der DSGVO werden personenbezogene Daten als Informationen definiert, die sich auf eine identifizierte oder identifizierbare Person beziehen. Damit wird deutlich, dass neben offensichtlichen Informationen, wie Name oder Adresse, auch weniger offensichtliche Informationen, wie etwa IP-Adressen, personenbezogen sein können.

Bei Unsicherheit sollte entsprechend eine Datenschutzexpertin oder ein Datenschutzexperte zur Einschätzung hinzugezogen werden.

Die Anonymisierung von Daten kann durch unterschiedliche Verfahren erfolgen. Zwei Beispiele sind die „Aggregation“ und die „k-Anonymität“, die aufeinander aufbauen. Bei der Aggregation werden Daten von der Individualebene zu Gruppen geclustert. So können beispielsweise bei Adressdaten Straßennamen und -nummern durch Postleitzahlen ersetzt werden, um den Rückschluss auf einzelne Personen zu verhindern. Dennoch kann es vorkommen, dass Gruppen zu klein sind und im ungünstigsten Fall aus einer Person bestehen, die dadurch identifizierbar bleibt.

Die sog. „k-Anonymität“ ist die Erweiterung der Aggregation. Bei dieser Methode wird sichergestellt, dass bei der Aggregation eine ausreichende Anzahl an Personen (wobei k die Anzahl der Personen darstellt) mit den gleichen Merkmalen pro Gruppe bestehen bleibt, damit ein Rückschluss auf die Individualebene nicht möglich ist.

Das Löschen einzelner Dateneinträge kann auch nach der Anwendung dieser Verfahren notwendig sein, wenn weiterhin ein Personenbezug besteht.

- ① Eine Handreichung zum Thema Anonymisierung finden Sie auf der Transparenz- und Offene-Daten-Plattform open.rlp.de.

Metadaten erstellen

Auf open.rlp.de werden Daten nicht direkt, sondern als Metadaten nach dem DCAT-AP.de-Standard veröffentlicht. Sorgfalt bei der Erstellung dieser Metadaten ist die Grundvoraussetzung für die Auffindbarkeit der Daten.

Nach dem DCAT-AP.de-Standard ist die Angabe eines Titels, einer Beschreibung und eines Herausgebers verpflichtend. Zusätz-

lich existieren optionale Metadatenfelder, die es ermöglichen, diesen Pflichtangaben weitere Informationen hinzuzufügen. Je mehr optionale Informationen den Metadaten hinzugefügt werden, desto besser ist ein Datensatz auffindbar. Die folgende Tabelle bietet eine Übersicht über die Metadatenfelder, die für eine Veröffentlichung auf open.rlp.de als Grundbestand definiert sind.

Metadatenfeld	Beschreibung
Titel*	Eindeutiger, menschenlesbarer Name des Datensatzes. Der Name sollte keine weiteren Metadaten wie zum Beispiel Datumsangaben enthalten, da diese Angaben in extra dafür vorgesehenen Metadatenfeldern abgelegt werden.
Beschreibung*	Präzise Beschreibung des veröffentlichten Datensatzes als Freitext
Schlagwort	Schlagworte/Schlüsselbegriffe zur Beschreibung des Datensatzes
Veröffentlichungsdatum	Datum der Herausgabe des Datensatzes nach ISO 8601
Aktualisierungsdatum	Datum der letzten Aktualisierung des Datensatzes nach ISO 8601
Kategorie	Die dem Datensatz zugewiesene Informationskategorie. Die Kategorisierung erfolgt einheitlich gemäß des EU DATA Theme Vokabulars .
Herausgeber*	Öffentliche Stelle oder Person, die für die Bereitstellung des Datensatzes verantwortlich ist
HVD-Kategorie	Gibt HVD-Kategorie des Datensatzes an und ist ausschließlich für Datensätze gemäß HVD-Durchführungsverordnung notwendig. Die Kategorisierung erfolgt einheitlich gemäß des EU DATA Theme Vokabulars .
Lizenz	Lizenz, unter der der Datensatz veröffentlicht ist
Format	Dateiformat des veröffentlichten Datensatzes

*Pflichtfeld

In der Regel stehen bei den Datenbereitstellenden oder Liefersystemen Eingabemasken zur Verfügung, die der Eingabe der Daten in das jeweilige System dienen. Oft definieren

Liefersysteme einen eigenen Bestand an Metadatenfeldern, die in den Metadatenstandard DCAT-AP.de übertragen werden können.

