



Deutsches Forschungsinstitut
für öffentliche Verwaltung

Abschlussbericht

2019/32

Friederike Bickmann / Zarina Feller / Marlene Kunst

Die Nutzung offener Verwaltungsdaten in Deutschland: Eine Anwendungs- und Bedarfsanalyse

Abschlussbericht

Friederike Bickmann / Zarina Feller / Marlene Kunst

Die Nutzung offener Verwaltungsdaten in Deutschland: Eine Anwendungs- und Bedarfsanalyse

Abschlussbericht

Deutsches Forschungsinstitut
für öffentliche Verwaltung
Gefördert durch die Bundesrepublik Deutschland

Inhalt

1	Hintergrund und Zweck der Studie	8
1.1	Offene Verwaltungsdaten	8
1.2	Forschungsstand und -lücken	9
1.3	Umsetzungsstand in Deutschland	10
1.4	Ziele und Inhalte der Studie	12
2	Methoden der Datenerhebung und -auswertung	13
2.1	Phase 1: Betrachtung der nationalen Metadatenplattform GovData.de	13
2.1.1	Metadaten	13
2.1.2	Logfiles und Matomo Analytics	14
2.1.3	Onlineumfrage	15
2.2	Phase 2: Analyse der Perspektiven verschiedener Nutzergruppen	15
2.2.1	Fokusgruppen	15
2.2.2	Auswertung von Statements aus der Perspektive von Nutzerinnen und Nutzern	19
3	Ergebnisse der Studie	21
3.1	Analyse des Datenangebots: Metadaten	21
3.1.1	Kategorien und Schlagwörter	21
3.1.2	Dateiformate und Lizenzen	23
3.1.3	Zulieferung und Bereitstellung von Datensätzen	25
3.1.4	Zusammenfassung	27
3.2	Analyse der Datennutzung: Matomo Analytics	27
3.2.1	Besuche	28
3.2.2	Durchgeführte Aktionen der Besucherinnen und Besucher	32
3.2.3	Übergänge der Besuche	36
3.2.4	Zusammenfassung	37
3.3	Analyse der Datennutzung: Logfiles	37
3.3.1	Zusammensetzung der Anfragen nach User-Client	38
3.3.2	Aufgerufene Suchergebnisseiten	39
3.3.3	Aufgerufene Detailseiten	47
3.3.4	Sonstige häufig aufgerufene Teilbereiche der Seite	53
3.3.5	Abfragen über die API	56
3.3.6	Betrachtung der Abfragen durch Bots und sonstige Systeme	61
3.3.7	Zusammenfassung	64
3.4	Analyse der Nutzerinnen und Nutzer: Onlineumfrage	65
3.4.1	Demographische Daten der Teilnehmerinnen und Teilnehmer	66
3.4.2	Nutzung von GovData.de	69
3.4.3	Bewertung von GovData.de	80
3.4.4	Zusammenfassung	87
3.5	Analyse der Perspektiven verschiedener Nutzergruppen: Ergebnisse der Fokusgruppen und der Auswertung von Nutzerstatements	88
3.5.1	Inhaltliche Aspekte	88
3.5.2	Nutzungsaspekte	94
3.5.3	Technische Aspekte	105
3.5.4	Rechtliche Aspekte	113
3.5.5	Zusammenfassung	119

4	Abschließende Betrachtungen	122
4.1	Limitationen der Studie	122
4.2	Zentrale Erkenntnisse	123
4.3	Ausblick und Optimierungsempfehlungen	125
	Literaturverzeichnis	128

