



Wikimedia Deutschland e. V., Tempelhofer Ufer 23-24, 10963 Berlin

Rheinland-Pfalz
Ministerium für Arbeit, Soziales, Transformation und Digitalisierung
per E-Mail

Stellungnahme

des Wikimedia Deutschland, Gesellschaft zur Förderung Freien Wissens e. V. zum Entwurf eines Open-Data-Gesetz Rheinland-Pfalz – ODGRP¹

Berlin, den 20. Juni 2025

Wikimedia Deutschland setzt sich für Chancengleichheit beim Zugang zu Wissen und Bildung ein und unterstützt die vielen Tausend Ehrenamtlichen, die in der Wikipedia und ihren Schwesterprojekten tagtäglich ihr Wissen mit allen Menschen teilen. Wiederverwendbare Dokumente und Daten, nicht zuletzt der öffentlichen Hand, sind ein wichtiger Beitrag zur Sicherung Freien Wissens. Wir danken für die Einbeziehung im Rahmen des Anhörungsverfahrens und nehmen zum Entwurf eines Open-Data-Gesetz Rheinland-Pfalz (im Folgenden auch ODGRP) wie folgt dargelegt Stellung.

In Stichpunkten

- ☛ Eine Präzisierung der gesetzlichen Regelungen für eine bessere Bereitstellung von Dokumenten im Besitz öffentlicher Stellen ist prinzipiell begrüßenswert. Die konkrete Umsetzung halten wir jedoch für verbesserungswürdig und in der vorgeschlagenen Weise nicht für zielführend.
- ☛ Aufbau und Inhalt des Gesetzesentwurfs folgen den Formulierungen des EGovG und den ihm seither folgenden Landesgesetzen – samt der darin enthaltenen, mittlerweile erkannten Schwächen und Defizite. Eine Schärfung, insbesondere im

¹ Bei Wikimedia Deutschland e.V. eingegangen per E-Mail am 2. Juni 2025. Eine öffentlich einsehbare Fassung ist im Hause nicht bekannt.

**Wikimedia Deutschland –
Gesellschaft zur Förderung
Freien Wissens e. V.**

Tempelhofer Ufer 23 – 24
10963 Berlin

Postfach 61 03 49
10925 Berlin

Tel.: +49 (0) 30 577 11 62 - 0
Fax: +49 (0) 30 577 11 62 - 99

wikimedia.de
politik@wikimedia.de

Spendenkonto:
IBAN: DE05 1002 0500 0003
2873 00

BIC: BFSWDE33XXX

Geschäftsführende Vorständin:
Franziska Heine

Eingetragen beim Amtsgericht
Berlin-Charlottenburg unter der
Nummer 23855 B

Hinblick auf die aktuelle Fassung der Richtlinie (EU) 2019/1024 (im Folgenden: ODD), ist nur teilweise zu erkennen.

- ☛ Positiv ist, dass der Entwurf nicht nur die klassischen Open-Data-Argumente hinsichtlich eines möglichen monetären Nutzens für die Wirtschaft sowie der Stärkung von Transparenz und Vertrauen in den Staat hervorhebt, sondern auch das stärkste Argument für Open Data würdigt: Den direkten und strategischen Nutzen für die öffentliche Verwaltung selbst. Dieser Aspekt verlangt jedoch nach einer Entsprechung im Gesetz, was aktuell leider nicht erkennbar ist. Der Gesetzentwurf lässt aus unserer Sicht die konkreten Schritte vermissen, diesen Nutzen auch praktisch zu erreichen.
- ☛ Die Begriffsdefinitionen sollten durchgehend überarbeitet und an die der ODD angeglichen werden. Im aktuellen Entwurf haben sie das Potenzial, die Bereitstellung offener Daten in der Praxis eher zu erschweren als zu fördern.
- ☛ Der vorgeschlagene Ansatz stärkt aus unserer Perspektive weder die externe noch die interne Wiederverwendbarkeit staatlicher Informationen. Wir befürchten, dass er in der Praxis vor allem Mehraufwände verursachen wird, während der positive Nutzen fraglich bleibt.
- ☛ Nicht nachvollziehbar ist die Unterscheidung zwischen „Dokumenten“ im Sinne des LTranspG und „Daten“, die durch das geplante Gesetz getroffen werden soll. Diese Unterscheidung steht im Widerspruch zur ODD und birgt entgegen der Gesetzesbegründung das Potenzial konkurrierender Auslegung, welche Norm nun einschlägig sein könnte.
- ☛ Wir halten den Entwurf in der vorliegenden Form für nicht empfehlenswert und raten stattdessen zu einer Überarbeitung des LTranspG, um dort den aktuellen Stand der ODD abzubilden. Um den direkten Nutzen für die Verwaltung zu fördern, sollten die geplanten Prozesse das tatsächliche Verwaltungshandeln unterstützen und dazugehörige Organisationspflichten festgelegt werden.

Im Einzelnen:

Zusammenspiel mit dem LTranspG und der Richtlinie (EU) 2019/1024, Begrifflichkeiten

Die ODD bezieht sich auf Dokumente, definiert als „jede[r] Inhalt unabhängig von der Form des Datenträgers“ oder „einen beliebigen Teil eines solchen Inhalts“. Diese Definition bezieht unter Berücksichtigung der Erwägungsgründe offenkundig explizit beliebige „Datensätze“ ein – hier nicht zwischen „Daten“ und „anderen Dokumenten“ zu unterscheiden ist aus praktischer Sicht nur richtig und konsequent.