- ① **Mehr Informationen und Beispiele zum Thema Metadaten finden Sie in der Handreichung zum Thema Metadaten auf der Transparenz- und Offene-Daten-Plattform open.rlp.de.**

Lizenzauswahl

Lizenzen legen fest, unter welchen Bedingungen ein Datensatz nachgenutzt werden kann. Für offene Daten ist die Bereitstellung unter einer offenen Lizenz ein entscheidendes Merkmal. Dadurch wird gewährleistet, dass die Daten von jedem ohne Einschränkung nachgenutzt und weiterverbreitet werden dürfen.

Es existiert eine Vielzahl an offenen Lizenzen, die dadurch unterschieden werden können, wie sie den Grad der Nutzung und Weiterverbreitung handhaben. Die bekannteste Lizenzgruppe bilden die „Creative Commons“ (CC) -Lizenzen. Speziell für

die Lizenzierung von deutschen Verwaltungsdaten existiert auch die „Datenlizenz Deutschland“ (DL DE). Beide Lizenzgruppen haben Ausprägungen, die eine bedingungslose Nutzung (Zero), eine Nutzung unter Namensnennung und eine eingeschränkte Namensnennung für nicht-kommerzielle Zwecke erlaubt.

Eine Liste der von DCAT-AP.de unterstützten Lizenzen finden Sie unter:

<https://www.dcat-ap.de/def/licenses/>

- ① **Eine Handreichung zum Thema offene Lizenzen finden Sie auf der Transparenz- und Offene-Daten-Plattform open.rlp.de.**

Formatauswahl

Neben der Wahl einer geeigneten offenen Lizenz lassen sich anhand des 5-Sterne-Modells und der gesetzlichen Vorgaben Anforderungen formulieren, die für eine möglichst hohe Zugänglichkeit und technische Verwendbarkeit des Datensatzes sorgen. Daher sollte für die Veröffentlichung ein standardisiertes, maschinenlesbares und nicht proprietäres Format gewählt werden.

Standardisierte Formate haben den Vorteil, dass die zugrundeliegende Datenstruktur vollständig dokumentiert und nachvollziehbar ist. Dies ist eine Grundvoraussetzung für die Anforderung der Maschinenlesbarkeit. Sie stellt sicher, dass die enthaltenen Daten durch Software automatisiert gelesen werden können. Der Nachteil maschinenlesba-

rer Formate ist, dass sie für Menschen in geringerem Maß zugänglich und nicht intuitiv verständlich sind. Abschließend stellt die Wahl eines nicht proprietären Formats sicher, dass für die Nachnutzung keine Softwarelizenzen benötigt werden, sodass niemand von ihr ausgeschlossen wird.

Werden diese Kriterien beachtet, ist sichergestellt, dass das maximale Potenzial von offenen Daten ausgeschöpft werden kann. Es existieren zahlreiche Datenformate, die die oben genannten Kriterien erfüllen. Die folgende Übersicht zeigt empfohlene Datenformate für die Veröffentlichung von offenen Daten: .csv, .html, .json, .xml, .yaml, .rdf.

- ❶ Mehr Informationen zu Datenformaten im Hinblick auf die Anforderungen des 5-Sterne-Modells finden sie in der Handreichung Datenformate auf der Transparenz- und Offene-Daten-Plattform open.rlp.de.





VERÖFFENTLI-
CHUNG VON
METADATEN AUF
OPEN.RLP.DE

An die Erstellung der Metadaten schließt sich ihre Veröffentlichung an, die in der Regel zunächst auf einer Fachdatenplattform einer rheinland-pfälzischen Behörde erfolgt. Diese Plattformen halten die offenen Daten zum Download vor und erfassen in der Regel die Metadaten in eigenen Katalogen.

Ein Beispiel dafür ist das „[GeoPortal.rlp](#)“ des Landesamts für Vermessung und Geobasisinformationen Rheinland-Pfalz. Damit eine zentrale Bereitstellung der Daten auf open.rlp.de und somit auch GovData erfolgen kann, müssen die Fachdatenplattformen als

Liefersysteme an open.rlp.de angebunden werden.

Für die technische Anbindung und somit die maschinelle Übertragung der Metadaten an open.rlp.de bestehen grundsätzlich zwei technische Ansätze:

- Die Möglichkeit, die Metadaten per sog. Harvester technisch abholen zu lassen oder
- die technische Anlieferung der Metadaten per standardisierter Schnittstelle (API) durchzuführen.

