
Stellungnahme zum Gesetzentwurf für ein Hessisches Open-Data-Gesetz

im Ausschusses für Digitales und Datenschutz des
Hessischen Landtags

Yannick Bungers, Marco Holz

Chaos Computer Club Darmstadt e.V.



Chaos Computer Club
Darmstadt e.V.

Inhaltsverzeichnis

1	Einleitung	4
2	Offene Daten	5
2.1	Was sind offene Daten?	5
2.2	Warum offene Daten?	7
2.3	Wer profitiert von offenen Daten?	8
3	Zur Gesetzesinitiative	9
3.1	Was muss sich in Hessen ändern?	9
3.2	Bewertung des vorliegenden Gesetzentwurfs	10
4	Praktische Hilfestellungen zur Veröffentlichung von Datensätzen	12
4.1	Datenformate & Maschinenlesbarkeit	12
4.2	Standards	13
4.3	Lizenzen	14
4.4	Muster-Vertragsklauseln für kommunale Datenhoheit	15

Der Text dieser Stellungnahme ist lizenziert unter der CC0 1.0 Universell (CC0 1.0) Public Domain Dedication, abzurufen unter <https://creativecommons.org/publicdomain/zero/1.0/deed.de>.

1 Einleitung

Sucht man auf dem Informationsportal Hessen¹ nach dem Begriff „Open Data“, so findet man neben der Ankündigung des Beitritts Hessens zum bundesweiten Open-Data-Portal GovData² lediglich eine Pressemitteilung zu einem Forschungsprojekt zur Identifizierung optimaler Standorte für Windkraft- und Solaranlagen mit Hilfe offener Geodaten³ sowie einige Absichtserklärungen zur Stärkung von Open Data. Viel passiert ist zu diesem Thema in Hessen bisher nicht. Zwei Jahre nach dem Beitritt Hessens zum GovData-Portal finden sich dort immer noch keine Datensätze aus Hessen.⁴ Das Versprechen aus dem Koalitionsvertrag, die Daten des Haushaltsplans als offene Datensätze bereitzustellen, ist drei Jahre nach dessen Inkrafttreten noch nicht erfüllt. Auch ein eigenes Open-Data-Portal oder eine zentrale Anlaufstelle für offene Verwaltungsdaten vermisst man in Hessen.

Das Hessische Landesamt für Bodenmanagement und Geoinformation experimentiert dagegen sogar bereits mit der besonders fortschrittlichen Form der Datenbereitstellung als Linked Open Data.⁵ Das Ministerium für Soziales und Integration stellt auf einer eigens geschaffenen Webseite Informationen zu Hessischen Corona-Testzentren unkompliziert in einem maschinenlesbaren Format zum Download bereit.⁶ Die Verfolgung einer koordinierten Strategie, deren Relevanz bereits in der Digitalstrategie Hessen anerkannt wurde,⁷ ist hier allerdings bislang nicht zu erkennen.

Wir begrüßen ausdrücklich das überparteiliche Interesse zur Bereitstellung von offenen Daten und zur Schaffung einer entsprechenden gesetzlichen Grundlage. Neben der ebenfalls sehr begrüßenswerten Regelung zur Bereitstellung offener Geodaten im gerade diskutierten Geobasisdaten-Gesetz (Drs. 20/5810) befürworten wir auch die Schaffung einer gesetzlichen Grundlage für die Bereitstellung von Datensätzen durch Hess. Behörden.

Ein Open-Data-Gesetz mit proaktiven Veröffentlichungspflichten ist wichtig und kann dabei helfen, die Bereitstellung von offenen Datensätzen in Hessen voranzubringen. Eine Bereitstellung von Datensätzen als offene Daten scheitert in der Praxis in den Behörden häufig an fehlenden Erfahrungen und nicht vorhandenem Know-How. Neben einer gesetzlichen Grundlage braucht es für eine erfolgreiche Verwirklichung der Ziele daher weitere Schritte, auf die wir im folgenden ebenfalls eingehen möchten.

¹<https://www.hessen.de/>

²<https://digitales.hessen.de/press/pressarchiv/hessen-tritt-govdata-portal-bei>

³<https://digitales.hessen.de/pressemitteilungen/open-data-hilft-beim-erreichen-der-klimaschutzziele>

⁴<https://www.govdata.de/web/guest/daten>

⁵https://www.geoportal.hessen.de/linked_open_data/

⁶<https://www.corona-test-hessen.de/?action=Daten>

⁷<https://digitales.hessen.de/digitale-zukunft/digitalstrategie-hessen>

2 Offene Daten

2.1 Was sind offene Daten?

Daten, die zur Veröffentlichung als Open Data geeignet wären, gibt es in fast jeder öffentlichen Behörde. Bei der Definition des Begriffs der Offenheit von Daten lohnt sich ein Blick in die *Open Definition* der Open Knowledge Foundation.⁸ Die Open Knowledge Foundation listet auf ihrer Webseite einige Beispiele für Offene Daten:⁹