Tabellenverzeichnis

Tabelle 1:	Die häufigsten gewählten Suchbegriffe bei Nutzung der Suchfunktion	41
Tabelle 2:	Die häufigsten gewählten Bestandteile von Suchbegriffen bei Nutzung der Suchfunktion	42
Tabelle 3:	Am häufigsten verwendete Schlagworte bei Nutzung der Suchfunktion	44
Tabelle 4:	Am häufigsten als Filter verwendete datenbereitstellende Plattformen	45
Tabelle 5:	Am häufigsten zur Filterung gewählte Lizenzen	47
Tabelle 6:	Bezeichnung der Datensätze zu denen am häufigsten Detailseiten aufgerufen wurden	49
Tabelle 7:	Abrufe von Detailseiten nach vergebener Kategorie	50
Tabelle 8:	Schlagwörter zu denen am häufigsten Detailseiten abgerufen wurden	51
Tabelle 9:	Abgerufene Detailseite nach datenbereitstellenden Plattformen	52
Tabelle 10:	Abgerufene Detailseiten nach Datensatz-Format	53
Tabelle 11:	Häufig abgerufene Informationsseiten	54
Tabelle 12:	Häufig abgerufene Blogeinträge	55
Tabelle 13:	Am häufigsten in den API-Abfragen verwendete Datensatzbezeichnungen	57
Tabelle 14:	Häufig in den abgefragten Datensatzbezeichnungen verwendete Begriffe	58
Tabelle 15:	Am häufigsten in den API-Abfragen verwendete Schlagwörter	59
Tabelle 16:	Am häufigsten in den API-Abfragen ausgewählte Dateiformate	61
Tabelle 17:	Bots bzw. Crawler mit den meisten erzeugten Abrufen	62
Tabelle 18:	Sonstige Systeme mit den meisten erzeugten Abrufen	63

Abbildungsverzeichnis

Abbildung 1:	Kategorisierung der Datensätze	22
Abbildung 2:	Die zehn häufigsten Schlagwörter	23
Abbildung 3:	Dateiformate	24
Abbildung 4:	Datenlizenzen	25
Abbildung 5:	Plattformen	26
Abbildung 6:	Datenlieferantinnen und -lieferanten	27
Abbildung 7:	Anteile neuer und wiederkehrender Besuche	29
Abbildung 8:	Monatliche Summe der Besuche	29
Abbildung 9:	Durchschnittliche Besuchszeit bei wiederkehrenden Besuchen	31
Abbildung 10:	Durchschnittliche Besuchszeit bei neuen Besuchen	31
Abbildung 11:	Absprungrate bei (neuen und wiederkehrenden) Besuchen	32
Abbildung 12:	Anteile der Aktionen von neuen und wiederkehrenden Besuchen	33
Abbildung 13:	Aktionen pro neuen und wiederkehrenden Besuchen	34
Abbildung 14:	Durchgeführte Aktionen im Gesamtzeitraum	35
Abbildung 15:	Einstiegsrouten der Besuche im Gesamtzeitraum	36
Abbildung 16:	Zusammensetzung der GET-Requests nach abfragendem System	39
Abbildung 17:	Ausschnitt aus einer Suchergebnisseite	40