Das LTranspG ist leicht diffuser und spricht von „amtliche[n] Informationen und Umweltinformationen, unabhängig von der Art ihrer Speicherung“ (Hervorhebung durch Wikimedia Deutschland). Es läge nahe, diese Formulierung im Kontext der früheren

Richtlinie 2003/98/EG (PSI-Richtlinie) so zu interpretieren, dass auch hiermit „Datensätze“ gleich welcher Art gemeint sind.

Der Gesetzesentwurf führt in „C. Alternativen“ an, dass sich die unaufgeforderte Veröffentlichung von Dokumenten als Folge des LTranspG lediglich auf Dokumente in für Menschen lesbarer und verständlicher Form beziehe und hier eine Regelungslücke bestehe. Dieser Gedankengang ist für uns nicht nachvollziehbar. Wir sehen durchaus, dass die Formulierungen im LTranspG insbesondere durch die Aufzählung in § 7 LTranspG viele Potenziale für die unaufgeforderte Veröffentlichung weiterer Dokumente im Besitz öffentlicher Stellen ungehoben lassen. Dieses Defizit ist unserer Meinung nach durch eine Aktualisierung des LTranspG deutlich sinnvoller zu beheben und würde eine ganzheitlichere Betrachtung der gewünschten Ziele erleichtern.

Eine separate gesetzliche Regelung durch ein Open-Data-Gesetz würde zudem die falsche Annahme stärken, dass es sich bei Open Data um ein Feld handle, das isoliert von einer Gesamtstrategie für die Verwaltungsdigitalisierung betrachtet und praktisch umgesetzt werden kann. Die fehlenden Organisationspflichten bis zu einem fehlenden Rechtsanspruch auf Zugang zu diesen Informationen bestärken den Eindruck, dass sich diese Annahme auch im vorliegenden Entwurf widerspiegelt.²

Auffällig ist, dass die Begriffsdefinitionen in den §§ 2, 3, 4 und 7 des Entwurfs inkonsistent sowohl zu den Erläuterungen in der Begründung, als auch zum LTranspG und der ODD sind:

Der Begriff „Daten“ in § 3 Abs. 1 ODGRP ist ausweislich der Gesetzesbegründung an die Begriffsbestimmung aus § 3 Nr. 3 des DNG angelehnt, der wiederum die nationale Umsetzung der Begrifflichkeiten aus der ODD vornehmen soll. § 5 LTranspG scheint dagegen viel näher an der Begrifflichkeit der PSI-Richtlinie und in der Folge auch der ODD orientiert zu sein. Warum bei Informationen im Besitz öffentlicher Stellen zwischen „Daten“ und „Nicht-Daten“ unterschieden werden soll, bleibt ungewiss – zumal nach der Definition des Gesetzesentwurfs auch die vorwiegend für die menschliche Interpretation gedachten Dokumente qua dieser Definition als „Daten“ einzuordnen sind, sobald sie digital gespeichert sind. Das angegebene Ziel einer Unterscheidbarkeit von „Daten“ und „Nicht-Daten“ kollidiert bereits hier mit der Unschärfe der Begriffsdefinitionen.

Der Begriff „Offen“ in § 3 Abs. 2 ODGRP verzichtet auf die Klarstellung, dass die Wiederverwendung zu jedem Zweck gestattet ist und hierfür anerkannte, standardisierte Lizenzen verwendet werden sollen. Die Gesetzesbegründung präzisiert dies. Es ist jedoch nicht nachvollziehbar, warum diese Formulierung nicht bereits in die konkrete Norm aufgenommen wird. Lediglich die Begründung zu § 7 ODGRP geht auf Einschränkungen und Lizenzen ein, wirkt in ihrer Formulierung jedoch praxisfern und nennt nicht die durch die EU empfohlene Lizenz CC-0 1.0, dagegen zwei Varianten der sogenannten „Datenlizenz

² siehe die Publikation „Daten als öffentliche Infrastruktur“ (2022), <https://www.boell.de/de/2022/09/15/daten-als-oeffentliche-infrastruktur>

Deutschland“ (ohne Versionsnummer), deren Rechtsgrundlage und Anwendbarkeit bereits im EU-Ausland fragwürdig ist.³

Der Begriff der „Maschinenlesbarkeit“ in § 3 Abs. 4 ODGRP fällt sowohl hinter die Begriffsbestimmung in § 5 Abs. 4 Nr. 1 LTranspG als auch die Definition in Artikel 2 Nr. 13 der ODD zurück. Nach wörtlicher Auslegung des Entwurfs erfüllt buchstäblich jedes digitale Speicherformat diese Anforderung – selbst ein Scan eines Schriftstücks, der danach in ein Office-Dokument eingefügt wurde. Die Begründung stellt die Anlehnung an § 3 Nr. 5 DNG klar, bezieht sich in der weiteren Ausführung aber wörtlich auf Erwägungsgrund 35 der ODD, der in der eigentlichen Norm des Entwurfs nicht ausformuliert wird. Artikel 2 Nr. 13 ODD stellt klar, dass mit „Maschinenlesbarkeit“ auch die maschinelle Interpretation von Sachverhalten und der internen Struktur von Dokumenten gemeint ist – also die automatisierte Interpretation der Semantik digital gespeicherter Informationen. Erwägungsgrund 35 führt zudem weiter aus: „Dokumente, die in einem Dateiformat kodiert sind, das eine automatische Verarbeitung einschränkt, weil die Daten nicht oder nicht ohne Weiteres aus ihnen extrahiert werden können, sollten nicht als maschinenlesbar gelten.“