- Geodaten sind Daten, die einer räumlichen Lage auf der Erdoberfläche zugeordnet werden können. Neben den sog. Geobasisdaten, welche die Lage (Georeferenz) von Objekten beschreiben (z.B. Ortskoordinaten oder Adressen), beschreiben Geofachdaten Sachverhalte aus einem bestimmten Fachgebiet (z.B. Demographie, Epidemiologie, Bodenkunde, Klimatologie, Wahlstatistik). Geodaten werden häufig als Grundlage für die Erstellung themenspezifischer Karten genutzt und über ein Geoportal¹⁰ bereitgestellt.
- Kulturdaten beinhalten Informationen über kulturelle Werke und Artefakte (z.B. Titel und Künstler:innen). Diese werden häufig von Galerien, Bibliotheken, Archiven und Museen gesammelt und bereitgestellt.

BEISPIEL: Coding da Vinci

Der Kulturdaten-Hackathon *Coding da Vinci* beschäftigt sich mit der innovativen Nutzung von offenen Kulturdaten.

→ Weitere Informationen zu Coding da Vinci: <https://codingdavinci.de/de>

- Mobilitätsdaten sind zum Beispiel Verkehrsdaten des lokalen Nahverkehrs wie Fahrpläne, Haltestelleninformationen¹¹ und Echtzeit-Statistiken wie Live-Standorte von Leihrädern¹² oder multimodale Routingdaten¹³. Das am weitesten verbreitete Format für Fahrplandaten ist GTFS¹⁴. Die Bereitstellung von Mobilitätsdaten als Open Data ermöglicht z.B. die vereinfachte Nutzung durch Drittanbietersoftware.

⁸<https://opendefinition.org/od/2.1/de/>

⁹https://okfn.de/themen/open_data/

¹⁰<https://de.wikipedia.org/wiki/Geoportal>

¹¹z.B. <https://opendata.rmv.de/>

¹²z.B. <https://ulm.dev/projects/openbike/>

¹³z.B. <https://ulm.dev/projects/digitransit/>

¹⁴https://de.wikipedia.org/wiki/General_Transit_Feed_Specification

BEISPIEL: Fahrplan-Apps

Die Open-Source-Projekte Transportr und Öffi ermöglichen beispielsweise Fahrplanauskünfte in vielen europäischen und internationalen Verkehrsverbänden in einer einzigen App.

→ Transportr: <https://transportr.app/>

→ Öffi-App: <https://oeffi.schildbach.de/>

- Rats- bzw. Parlamentsinformationssysteme enthalten Daten der (lokal)politischen Gremienarbeit im Gemeinderat, Kreistag, Landtag oder Bundestag. Eine maschinenlesbare Bereitstellung von Daten aus parlamentarischen Informationssystemen ermöglicht der OParl-Standard¹⁵.

BEISPIEL: Parlamentsinformationen

Projekte wie *Politik bei uns* nutzen die OParl-Schnittstelle der bestehenden Ratsinformationssysteme, um die darin befindlichen öffentlichen - und nur die öffentlichen - Daten der Kommune abzurufen und aufzubereiten. Auch das Informationssystem des Bundestages (DIP) stellt seit diesem Jahr eine Schnittstelle (API) zum Datenabruf bereit.

→ OParl-Standard: <https://oparl.org/>

→ Politik bei uns: <https://politik-bei-uns.de/>

→ API des Bundestag: <https://dip.bundestag.de/über-dip/hilfe/api>

BEISPIEL: kleineAnfragen.de

Das ehrenamtliche Projekt *kleineAnfragen.de* stellte über 5 Jahre lang die kleinen und großen Anfragen der Landesparlamente und des Bundestages zusammen und veröffentlichte sie in maschinenlesbarer Form, wurde jedoch aufgrund fehlender Schnittstellen und einer schlechten technischen Umsetzung der Informationssysteme der Landesparlamente aus Kapazitätsgründen eingestellt.

→ Zum Projekt: <https://kleineanfragen.de/info/stilllegung>

¹⁵<https://oparl.org/>

BEISPIEL: Open Legal Data

Das ebenfalls ehrenamtliche Projekt *Open Legal Data* stellt Gesetzestexte und Urteile für alle kostenlos und maschinenlesbar zur Verfügung. Die freie juristische Datenbank veröffentlicht diese Dokumente über eine Schnittstelle (API) und ermöglicht so unter anderem wissenschaftliche Forschung zu Gesetzestexten.