① **Nutzen Sie gern das Kontaktformular auf open.rlp.de, wenn Sie eine Anbindung Ihrer Plattform an die Transparenz- und Offene-Daten-Plattform open.rlp.de anstreben oder Daten bereitstellen möchten.**

Harvesting: Abholung von Metadaten

Der Metadatenkatalog von open.rlp.de setzt sich aus den verschiedenen Metadatenkatalogen einzelner Fachdatenportale, den Liefersystemen, zusammen. Ein Abruf dieser Kataloge findet zeitgesteuert und automatisiert über „Harvester“ (engl. für „Erntemaschinen“) statt, sodass neue oder geänderte Metadatensätze in regelmäßigen Abständen auf open.rlp.de übertragen bzw. aktualisiert werden. Ebenso wird ein Metadatensatz auf open.rlp.de gelöscht, wenn er aus dem Liefersystem entfernt wird.

Um den Abruf per Harvester zu ermöglichen, wird ein API-Endpunkt oder Export in einem Austauschverzeichnis benötigt, der durch den Betreiber des Liefersystems bereitgestellt werden muss und den Metadatenkatalog des Liefersystems als

DCAT-AP.de-konformes RDF/XML bereitstellt.

Werden die Metadaten über einen Export in einem Austauschverzeichnis für das Harvesting exponiert, erfolgt die Bereitstellung asynchron. Dabei muss der Metadatenkatalog durch das Fachportal in regelmäßigen zeitlichen Abständen an den entsprechenden Ablageort übertragen werden. Der Nachteil bei diesem Verfahren ist, dass zwischenzeitliche Änderungen an den Metadaten erst bei einer erneuten Übertragung in das Austauschverzeichnis berücksichtigt werden.

Bei der Bereitstellung einer API für das Harvesting erfolgt der Abruf der Metadaten synchron mit dem Datenbestand des Fach-

portals. Das heißt, dass immer der aktuelle Stand der Metadaten aus der Datenbank des Fachportals abgerufen wird. Zusätzlich erfolgt die Kommunikation zwischen Liefersystem und dem Harvester von open.rlp.de bidirektional in Form eines Request-Response-Zyklus.

Für den Fall, dass ein Liefersystem die Metadaten nicht als standardkonformes RDF/XML bereitstellen kann, verfügt open.rlp.de über eine Middleware. Eine Middleware ist ein Programm, das zwischen zwei Anwendungen vermittelt. Im Rahmen von open.rlp.de führt die Middleware ein Mapping des gelieferten Schemas auf den Metadatenstandard DCAT-AP.de durch. Diese Middleware kann somit für neue Liefersysteme

angepasst werden. Der dafür notwendige Entwicklungsaufwand ist vom anzubindenden Portal abhängig und muss im Vorfeld mit den Verantwortlichen von open.rlp.de abgestimmt werden. Der Vorteil dieses Verfahrens ist, dass die Integration der Metadaten in open.rlp.de damit standardisiert wird. Dennoch stellt die Verwendung der Middleware eine Ausweichmöglichkeit dar, weil ein Mapping immer das Risiko eines Datenverlustes beinhaltet und zusätzliche Felder, die nicht Bestandteil des Standards sind, nicht berücksichtigt werden können. Daher ist in jedem Fall die Bereitstellung der Datensätze in einer DCAT-AP.de-kompatiblen Form durch das jeweilige Liefersystem zu bevorzugen.

Application Programming Interface (API)

Über eine API (auch Programmierschnittstelle oder Endpunkt genannt) können Systeme oder Dienste unabhängig von ihrer Implementierung untereinander kommunizieren und Daten austauschen. Standards für APIs umfassen Sammlungen von Regeln, Protokollen und Best-Practices. Sie legen fest, wie Schnittstellen entwickelt, aufgebaut und verwendet werden sollen. Gängige Standards sind zum Beispiel REST oder SOAP.

Eine Webanwendung, die über einen API-Endpunkt verfügt, ermöglicht den automatisierten Abruf der bereitgestellten Daten und fördert insbesondere im Kontext von Open Data die Nachnutzung, da Anwendungen auf Basis der bereitgestellten Daten entwickelt werden können.

Zusätzlich können Schnittstellen auch schreibende Zugriffe ermöglichen, um eigene (Meta-)Daten in eine Anwendung zu übertragen. Diese Zugriffe sind authentifizierten Nutzenden vorbehalten, d.h. Nutzenden, die vom System als zugelassen identifiziert werden. Eine Authentifizierung erfolgt heute im Regelfall über einen API-Schlüssel. Dabei handelt es sich um eine alphanumerische Zeichenfolge, die von der Anwendung generiert wird, die die Schnittstelle bereitstellt. Bei der Übermittlung einer Anfrage wird der API-Schlüssel mitübertragen und das Zielsystem überprüft, ob der entsprechende Schlüssel im System bekannt ist.