Die Mehrdeutigkeit des Begriffs „Daten“ wird hier offenkundig nicht aufgelöst und kann zu verschiedener Auslegung einladen. Nach dem Wortlaut des Entwurfs kann der Begriff „Daten“ wahlfrei für einzelne maschinenlesbare Informationstapel oder lediglich für die digitale Codierung beliebiger Inhalte in einem Bitstrom angewandt werden. Das bedeutet auch, dass das LTranspG höhere Anforderungen an die maschinelle Interpretation der Semantik von Informationen in Besitz öffentlicher Stellen stellt als das ODGRP. Wenn der ausweisliche Zweck des ODGRP sein soll, gerade die Verfügbarkeit automatisiert maschinell auswertbarer Informationsbestände des Staats zu regeln, während sich das LTranspG lediglich auf Informationen beziehe, die für die Rezeption und Interpretation durch Menschen bestimmt sei, ist diese inkonsistente Begriffsverwendung von Maschinenlesbarkeit nicht nachvollziehbar.

Der Begriff der „Erhebung“ in § 3 Abs. 5 ODGRP wird als „aktives Beschaffen von Daten“ definiert. In der Begründung wird weiter präzisiert, dass es sich dabei um eine „erstmalige“ Beschaffung handle, nicht jedoch die Aufbereitung bereits vorhandener Daten. Diese Begriffsbestimmung wirft eine ganze Reihe von Fragen auf.

Fraglich ist beispielsweise, ob die Errichtung digitaler Infrastrukturen und Prozesse, in deren Folge einer öffentlichen Stelle ohne weiteres Zutun laufend Informationen zukommen, unter diese Bestimmung fällt. Schließlich war zwar die Einrichtung dieser Prozesse und Infrastrukturen ein aktiver Prozess, nicht jedoch der Zulauf von Informationen als Folge dieser Handlung – obgleich es sich hierbei ganz eindeutig um Informationen sowohl im Sinne von Artikel 2 Abs. 6 ODD und § 5 Abs. 1 LTranspG handelt,

³ vgl. hierzu die Publikation „Öffentliche Hand darf und sollte CC-Lizenzen nutzen“ und das darin wiedergegebene juristische Gutachten „Nutzung von Open Data-Lizenzen durch die öffentliche Hand“ vom 23.02.2024:

https://upload.wikimedia.org/wikipedia/commons/9/9e/2024-04-04_Nutzung_von_Open-Data-Lizenzen_durch_die_oeffentliche_Hand.pdf

zu denen es wenigstens auf Antrag einen Zugangs- und Lizenzierungsanspruch durch diese beiden Regulierungen geben könnte.

In Verbindung mit der Begriffsdefinition „unbearbeiteter“ Daten (hier den häufig irreführenden Begriff von „Rohdaten“ verwendend) liegt bei konsequenter Auslegung des Entwurfs der Schluss nahe, dass bei der Einrichtung neuer, digitaler Informationsbestände öffentlicher Stellen alleine der erstmalige Erhebungszustand öffentlich verfügbar gemacht wird, nicht jedoch laufende Aktualisierungen der Informationen. Weder der Wortlaut des Entwurfs noch die Begründung lassen einen gesetzgeberischen Willen erkennen, den Ist-Zustand von Informationen im Besitz öffentlicher Stellen unaufgefordert als Open Data zu veröffentlichen.

Die Absicht hinter der Begriffsdefinition „Unbearbeitet“ in § 3 Abs. 6 ODGRP ist nicht nachvollziehbar. Zwar wird weiter ausgeführt, dass der Begriff nicht für die Fälle gelte, in denen eine Veränderung (hier als „Bearbeitung“ bezeichnet) lediglich der Fehlerbereinigung oder „aus rechtlichen oder tatsächlichen Gründen“ erfolgt sei. Unklar ist jedoch auch nach Berücksichtigung der Begründung, ob hiermit *händische* Eingriffe des „Editierens und Formatierens“ gemeint sind, oder ob hiermit oder bei „Konvertierungen, die ihre Struktur oder Granularität verändern“ auch automatisierte Prozesse gemeint sein könnten. In der Logik einer Einbettung automatisierter Open-Data-Veröffentlichung in allgemeine Ziele der Verwaltungsdigitalisierung sollte die Maßgabe sein, dass maschinenlesbare Wissensbestände im Besitz öffentlicher Stellen so automatisiert wie möglich für die interne Wiederverwendung optimiert werden, aber auch die Veröffentlichung der veröffentlichungsfähigen Teile davon möglichst ohne händische Bearbeitung stattfindet.