→ Zum Projekt: <https://de.openlegaldata.io/>

- Finanzdaten sind beispielsweise Haushaltsdaten der öffentlichen Hand (Ausgaben, Einnahmen) oder Informationen zu Finanzmärkten (Aktien, Anleihen).

BEISPIEL: Offener Haushalt

Das Projekt *Offener Haushalt* hat aufgezeigt, wie offene Haushaltsdaten in einem einheitlichen Format bereitgestellt und visualisiert werden können.

→ Zum Projekt: <https://offenerhaushalt.de/>

- Wissenschaftliche Daten werden zu Forschungszwecken erhoben, generiert und verwendet. Die meisten Universitäten und Hochschulen betreiben ein eigenes Forschungsdatenmanagement. In Hessen werden diese Aktivitäten durch die Landesinitiative „Hessische Forschungsdateninfrastrukturen“ (HeFDI) koordiniert¹⁶.

Zur Beschreibung von Datensätzen selbst dienen Metadaten. Diese können Informationen zu Struktur, Datenformat, Herkunft, Zeitbezug und Entstehungszeitpunkt beinhalten. Auf dem Datenportal GovData finden sich Informationen zum standardisierten Metadatenmodell *DCAT-AP.de* für offene Daten der öffentlichen Verwaltung in Deutschland.¹⁷

2.2 Warum offene Daten?

Daten sind eine wichtige Grundlage für viele digitale Anwendungen. Die Verfügbarkeit offener und lizenzkostenfreier Daten ist gerade für Organisationen und Firmen mit geringen finanziellen Mitteln wichtig. Um in der Lage zu sein neue Anwendungsfelder und Nutzen zu erschließen, ist es deswegen

¹⁶<https://www.forschungsdaten.info/fdm-im-deutschsprachigen-raum/hessen/hefdi/>

¹⁷<https://www.govdata.de/web/guest/metadaten-schema>

von besonderer Wichtigkeit, dass diese zum Experimentieren und Verwenden unmittelbar zu Verfügung stehen. Müssen zu diesem Zweck erst alle Verwaltungsstrukturen durchlaufen werden, die Verwaltung davon überzeugt werden, dass die Daten auch wirklich benötigt werden und dann die Bereitstellung umgesetzt werden, ist eine sehr langsame Entwicklung vorprogrammiert. Dadurch sind Innovationen nur schwer möglich und Konzepte können nur aus anderen Regionen mit weniger Restriktionen auf den Daten übernommen werden.

Die von der Hess. Verwaltung erhobenen Datensätze sind **von der Gesellschaft finanziert** und zum Wohle der Allgemeinheit erhoben. Daher sollte es eigentlich eine Selbstverständlichkeit sein, dass diese Daten analog zum Prinzip „Öffentliches Geld - Öffentliches Gut“¹⁸ auch der Gesellschaft zur Verfügung gestellt werden.

Daten sind nicht nur **wichtige Informationsquellen** für interessierte Bürger:innen, sondern genauso wichtig für Akteure aus Politik, Wissenschaft und Wirtschaft: Sie bieten eine Grundlage für fundierte, faktenbasierte Entscheidungen und sind damit ein elementarer Bestandteil einer lebendigen Demokratie.

2.3 Wer profitiert von offenen Daten?

Eine Bereitstellung von Daten ist eine besonders nachhaltige Form der Verwaltungsdigitalisierung. Statt zeitlich begrenzten Förderprojekten oder der aufwändigen Entwicklung und Pflege von Dashboards zur Datenvisualisierung ermöglicht Open Data die Auswertung, Visualisierung und Verknüpfung von Datensätzen durch Dritte. So können die **Kapazitäten der Verwaltung möglichst effizient genutzt** werden.

Auch die Verwaltung selbst profitiert von der Bereitstellung von offenen Daten. Durch eine bessere Auffindbarkeit und Abrufbarkeit von Datensätzen sinken Koordinierungs- und Abstimmungsaufwände zur Datennachnutzung innerhalb der Verwaltung enorm.

Die Bereitstellung von offenen Daten bricht zudem die herrschende **Bevorteilung großer Digitalkonzernen** im Umgang mit Daten auf. Wenn Großkonzerne an der Nutzung bestimmter Datensätze interessiert sind, spielen Kosten für einen exklusiven Datenzugang meist keine Rolle. Kleinere Unternehmen, Startups und die Zivilgesellschaft können sich entsprechende Deals dagegen nicht leisten.