API : Anlieferung von Metadaten

Kann kein geeigneter Endpunkt für ein Harvesting durch das Liefersystem bereitgestellt werden, ist in Ausnahmefällen eine Anlieferung von Daten über die standardisierte CKAN-API von open.rlp.de möglich.

Für die Anlieferung von Metadaten muss eine Berechtigung durch den Betreiber von open.rlp.de vorgenommen werden, damit ein Zugriff auf die notwendigen Funktionen des Servers von open.rlp.de erfolgen kann. Zusätzlich ist eine Authentifizierung erforderlich. Der Schlüssel wird bei Bedarf seitens des Betreibers von open.rlp.de zur Verfügung gestellt.

Eine Nutzung der API hat den Nachteil, dass die Verantwortung für eine regelmäßige Anlieferung beim Liefersystem liegt. Anpassungen bei open.rlp.de können zu einer notwendigen Anpassung der Schnittstelle des Liefersystems führen und damit Mehraufwände auf Seiten des Liefersystems verursachen. Schlimmstenfalls kann es zum temporären Ausfall der automatisierten Metadatenanlieferung kommen, wenn die Anpassungen nicht synchronisiert sind. Damit ist diese Anbindung an open.rlp.de in der Regel nicht präferiert.

Einstellsystem

Für den Fall, dass Sie Ihre Daten nicht über ein Fachportal bereitstellen können, steht außerdem das Einstellsystem (upload-tpp.rlp.de) bereit. Das Einstellsystem ermöglicht es allen Behörden der Landesverwaltung inklusive der Kommunen und anderen Akteuren, einzelne Datensätze händisch zu veröffentlichen, die dann per Harvester auf open.rlp.de übertragen werden. Derzeit ist dies ausschließlich für Transparenzinformationen möglich. Perspektivisch soll die

Plattform um ein Formular für offene Daten erweitert werden.

Ein Zugriff ist in der Regel nur aus dem Rheinland-Pfalz Netz möglich. In Ausnahmefällen kann auch ein Zugriff von außerhalb per Whitelisting ermöglicht werden.



WEITERE HINWEISE

Neben dem Open-Data-Handbuch stehen weitere Materialien und Möglichkeiten zur Verfügung, um sich zum Thema Transparenzinformationen und offene Daten zu informieren. Außerdem bestehen verschiedene Anlaufstellen für inhaltliche und technische Fragen.

- Die „[Open-Data-Strategie des Landes Rheinland-Pfalz](#)“ setzt den Handlungsrahmen für die Bereitstellung von offenen Daten durch die Landesverwaltung sowie für deren Nutzung.
- Die „[Digitalstrategie für das Land Rheinland-Pfalz](#)“ beschreibt das gesellschaftliche Zielbild für das Jahr 2030 und definiert Handlungsfelder und Quer-

schnittsthemen, die durch die Strategie gestaltet werden sollen.

- Bei datenschutzrechtlichen Fragen oder Unklarheit wenden Sie sich bitte an die Datenschutzbeauftragten Ihrer Behörde.
- Die inhaltliche Verantwortung obliegt der Abteilung 63 – Digitalisierung des MASTD. Nutzen Sie für Anfragen zu inhaltlichen Themen gerne unser Kontaktformular oder schreiben Sie eine E-Mail an cc-od@open.rlp.de.
- Die technische Verantwortung obliegt dem Landesbetrieb für Daten und Information. Neben dem Kontaktformular steht Ihnen für technische Anliegen die Mailadresse helpdesk@ldi.rlp.de zur Verfügung.

Bildnachweise

iStock/jacoblund (S. 6)
iStock/Wirestock (S. 11)
iStock/Jacob Wackerhausen (S. 13)
iStock/Paperkites (S. 21)
iStock/Moyo Studio (S. 23)
iStock/shapecharge (S. 24)
iStock/courtneyk (S. 27)
iStock/Delmaine Donson (S. 34)
iStock/puhhha (S. 35)
iStock/mixetto (S. 39)

Impressum

Herausgeber

Open-Data-Kompetenzzentrum Rheinland-Pfalz im
Ministerium für Arbeit, Soziales, Transformation
und Digitalisierung des Landes Rheinland-Pfalz

Bauhofstr. 9
55116 Mainz

E-Mail: cc-od@open.rlp.de

Gestaltung

]init[AG für digitale Kommunikation



Rheinland-Pfalz

MINISTERIUM FÜR ARBEIT,
SOZIALES, TRANSFORMATION
UND DIGITALISIERUNG