Zudem werden in der Begründung Informationen aus verwaltungsinternen Prozessen, z.B. „Gesprächsvermerke, Stellungnahmen, Voten, Sprechzettel, Entwürfe für Gesetze oder Verordnungen, dienstinterne Anweisungen oder vorbereitende Unterlagen“ aus der Definition „unbearbeiteter“ Daten ausgenommen. Hier zeigt sich erneut die Schwäche des interpretationsoffenen Datenbegriffs in Konkurrenz zum LTranspG.

Auch der Begriff des „Bewertens“ oder „inhaltlichen Interpretierens“ trägt nur zur unnötigen Ausweitung des in der Begründung angeführten Ermessensspielraums der Informationen besitzenden öffentlichen Stelle bei. Wird ein Dokument bereits „bearbeitet“, indem es durch stille Betrachtung „bewertet“ oder „inhaltlich interpretiert“ und sodann ohne weiteren Eingriff in einem anderen Format gespeichert wird, was möglicherweise als „Veränderung der Struktur oder Granularität“ interpretiert werden könnte? Dies bleibt unklar, ebenso wie die Intention dieser Definition.

Keine strategische Zielsetzung für einen behördeninternen Nutzen erkennbar

In Verbindung der im vorangegangenen Abschnitt dargelegten Unschärfen bei den Begrifflichkeiten mit den **Definitionen zur Bereitstellung von Daten** in § 4 ODGRP ist ein zu erwartender Nutzen für die Informationen besitzenden Stellen unwahrscheinlich.

Einerseits bestehen weder Verpflichtung noch Rechtsanspruch auf die Bereitstellung von „maschinenlesbaren Daten“. Hiermit kann grundsätzlich in Frage gestellt werden, welchen Mehrwert dieses zusätzliche Gesetz gegenüber dem bestehenden LTranspG und den bestehenden Regelungen durch die ODD hätte. Durch die Bereitstellungspflicht lediglich „im Rahmen der verfügbaren Ressourcen“ und den fehlenden Rechtsanspruch wird ein enorm breiter Auslegungsspielraum geschaffen, der letztlich nur einem Wunsch an die öffentlichen Stellen in Rheinland-Pfalz entspricht. Warum es dafür eines Gesetzes bedarf, ist nicht nachvollziehbar.

Insbesondere die Regelungen zu „unbearbeiteten Daten“, der „Erhebung“ und der „Maschinenlesbarkeit“ verleiten geradezu dazu, ausschließlich die „Rohdaten“ zu veröffentlichen, die nicht etwa Ergebnis einer internen Verarbeitungskette oder gar eingerichteter Automatisierungsprozesse sind, sondern lediglich die erste Fassung *am Anfang* einer solchen Bearbeitung.

Durch die fehlenden Organisationspflichten und die im Vergleich zur ODD und dem LTranspG aufgeweichten Definitionen wird auch keinerlei Anreiz geschaffen, eine den Regeln der Verwaltung angemessene regel- und logikbasierte automatisierte Auswertbarkeit und Wiederverwendbarkeit durch die Wahl der Speicherung von Informationen zu ermöglichen. Sowohl die Einleitung als auch die Begründung, insbesondere der Abschnitt A VI. „Demografische Entwicklung“ nehmen zwar Bezug auf „KI“, scheinen damit aber lediglich probabilistische Systeme nach dem konnektionistischen Modell wie Large Language Models zu meinen, für die die Datenbestände der öffentlichen Hand als „Trainingsdaten“ (Begründung zu § 2) dienen sollen. Die Möglichkeiten einer besseren deterministischen Auswertbarkeit der Informationsbestände durch symbolische KI-Systeme, aber auch ganz klassische regelbasierte Verfahren scheinen hier nicht berücksichtigt worden zu sein.⁴ Ein Wirkmodell, wie der Fokus auf probabilistische und damit zwangsläufig fehlerbehaftete KI-Modelle zu dem erwarteten Effizienzgewinn und dadurch auch zu einer Reduzierung des Fachkräftebedarfs in der öffentlichen Verwaltung führen kann, wird im Entwurf nicht erläutert oder belegt.

Vielmehr halten wir es für geboten, Umfang und Qualität automatisierter Bereitstellung maschinell auch auf semantischer Ebene auswertbarer Informationen in Besitz öffentlicher Stellen als *Indikator* für eine gelungene Verwaltungsmodernisierung zu betrachten, deren primärer Nutznießer der Staat selbst ist. Anstatt unnötige Bereitstellungsprozesse für „Rohdaten“ und einen damit verbundenen Mehraufwand zu schaffen, müssen Prozessketten

⁴ Siehe hierzu auch unsere Veröffentlichung „generative KI für die Verwaltung?“ von 2025: <https://www.wikimedia.de/wp-content/uploads/2025/06/Generative-KI-in-der-Verwaltung-Einzelseiten.pdf>

für eine bessere Wiederverwendung bereits vorhandener Informationen durch den Staat selbst entstehen. Damit werden gleichzeitig die Basisinfrastrukturen geschaffen, die auch Grundlage für strategische Ziele wie die Registermodernisierung und die Umsetzung des OZG sind.