Ein konkretes Beispiel für eine solche Bevorzugung von international agierenden IT-Konzernen ist der exklusive Deal zwischen Google und der Deutschen Bahn AG zu einem exklusiven Zugang zu Echtzeitdaten der Bahn.^{19 20} Statt diese ohnehin maschinenlesbar vorliegenden Daten für alle zur

¹⁸<https://www.wikimedia.de/oeffentliches-gut/>

¹⁹https://www.deutschebahn.com/de/presse/pressestart_zentrales_uebersicht/Google-und-DB-zeigen-Live-Auskuenfte-zu-Zuegen-und-vereinfachen-Ticketbuchung-6294046

²⁰<https://blog.google/intl/de-de/produkte/suchen-entdecken/google-maps-macht-reiseplanung-fur-bahnfahrende-noch-einfacher/>

Verfügung zu stellen, verkauft die Bahn diese Daten zum Leid von Startups und der Zivilgesellschaft an die meistbietenden IT-Großkonzerne.²¹

3 Zur Gesetzesinitiative

3.1 Was muss sich in Hessen ändern?

Neben der im Gesetzentwurf geforderten Etablierung von Ansprechpersonen, sollten unbedingt auch **unterstützende Prozesse zur Datenbereitstellung** etabliert werden. Um eine gewisse Kontinuität bei der Datenbereitstellung zu garantieren und personelle Ressourcenengpässe zu vermeiden, sollten Daten nicht händisch, sondern möglichst vollständig automatisiert bereitgestellt werden können. Hierzu ist eine Weiterentwicklung bestehender IT-Systeme nötig, die beispielsweise über Fördergelder für interkommunale Zusammenarbeit (IKZ) finanziert werden könnte. Die Finanzierung von IT-Basisinfrastruktur ist eine der besten Maßnahmen zur Effizienzsteigerung der öffentlichen Verwaltung. Eine **automatisierte Bereitstellung von Verwaltungsdaten** schafft zudem die Grundlage für eine Reihe von weiteren datengetriebenen Anwendungen.

Die Benennung von Open-Data-Koordinator:innen in den Ministerien ist ein erster Schritt hin zu einem koordinierten Datenmanagement und kann zudem einen ressortübergreifenden Wissens- und Erfahrungsaustausch stärken. Entscheidend für den Erfolg der Open-Data-Koordinator:innen ist neben ausreichenden zeitlichen Ressourcen für die Unterstützung bei Anfragen aus dem eigenen Haus auch die Vernetzung untereinander. Aufgrund des hohen Interesses und des enormen Erfahrungswissens in der deutschen Open-Data-Szene empfiehlt sich insbesondere auch ein **gegenseitiger Erfahrungsaustausch mit zivilgesellschaftlichen Akteur:innen**.

Zur Steigerung der Datenkompetenz (Data Literacy²²) innerhalb der Hess. Verwaltung könnte die Etablierung einer **Hess. Serviceeinheit für Open Data** ein geeignetes Mittel sein. Ein solche Serviceeinheit könnte ressortübergreifend bei der Bewertung, Aufbereitung, Bereitstellung und Interpretation von Daten sowie bei urheberrechtlichen Fragestellungen unterstützen, aber auch eine erste Anlaufstelle für Kommunen sein. Entscheidend für den Erfolg einer solchen Organisation ist jedoch, dass diese nicht nur eine koordinierende Rolle einnimmt, sondern auch selbst **mit tiefem fachlichem und technischem Know-How** die Bearbeitung von Herausforderungen bei der Veröffentlichung von Daten übernimmt und **aktiv geeignete Lösungen erarbeitet und umsetzt**. Neben der Bereitstellung von **Tools und Ressourcen zur Veröffentlichung von Daten in maschinenlesbaren Formaten** könnten **Schulungsangebote für die Landesverwaltung und Kommunen** sowie die Kollaboration mit zivilgesellschaftlichen Akteur:innen zu den Kernaufgaben der Organisationseinheit zählen. Vorbild könnte

²¹<https://www.handelsblatt.com/technik/it-internet/kooperation-mit-google-digitalisierung-auf-kosten-von-start-ups-gruender-sauer-auf-die-deutsche-bahn/27435250.html?ticket=ST-3627762-mefktBhFrhgegsa7fzj4-ap6>

²²<https://de.wikipedia.org/wiki/Datenkompetenz>

beispielsweise das TUdata-Team der Technischen Universität Darmstadt²³ sein, dass auf dem Gebiet offener Forschungsdaten wissenschaftliche Projekte bei der Veröffentlichung von Daten berät und unterstützt, aber auch eigene Softwarelösungen zur Unterstützung dieses Prozesses entwickelt.

Die **Gewinnung von hochqualifiziertem IT-Personal** stellt nicht nur Kommunen vor große Herausforderungen,²⁴ sondern wird auch eine zentrale Herausforderung beim Aufbau einer solchen Unterstützungsorganisation sein. Um für fachlich qualifiziertes Personal attraktiv zu werden, braucht es entsprechende Finanzierung, aber auch nicht-monetäre Anreize, wie beispielsweise weitreichende Gestaltungsspielräume und Mut zur Veränderung.