Beispiele für eine normtreue Umsetzung des Gesetzesentwurfs, die nicht im Sinne des Gesetzgebers liegen könnten

Um die sicher unbeabsichtigten möglichen Folgen des Entwurfs besser greifbar zu machen, möchten wir einige realistische und praktische Beispiele von Informationsflüssen aufzeigen, die sich strikt an die Regelungen des Gesetzesentwurfs halten.

Fall 1: Aggregation und Aufbereitung

Ein typischer Fall behördlicher Aggregation von Informationen ist, dass eine öffentliche Stelle von nachgeordneten Behörden Informationen übermittelt bekommt, die sie in einen gemeinsamen Bericht – z.B. auf Landesebene – zusammenfasst. Dieser Bericht könnte einerseits sowieso zur Veröffentlichung bestimmt sein, andererseits könnte es durch gesetzliche Regelungen zur Informationsfreiheit oder das LTranspG zumindest auf Antrag einen Anspruch auf Zugang geben.

Nach den vorgesehenen Definitionen und Regeln des ODGRP würde solch eine Zusammenfassung nicht vom ODGRP erfasst sein. Einerseits ist offen, ob es sich hierbei um „aktives Beschaffen“ handelt, wenn nachgeordnete Behörden aufgrund rechtlichen Auftrags diese Informationen an die übergeordnete Behörde liefern – die übergeordnete Behörde handelt hier schließlich nicht aktiv. Andererseits handelt es sich um Informationen, die eindeutig das Ergebnis einer Bearbeitung anderer Daten von Behörden des Landes sind. Landkreise sind beispielsweise nach § 1 LKO untere Behörde der allgemeinen Landesverwaltung und nehmen Aufgaben der öffentlichen Verwaltung im Sinne des § 2 LVwVfG wahr, auf den sich § 1 Abs. 1 ODGRP bezieht. Bei derart entstehenden Berichten auf Landesebene müssten folgerichtig alle unteren Behörden der allgemeinen Landesverwaltung die sogenannten „Rohdaten“ – sofern dies „im Rahmen der verfügbaren Ressourcen“ möglich ist – bereitstellen und dafür eigene Infrastrukturen bereitstellen und betreiben, auf denen diese „Rohdaten“ langfristig abrufbar bleiben (§ 6 Abs. 2 Satz 2 ODGRP).

An den eigentlich zu erstellenden Bericht, in dem diese Informationen aggregiert werden, werden sicherlich Gütekriterien hinsichtlich der Richtigkeit, dem Verständnis dienenden Interpretationen, Summen und weiteren Bearbeitungsschritten gestellt. Derartige Berichte werden häufig als menschenles- und interpretierbare Dokumente, beispielsweise im PDF-Format erstellt, wobei es in der Regel Zwischenschritte für die Aggregation und Bearbeitung der verwendeten Ausgangsinformationen wie beispielsweise tabellarische Auflistungen von Messgrößen gibt, die in dieses Dokument einfließen.

Bei folgerichtiger Auslegung des Gesetzentwurfs würde nun eine Situation entstehen, in der die nachgeordneten Behörden, die die Ausgangsinformationen erheben und beschaffen, diese „Rohdaten“ ohne weitere Plausibilitäts- oder Korrektheitsprüfung veröffentlichen sollen – aber nur, sofern ihre Ressourcen dies erlauben. In der Praxis könnte dies zu einem Flickenteppich führen, bei dem diese „Rohdaten“ *teilweise* und in unklarer Qualität verfügbar sind, aber nicht vollständig. Mögliche Zugangsansprüche auf diese Informationen bestehen parallel und unabhängig davon.

Die eigentlich relevante Information – nämlich die Zusammenführung der verteilten Informationen – wird möglicherweise auf Grundlage des LTranspG oder unabhängig davon freiwillig veröffentlicht. Möglicherweise besteht auch ein Zugangsanspruch auf Antrag zu dieser Information, zumindest im PDF-Format. Wir gehen davon aus, dass durch die Veröffentlichung dieses Dokuments oder eines gesetzlichen Anspruchs auf Zugang darauf auch ein Anspruch auf Erteilung einer Lizenz für die Weiterverwendung der Informationen im Sinne von Artikel 4 ODD i.V.m. § 2 Abs. 1 DNG besteht. Komplett absurd wäre es jedoch, wenn am Ende dieses Prozesses durch eine enge Auslegung des ODGRP ohne Berücksichtigung der ODD und des DNG die Erteilung einer Freien Lizenz für das PDF stünde, nicht jedoch für die aggregierten („bearbeiteten“, als „Ergebnis einer Bearbeitung anderer Daten durch eine Behörde des Landes“ entstandenen) *maschinell verarbeitbaren* Informationen selbst, die Grundlage für dieses PDF waren.

Fall 2: Erstmalige Erhebung vs. Fortschreibungen

Die Begriffsdefinitionen des Entwurfs lassen die Auslegung zu, dass bei der systematischen Erfassung neuer Informationsbestände in strukturierter Form lediglich die unbearbeitete *Ersterhebung* dieser Informationen veröffentlicht werden soll. § 6 Abs. 1 ODGRP bestärkt diesen Eindruck. Folgende Fortschreibungen dieser Informationen könnten jedoch sowohl bei wörtlicher als auch aus der Begründung folgender teleologischer Auslegung der Normen nicht vom ODGRP erfasst sein.