3.2 Bewertung des vorliegenden Gesetzentwurfs

Alles in allem bietet der vorliegende Gesetzentwurf eine gute Grundlage für ein Hess. Open-Data-Gesetz. Im folgenden Abschnitt möchten wir auf besonders begrüßenswerte Aspekte des vorliegenden Gesetzentwurfs eingehen und einige Empfehlungen zu dessen Fortschreibung geben.

Begrüßenswert und essentiell für den Erfolg des Gesetzes ist insbesondere die Verpflichtung zur proaktiven Veröffentlichung aller neu erhobener Datensätze. Von den positiven Nebeneffekten einer Veröffentlichung von Datensätzen (z.B. die Etablierung einheitlicher Qualitätsstandards zum Datenmanagement durch Anreicherung um Metadaten und etablierten Prozessen zur Sicherstellung der Datenverfügbarkeit, einer erhöhten Auffindbarkeit und besseren Nachnutzbarkeit durch andere Behörden) profitiert nicht nur die datenbereitstellende Behörde, sondern letztlich auch die gesamte Hess. Verwaltung. Innovative Datennutzungen durch Wissenschaft, Startups, journalistische Medien und die Zivilgesellschaft erfordert Kreativität. Eine selektive Auswahl von Datensätzen durch die Hess. Verwaltung (wie bisher bereits möglich) ist dagegen nicht zielführend.

Auch die Regelung zur entgeltfreien und uneingeschränkten Weiterverwendung ohne verpflichtende Registrierung in § 4a Abs. 5 zur Erfüllung des Ziels einer breiten Verfügbarkeit und Nutzbarkeit von offenen Daten sind sehr zu begrüßen.

Ebenfalls begrüßenswert sind die Regelungen aus § 4a Abs. 6 zur maschinenlesbaren Bereitstellung in offenen Formaten über Programmierschnittstellen (APIs) sowie die Bereitstellung von Metadaten auf einem Open-Data-Portal des Landes sowie dem nationalen Metadatenportal GovData. Die getroffene Regelung orientiert sich an den aus dem wissenschaftlichen Forschungsdatenmanagement bekannten FAIR-Data-Prinzipien²⁵.

Eine Umsetzung des Open-Data-Gesetzes im Hess. E-Government-Gesetz (HEGovG) nimmt einige Institutionen von den Regelungen zur Bereitstellung von offenen Daten aus. Insbesondere nicht im

²³https://www.tu-darmstadt.de/tudata/tudata/digitale_forschungsdaten_an_der_tu/index.de.jsp

²⁴vgl. <https://www.chaos-darmstadt.de/2020/Stellungnahme-zum-Kommunalbericht-2019.html>

²⁵https://www.ruhr-uni-bochum.de/researchdata/de/fair_data.html

Geltungsbereich des HEGovG sind beispielsweise die Wasser- und Bodenverbände nach dem Wasserverbandsgesetz. Die von diesen Organisationen erfassten Umweltdaten sind als Grundlage für viele spannende Open-Data-Projekte und -Anwendung von besonderem Wert. Wünschenswert wäre daher eine explizite Aufnahme dieser Organisationen in den Geltungsbereich des HEGovG in Bezug auf die neu geschaffenen Paragraphen § 4a und § 4b.

Kritisch zu sehen ist auch die vollständige Ausnahme von Kommunen aus der Pflicht zur Bereitstellung von offenen Daten in § 4a Abs. 1. Eine ausführliche Begründung für die Relevanz offener Daten für Kommunen haben wir bereits im vergangenen Jahr in unserer Stellungnahme zum Kommunalbericht dargelegt.²⁶ „Landesbehörden und Kommunen sollen zunehmend Daten in einem nutzenorientierten Maß im GovData-Portal bereitstellen und der Öffentlichkeit zugänglich machen“, heißt es hierzu in der Digitalstrategie Hessen der Landesregierung. Dieses Ziel wird mit dem vorliegenden Gesetzentwurf in Bezug auf kommunale Datenbereitstellung nicht zu erfüllen sein. Wünschenswert und zielführend wäre stattdessen eine Verpflichtung von Kommunen zur Veröffentlichung aller neu erhobenen Datensätze; mindestens jedoch die Verpflichtung zur Veröffentlichung von Daten im Gültigkeitsbereich des Gesetzes durch formlosen Antrag.

Die in § 4a Abs. 2 getroffene Einschränkung auf Daten, die „außerhalb der Behörde liegende Verhältnisse betreffen“ ist zu pauschal. OpenData sollte auch als Instrument für Haushaltskennzahlen und Performance-Indikatoren für die Verwaltung verstanden werden, etwa bei der Umsetzung von OZG-Prozessen. Statt der genannten Ausnahme bietet es sich stattdessen an, die in der Gesetzesbegründung genannten konkreten Ausnahmeregelungen in Abs. 3 aufzunehmen.