Verändert sich die Situation in der realen Welt, die Grundlage für die Beobachtung, Messung, Zählung, Berechnung oder Schätzung, die Übermittlung von Auskünften oder eine eigene Beobachtung ist, ist es sinnvoll und geboten, das in Daten abgebildete Modell der Realität auf diese anzupassen. Gemäß der Gesetzesbegründung ist eine „Erhebung“ jedoch ausdrücklich nur die *erstmalige aktive Beschaffung* solcher Informationen. Eine Aktualisierung bereits vorhandener Daten wird nicht erwähnt. Eine „Aufbereitung bereits vorhandener Daten“ wird explizit vom Begriff der „Erhebung“ ausgenommen. Es könnte auch der Eindruck entstehen, dass die Aktualisierung einmal erhobener Informationen etwas anderes sei als eine Fehlerbehebung. Denn wenn zum Zeitpunkt der erstmaligen Beschaffung die Information korrekt war, ist eine spätere Angleichung der Information auf den Ist-Zustand keine Korrektur eines Fehlers – denn die erstmalige Beschaffung war ja korrekt, es hat sich „lediglich“ die abzubildende Realität in der Zwischenzeit geändert.

Eine normtreue Umsetzung des ODGRP könnte bedeuten, dass jeweils bei der erstmaligen Ver-Datung von Sachverhalten im Land eben dieser Ausgangszustand unaufgefordert – sofern die Ressourcen dies zulassen – veröffentlicht werden soll. Fortschreibungen *können* veröffentlicht werden, sind jedoch freiwillige Zusatzleistungen über das ODGRP hinaus. Wie

in Fall 1 könnten jedoch durchaus wenigstens über die ODD i.V.m. § 2 Abs. 1 DNG Ansprüche auf Erteilung passender Lizenzen für die freie Wiederverwendung auch der Fortschreibungen der Informationen bestehen. Welchen praktischen Mehrwert und Nutzen das ODGRP in diesem Kontext bieten soll, bleibt unklar.

Fall 3: Veröffentlichung unbearbeiteter Rohdaten, jedoch nicht veröffentlichter Informationen in wiederverwendbarer Form

Analog zu Fall 1 ist auch die weitere Bearbeitung von Informationsbeständen selbst innerhalb einer Behörde (über Sachverhalte, die außerhalb der jeweiligen Behörde liegenden Verhältnisse betreffen) inklusive einer Bewertung, Interpretation, Bereinigung, Plausibilitätsprüfung, Filterung, Berechnung oder Zusammenfassung ein vollkommen normaler und alltäglicher Schritt bei der Aufbereitung von Informationen zur Veröffentlichung dieser Sachverhalte beispielsweise in einem schriftlichen Bericht.

Auch hier entsteht am Ende der Verarbeitungskette ein Dokument, das möglicherweise aufgrund gesetzlicher Bestimmungen oder auf sonstige Weise öffentlich oder zur ausschließlichen Nutzung bereitgestellt wird oder zu dem ein gesetzlicher Anspruch auf Zugang besteht.

Das heißt erneut, dass möglicherweise ein Anspruch auf Erteilung einer Lizenz für die Wiederverwendung wenigstens auf Antrag gemäß Artikel 4 ODD i.V.m. § 2 Abs. 1 DNG besteht – zumindest für das vorwiegend für menschliche Interpretation gedachte Dokument, möglicherweise aber auch für zugrunde liegende strukturierte Informationen beispielsweise in Tabellen oder Listen.

Das ODGRP würde sich hier allein auf die Ausgangsinformationen *vor* jeglicher Bearbeitung beziehen. Das hieße, dass am Ende einer Prozesskette Dokumente stünden, die sich nicht für automatisierte Weiterverarbeitung eignen, sondern in ihrer Aufbereitung auf menschliche Auswertung optimiert sind, und dass zusätzlich die „Rohdaten“ ganz am *Beginn* der Aufbereitungskette nach Möglichkeit als Open Data veröffentlicht werden sollen. Die eigentliche „Nutzinformation“, die in das Dokument am Ende der Prozesskette eingeflossen ist, ist jedoch überhaupt nicht zur unaufgeforderten Veröffentlichung vorgesehen, um es reibungsarm wiederverwenden zu können.

Das heißt in letzter Konsequenz auch, dass Behörden unnötige Zusatzaufwände zur unaufgeforderten Veröffentlichung von „Rohdaten“ mit unterschiedlichem praktischen Nutzen betreiben sollen, während die eigentlich sinnhaltigen und nützlichen Aufbereitungen dieser Informationen erst auf Antrag mit damit einhergehenden Verwaltungsaufwänden zugänglich werden.

Weitere Anmerkungen

Der Entwurf weist aus unserer Sicht weitere Mängel auf, die teilweise auf bislang ungehobenes Potenzial im LTranspG zurückzuführen sind.

Geltungsbereich

Der Entwurf wie auch das LTranspG beziehen sich auf „Behörden des Landes, soweit sie in öffentlich-rechtlicher oder privatrechtlicher Form Verwaltungstätigkeit ausüben“. Wir interpretieren dies so, dass hier – wie auch im LTranspG – beispielsweise berufsständische Kammern, staatlich kontrollierte Unternehmen, der öffentlich-rechtliche Rundfunk, Sparkassen und Landesbanken sowie staatliche Stiftungen und Anstalten des öffentlichen Rechts nicht oder nur teilweise betroffen sind.

Der Landtag, Gerichte, Strafverfolgungs- und -vollstreckungsbehörden sowie Hochschulen sind nur im Bereich öffentlicher Verwaltungsaufgaben betroffen. Selbst ohnehin stattfindende statistische Erhebungen dieser Stellen, die teilweise auch als Dokumente veröffentlicht werden, sind damit vom Geltungsbereich des Gesetzes ausgenommen, was wir nicht nachvollziehen können.

Ausnahmeregelungen

Der Entwurf des ODGRP soll zwar vom LTranspG aufgrund der angenommenen Besonderheit von „Daten“ abgegrenzt werden, bezieht sich jedoch in § 5 ausdrücklich auf die Ausnahmeregelungen der §§ 14 bis 16 des LTranspG.

§ 14 Abs. 1 Ziffer 11 erwähnt beispielsweise als Ausschlussgrund für Anträge auf Informationszugang, wenn es sich um „noch nicht aufbereitete Daten“ handle. Der Entwurf des ODGRP bestimmt dagegen in § 4 Abs. 3 Ziffer 4, dass er sich *nur* auf eben solche *nicht aufbereiteten* Daten beziehe. Möglicherweise könnte das in der Praxis bedeuten, dass die öffentliche Stelle die „Rohdaten“ vor der Aufbereitung sogleich unaufgefordert (per ODGRP) veröffentlichen soll, es jedoch keinen Anspruch auf Zugang auf Antrag (gemäß LTranspG) auf das Ergebnis gebe, bis die Informationen „aufbereitet“ sind. Da sich die datenhaltende Stelle jedoch ohnehin jederzeit aufgrund § 4 Abs. 1 ODGRP auf angeblich mangelnde Ressourcen berufen kann, bleibt der Beitrag des ODGRP unklar.

Die Einordnung der Ausschlussgründe nach § 15 LTranspG in die Systematik unbearbeiteter Fakteninformationen bleibt unklar. Gemäß der Begründung zu § 3 Abs. 6 ODGRP seien „Gesprächsvermerke, Stellungnahmen, Voten, Sprechzettel, Entwürfe für Gesetze oder Verordnungen, dienstinterne Anweisungen oder vorbereitende Unterlagen“ ohnehin keine „unbearbeiteten Daten“ im Sinne des Gesetzes. Bei den Erhebungsmethoden für „unbearbeitete Daten“ handelt es sich ausweislich der Gesetzesbegründung um Quantifizierungsmethoden anhand regelhafter Vorgehensweisen wie „beobachten, messen, zählen, berechnen, schätzen, Auskünfte oder eigene objektive Festlegungen.“ Der Bezug auf § 15 LTranspG legt die Interpretation nahe, dass die Bekanntgabe von Informationen über objektive, regelhaft vorgenommene Quantifizierungen der realen Welt den Erfolg von Entscheidungen oder die Vertraulichkeit von Beratungen haben könnten. Die Erläuterung realitätsnaher Beispiele, wie dies in der Praxis geschehen könnte, fänden wir sehr spannend. Allein die Formulierung, dass „Datensammlungen“ anhand „einer bewertenden oder interpretierenden Bearbeitung von Daten entstehen“ können, weicht von der Definition einer möglichst objektiven Abbildung realer Verhältnisse ab – dies scheint jedoch

vorwiegend einer zirkelschluss-artigen Mehrfachdefinition des „Daten“-Begriffs zu entspringen, die wir bereits als überarbeitenswert hervorgehoben hatten.

Auch der einbezogene Ausschlussgrund nach § 16 Abs. 1 Ziffer 1 LTranspG wirft Fragen auf, zumal es hier ausdrücklich um *Daten* als möglichst objektive Abbildung von Sachverhalten gehen soll, an denen aufgrund der offenkundig nicht vorhandenen individuellen Schöpfungshöhe regelmäßig kein Urheberrecht begründet werden kann.

Immaterialgüterrechte an Sammlungen solcher Informationen könnten allenfalls durch das europäische Leistungsschutzrecht für Datenbankhersteller begründet werden. Eine sinnvolle Regelung könnte hier sein, mögliche Einschränkungen der Weiterverwertung als Erweiterung von Artikel 1 Abs. 6 ODD zu präzisieren, beispielsweise für eine Nicht-Ausübung dieses Leistungsschutzrechts, eine vertragliche Sicherung, dass beauftragte Stellen dieses Recht nicht in Anspruch nehmen, und eine standardmäßige Veröffentlichung solcher Informationsbestände mit der Aufgabeerklärung und Lizenz CC-0 1.0, die den Verzicht auf die Ausübung dieser Rechte international nachvollziehbar klar stellt.