In Bezug auf § 4a Abs. 3 erscheint die Übernahme der Ausnahmeregelung „soweit ein rein wirtschaftliches Interesse an den Informationen besteht“ aus § 83 des Hessischen Datenschutz- und Informationsfreiheitsgesetzes (HDSIG) zur Erfüllung des Ziels einer breiten Verfügbarkeit und Nutzbarkeit von offenen Daten nicht angemessen. Dies widerspricht dem erklärten Ziel der Förderung neuer, datengetriebener Geschäftsmodelle. Zudem stellt sich in diesem Zusammenhang die Frage, wie eine solche Bewertung von Datensätzen im Vorhinein möglich sein soll.

Die Übernahme der Regelungen aus § 84 HDSIG zum Schutz behördlicher Entscheidungsprozesse erscheint im Kontext von Datensätzen nicht angemessen und sollte daher ebenfalls gestrichen werden. Wie auch Ergebnisse der Beweiserhebung sowie Gutachten oder Stellungnahmen Dritter gemäß § 84 HDSIG stellen erhobene Datensätze keine unmittelbare Vorbereitungen einer Entscheidung dar. Alternativ wäre insofern auch eine entsprechende Ausnahmeregelung für erhobene Datensätze in § 84 Abs. 1 Satz 2 HDSIG denkbar.

Die Einschränkung des § 4a Abs. 8 auf Landesbehörden nimmt Kommunen von der Verpflichtung zur Optimierung von Verwaltungsabläufen sowie von der Berücksichtigung von Anforderungen an das Bereitstellen von Daten beim Abschluss vertraglicher Regelungen und der Beschaffung von informa-

²⁶<https://www.chaos-darmstadt.de/2020/Stellungnahme-zum-Kommunalbericht-2019.html>

tionstechnischen Systemen explizit aus. Hier wird die Chance vertan, Kommunen zur nachhaltigen Digitalisierung anzuhalten.²⁷

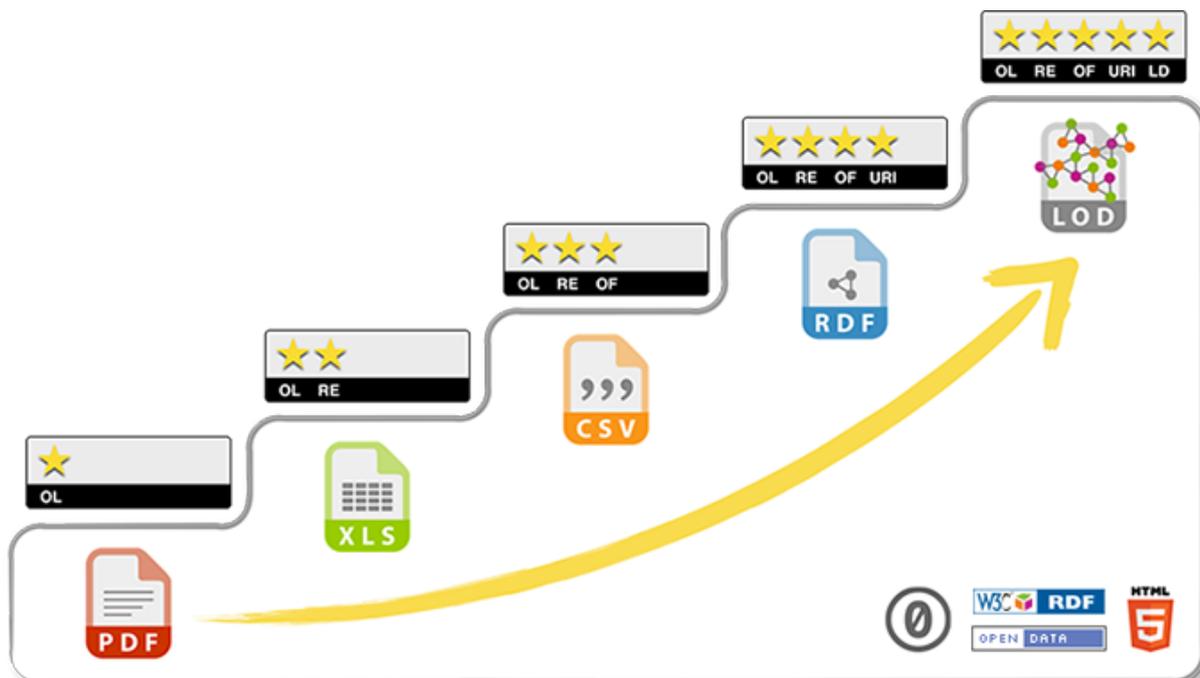
Eine Veröffentlichung eines Open-Data-Berichts ein bis zweimal pro Legislaturperiode erscheint wenig ambitioniert. Im Hinblick auf § 4b Abs. 1 ist zu empfehlen, den Veröffentlichungszyklus des neu geschaffenen Open-Data-Berichts von drei Jahren auf eine jährliche Veröffentlichung zu verkürzen.

4 Praktische Hilfestellungen zur Veröffentlichung von Datensätzen

Im Folgenden möchten wir abschließend noch einige wissenswerte Informationen zum Thema Open Data und Hinweise zur Veröffentlichung von Daten geben.

4.1 Datenformate & Maschinenlesbarkeit

Ein Schema zur Bewertung der Maschinenlesbarkeit von offenen Daten bietet das 5-Sterne-Modell von Tim Berners-Lee.²⁸ Im Folgenden möchten wir auf die unterschiedlichen Qualitätsniveaus bei der Datenbereitstellung eingehen.



Eine Bereitstellung von Daten unter einer offenen Lizenz in einem unstrukturierten Format (beispielsweise die Veröffentlichung von gescannten Tabellen als PDF) stellt Datennutzende vor große Heraus-

²⁷vgl. <https://www.chaos-darmstadt.de/2020/Stellungnahme-zum-Kommunalbericht-2019.html>

²⁸<https://5stardata.info/de/>

forderungen. Diese Daten sind sehr einfach bereitzustellen, allerdings erfordert es einen sehr hohen, oft manuellen, Aufwand, um eine Daten in ein maschinenlesbares Format zu transformieren, um sie anschließend weiterverarbeiten zu können. Eine Datenbereitstellung in einem unstrukturierten Format entspricht der ersten von insgesamt fünf Stufen des 5-Sterne Open-Data-Modells.

Die zweite Stufe des Modells entspricht der Bereitstellung in einem strukturierten, maschinenlesbaren, aber proprietären, d.h. herstellerebenen, Datenformat (z.B. Excel oder nicht standardisierte Geodatenformate). Eine Weiterverarbeitung ist meist nur mit Software eines bestimmten Herstellers möglich. Ein Beispiel für die Veröffentlichung von Daten in einem proprietären Datenformat stellen beispielsweise die Tabellen im Hess. Statistikportal als Excel-Dokumente dar.²⁹

Mit der Bereitstellung von Daten in einem offenen, strukturierten Datenformat wie beispielsweise CSV, JSON oder XML wird die dritte Stufe des 5-Sterne-Modells erreicht. Ein Beispiel hierfür stellt das Baumkataster der Stadt Frankfurt dar.³⁰ Mit der Bereitstellung in einem offenen Format, entspricht der Datensatz den Kriterien der *Open Definition* der Open Knowledge Foundation.³¹ Diese Stufe sollte den Mindeststandard für die Bereitstellung offener Daten darstellen.

Werden Datensätze unter einer eindeutigen, möglichst stabilen (d.h. zeitlich unveränderlichen) Internetadresse (URL) verfügbar gemacht, dann wird die nächste Stufe des 5-Sterne-Modells erreicht. Als Infrastruktur für eine derartige Datenbereitstellung können Open-Data-Portale dienen.

Von verlinkten Daten, auch Linked Open Data genannt, spricht man, wenn verschiedene Datensätze untereinander durch Links semantisch verknüpft werden. Das Wikipedia-Schwesterprojekt Wikidata hat sich zum Ziel gesetzt, auf diese Weise das gesamte Wissen der Welt strukturiert zugänglich zu machen. Dies eröffnet sehr weitreichende Möglichkeiten der Datenauswertung. Beispielsweise lässt sich über den maschinenlesbaren Eintrag der Stadt Wiesbaden³² vollständig automatisiert herausfinden, welche beiden Flüsse sich in Wiesbaden befinden. Über die verknüpften Datensätze sind wiederum Fotos der Flüsse, die Größe deren Wassereinzugsgebiete sowie Zuflüsse aus weiteren Flüssen maschinenverständlich abrufbar. Entspricht ein Datensatz diesen Kriterien, erreicht er die höchste Stufe des 5-Sterne-Modells für offene Daten.