International gültige und anerkannte Lizenzen

Wie im Abschnitt zu den unklaren Begrifflichkeiten ausgeführt, erwähnt der Entwurf nur am Rande in der Begründung die passenden Lizenzen, die Grundlage für eine interoperable Wiederverwendbarkeit sind. Genannt werden ausschließlich die Lizenz „Creative Commons Namensnennung“ (hier „Creative-Commons BY“ genannt) ohne Versionsangabe, sowie zwei Fassungen der sogenannten „Datenlizenz Deutschland“, ebenfalls ohne Versionsangabe. Die insbesondere für Datensätze besonders geeignete Aufgabeerklärung und Lizenz „Creative Commons Zero 1.0“ wird nicht erwähnt.

Aufgrund der rechtlichen Mängel der sog. Datenlizenz Deutschland raten wir dringend davon ab, diese zu verwenden oder ihre Verwendung zu empfehlen. Wir raten stattdessen zur Verwendung der international anerkannten Lizenzen der Creative-Commons-Familie. Insbesondere für Faktendaten sollte die Lizenz und Freigabeerklärung Creative Commons Zero 1.0 empfohlen werden, für individuelle Schöpfungen die Lizenz Creative Commons Namensnennung 4.0. Lizenzen mit verpflichtender Namensnennung sollten ausschließlich dann empfohlen werden, wenn die *verpflichtende* Namensnennung unabdingbar geboten ist. Insbesondere für Datensätze sollte sie nur höchst sparsam verwendet werden und nur dann, wenn auch tatsächlich ein urheberrechtlicher Schutz besteht. In diesem Fall sollte eine ergänzende Klausel verwendet werden, die eine Namensnennung auch auf einer gemeinsamen Übersichtsseite oder in einer ähnlichen praxisgeeigneten Form gestattet.⁵

Mangelnde Aussagekraft einer Evaluation

In § 12 des Entwurfs wird eine Berichtspflicht gegenüber dem Landtag eingeführt. Hierfür sollen die Anzahl der „Datensätze“, die Anzahl der Datenabrufe, die Entwicklung all dieser

⁵ Siehe zu diesem Abschnitt und zu Formulierungsvorschlägen für Namensnennungsklauseln erneut das Gutachten „Nutzung von Open Data-Lizenzen durch die öffentliche Hand“ vom 23.02.2024: https://upload.wikimedia.org/wikipedia/commons/9/9e/2024-04-04_Nutzung_von_Open-Data-Lizenzen_durch_die_oeffentliche_Hand.pdf

Kenngrößen, Zugriffszahlen, sowie Hindernisse bei der Datenbereitstellung erhoben werden.

In der Begründung wird sogleich ausgeführt, dass die Anzahl tatsächlicher Datenabrufe durch die im Gesetz vorgesehene Struktur der Bereitstellung gar nicht erfasst werden kann, sondern lediglich die Zugriffe auf Metainformationen. Tatsächlich scheinen diese Messgrößen ungeeignet, Erfolg oder notwendige Anpassungen der Regelungen tatsächlich messen zu können.

Eine skalierungsfähige, produktive Wiederverwendung von Datensätze der öffentlichen Hand zeichnet sich durch eine hochgradige Automatisierung dieser Wiederverwendung aus. Idealerweise findet eine Erkundung des Informationsbestands zu Beginn einer solchen Automatisierung statt, in der Folge werden nur mehr die „Nutzdaten“ von der bereitstellenden Stelle abgerufen. Dies geschieht in der Erwartung, dass die veröffentlichende Stelle dabei etablierten Standards folgt und den aktuellsten Stand eines Datensatzes immer über denselben, langfristig stabilen URL anbietet, während ältere Fassungen versioniert unter ebenfalls kanonischen URL zugänglich bleiben. Wachsende Zugriffszahlen auf das Portal des Landes könnten daher zwei gegensätzliche Interpretationen zulassen: Sie könnten entweder eine steigende Nachfrage allgemein anzeigen, oder sie könnten ein Indikator dafür sein, dass sich die Daten bereitstellenden Stellen *nicht* an diese bewährte Praxis halten und für einen verlässlichen Abruf stets unnötigerweise der Zugang über das Portal des Landes notwendig ist.

Auch die Aussagekraft der „Anzahl der Datensätze“ ist fraglich. Wir haben u.a. beim bundesweiten Portals Govdata beobachtet, dass solch eine Messgröße Anreize schafft, einzelne Bestandteile eigentlich zusammenhängender Informationsbestände ohne fachlichen Grund als jeweils einzelnen „Datensatz“ zu veröffentlichen. Dies führt in der Folge zwar zu einer vermeintlich positiven Entwicklung bei der Zahl der Datensätzen, läuft jedoch dem Ziel einer besseren Wiederverwendbarkeit der Informationen zuwider.

Auch wie die Hindernisse bei der Bereitstellung systematisch erhoben und kategorisiert werden sollen, bleibt unklar – zumal hierbei nicht ausgeführt wird, ob es rein um den Prozess der Integration in die Plattform des Landes gehen soll, oder um vorgelagerte Prozessketten ab der Genese und das Zusammenspiel mit einer ganzheitlich gedachten Digitalisierungsstrategie.

Für weiteren Austausch zu einer möglichen Weiterentwicklung der Open-Data-Bestrebungen steht Ihnen das Team Politik und öffentlicher Sektor von Wikimedia Deutschland unter politik@wikimedia.de jederzeit gerne zur Verfügung.