4.2 Standards

Um eine Heterogenität von Datenformaten zu vermeiden und auch Datenbestände unterschiedlicher Behörden oder Kommunen einfach zusammenzuführen und in behördenübergreifenden Diensten weitzunutzen zu können, ist es von besonderer Wichtigkeit, dass Daten in offenen und einheitlichen Standards vorliegen. Hier ist nicht nur das Datenformat (JSON, CSV, XML) von Bedeutung, sondern

²⁹<https://statistik.hessen.de/zahlen-fakten/bevoelkerung-gebiet-haushalte-familien/bevoelkerung/tabellen>

³⁰<https://www.offenedaten.frankfurt.de/dataset/baumkataster-frankfurt-am-main>

³¹<https://opendefinition.org/od/2.1/de/>

³²<https://www.wikidata.org/wiki/Q1721>

auch eine einheitliche Benennung und Struktur von Datenfeldern. Ein wichtiges Anliegen einer koordinierenden Stelle sollte es sein, bereits existierende Standards zu kennen, wenn nötig auf sie zu verweisen und bei Bedarf ggf. auch neue Standards in Zusammenarbeit mit mehreren Beteiligten Stellen zu entwickeln. Für viele Arten von Daten bestehen bereits offene Standards, die verwendet werden können. Hier sind zum Beispiel Initiativen wie OParl zu nennen, die sich für offene Standards für die Ratsinformationssystemen in Kommunen einsetzen.

4.3 Lizenzen

In Bezug auf die Festlegung von Nutzungsbedingungen auf Grundlage des Urheberrechts ist zu bedenken, dass für reine Faktendaten aufgrund der nicht vorhandenen „persönlichen geistigen Schöpfung“ (vgl. § 2 UrhG) keine Rechtsgrundlage zur Einschränkung einer Datennutzung vorliegt. Faktendaten beschreiben reale Sachverhalte und basieren daher nicht auf einer persönlichen, kreativen Leistung, sondern stellen ein Abbild der Realität ohne Spielraum für kreative Auslegung dar. Entsprechende Datensätze sind damit keine Werke im Sinne des Urheberrechts und sind somit als gemeinfrei einzustufen. Gleiches gilt für amtliche Werke gemäß § 5 UrhG.

Zur Kennzeichnung der Gemeinfreiheit von Datensätzen eignet sich die „CC0 1.0 Universell“-Verzichtserklärung (CC0 1.0)³³, wie beispielsweise die Stadt Bonn zeigt.³⁴ Auch der Verein Open Knowledge International empfiehlt die Veröffentlichung als CC0, u.a. zur Vermeidung von Datensilos.³⁵ Hinweise zur Kennzeichnung der Gemeinfreiheit von Daten liefert das Fact Sheet „Creative Commons & Open Science“ von Creative Commons UK.³⁶

Beinhalten die zu veröffentlichen Datensätze neben reinen Faktendaten auch urheberrechtlich geschützte Werke (beispielsweise Texte, Bilder, Musikstücke oder Videoclips), so empfiehlt sich eine Veröffentlichung der Werke unter der international anerkannten „Creative-Commons Namensnennung 4.0 International“-Lizenz (CC BY 4.0)³⁷. Die Creative-Commons-Lizenzfamilie ist ab Version 2 in einer offiziellen deutschen Übersetzung vorhanden und in Version 4 insbesondere auf die urheberrechtlichen Besonderheiten von Datenbankwerken angepasst. Zudem ist diese international anerkannte Lizenzfamilie (auch) auf das europäische Urheberrecht angepasst und damit ideal zur Verwendung in Deutschland geeignet. Gleichzeitig erleichtert die Creative-Commons-Lizenzfamilie eine internationale Nutzung der bereitgestellten Werke.

Detaillierte Antworten zu Fragestellung rund um Datenlizenzierung gibt der Leitfaden „Rechtsfragen bei Open Science“ von Till Kreutzer und Henning Lahmann.³⁸

³³<https://creativecommons.org/publicdomain/zero/1.0/deed.de>

³⁴<https://twitter.com/OpenDataBonn/status/1426071748085227523>

³⁵<https://research.okfn.org/avoiding-data-use-silos/>

³⁶<https://zenodo.org/record/840652>

³⁷<https://creativecommons.org/licenses/by/4.0/deed.de>

³⁸<https://blogs.sub.uni-hamburg.de/hup/products-page/publikationen/169/>

4.4 Muster-Vertragsklauseln für kommunale Datenhoheit

Datenverfügbarkeit und die Möglichkeiten der Datennutzung sind zentrale Smart-City-Herausforderungen und hängen wesentlich von den getroffenen Vereinbarungen mit IT-Dienstleistern ab. Fehlende Vereinbarungen zu Datenstandards und Schnittstellen für einen Datenzugriff in Verträgen mit Fachverfahrensherstellern oder anderen externen Dienstleistern können die Möglichkeiten der Kommunen zur Datennutzung einschränken.

Die Städte Bonn und Münster haben hierzu gemeinsam Muster-Vertragsklauseln entwickelt, welche als Bestandteil für Vergabe-Leistungsverzeichnisse bei der IT-Beschaffung und Vertragsergänzung für z.B. EVB-IT-Verträge genutzt werden können.³⁹

³⁹<https://github.com/od-ms/datennutzungsklauseln-muster>