Abbildung 18: Anzahl an Such-Requests pro Tag	40
Abbildung 19: Anzahl der Requests mit gewählter Filterung nach Kategorie	43
Abbildung 20: Am häufigsten als Filter verwendetes Datenformat in der Suchfunktion	46
Abbildung 21: Beispiel einer Detailseite auf GovData.de	48
Abbildung 22: Abgerufene Detailseiten pro Tag	48
Abbildung 23: Zusammensetzung der API-Abfragen nach abfragendem System	56
Abbildung 24: Anzahl der in den API-Abfragen gewählten Kategorie zur Auswahl der Daten	60
Abbildung 25: Zusammensetzung der Bot- bzw. Crawler-Abfragen	63
Abbildung 26: Zusammensetzung der Abfragen durch sonstige Systeme	64
Abbildung 27: Alter der Befragten	66
Abbildung 28: Geschlecht der Befragten	67
Abbildung 29: Wohnort der Befragten	68
Abbildung 30: Arbeitsort der Befragten	69
Abbildung 31: „Wie häufig haben Sie GovData in den letzten zwölf Monaten genutzt?“	70
Abbildung 32: „Wie sind Sie auf GovData aufmerksam geworden?“	71
Abbildung 33: „Nutzen Sie GovData.de beruflich?“	72
Abbildung 34: „In welchem Bereich sind Sie berufstätig?“	73
Abbildung 35: „Auf welcher Verwaltungsebene sind Sie tätig?“	74
Abbildung 36: „In welchem Wirtschaftszweig sind Sie tätig?“	75
Abbildung 37: „Aus welchen beruflichen Gründen nutzen Sie GovData?“ (Beschäftigte in der Wirtschaft)	76
Abbildung 38: „Aus welchen beruflichen Gründen nutzen Sie GovData?“ (Beschäftigte in der öffentlichen Verwaltung)	77
Abbildung 39: „Nutzen Sie GovData.de privat?“	78
Abbildung 40: „Aus welchen privaten Gründen nutzen Sie GovData.de?“	79
Abbildung 41: Zivilgesellschaftliche Organisationen, für die GovData.de genutzt wurde	80
Abbildung 42: „Haben Sie den gesuchten Datensatz/die gesuchten Datensätze auf GovData.de gefunden?“	81
Abbildung 43: „Entsprach der Datensatz Ihren Erwartungen?“	82
Abbildung 44: „Wie hoch schätzen Sie den Nutzen von GovData.de für Ihre Zwecke ein?“	83
Abbildung 45: „Bitte bewerten Sie die folgenden Aspekte der Open Data-Plattform GovData.de“	84
Abbildung 46: „Wünschen Sie sich die Bereitstellung weiterer Daten?“	85
Abbildung 47: „In welchen Bereichen wünschen Sie sich weitere Daten?“	86

Abkürzungsverzeichnis

APIApplication Programming Interface
BMIBundesministerium des Innern, für Bau und Heimat
CCODKompetenzzentrum Open Data
CSVCharacter-seperated values
D64D64 - Zentrum für Digitalen Fortschritt e. V.
dbbdbb beamtenbund und tarifunion
DCATData Catalogue Vocabulary
DNGGesetz für die Nutzung von Daten des öffentlichen Sektors
EGovGE-Government-Gesetz

EU	Europäische Union
EU27	Mitgliedstaaten der Europäischen Union
FÖV	Forschungsinstitut für Öffentliche Verwaltung
GDI-DE	<i>Geodateninfrastruktur Deutschland</i>
HTML	<i>Hypertext Markup Language</i>
INKAR	<i>Indikatoren und Karten zur Raum- und Stadtentwicklung</i>
INSPIRE	<i>Infrastructure for Spatial Information in the European Community</i>
JSON	<i>JavaScript Object Notation</i>
kfz	<i>Kraftfahrzeug</i>
NABU	<i>Naturschutzbund Deutschland</i>
OKF	<i>Open Knowledge Foundation Deutschland e. V.</i>
ÖPNV	<i>Öffentlicher Personennahverkehr</i>
PDF	<i>Portable Document Format</i>
PSI	<i>Public Sector Information</i>
RLP	<i>Rheinland-Pfalz</i>
RSS	<i>Really Simple Syndication Feed</i>
URL	<i>Uniform Resource Identifier</i>
WMS	<i>Web Map Service</i>
XML	<i>Extensible Markup Language</i>

1 Hintergrund und Zweck der Studie

1.1 Offene Verwaltungsdaten

Die Veröffentlichung von Verwaltungsdaten wird in verschiedenen Kontexten, wie im Rahmen des offenen Regierungs- und Verwaltungshandelns (Open Government) oder des E-Governments, diskutiert. Sie nimmt mittlerweile in vielen Regierungen und Verwaltungen weltweit, darunter Deutschland, sowie auch innerhalb der EU einen hohen Stellenwert ein.

Unter offenen Verwaltungsdaten können trotz der bereits vielfach vorgenommenen Eingrenzungen und Definitionen unterschiedliche Eigenschaften von Daten verstanden werden. In dieser Untersuchung werden als offene Verwaltungsdaten „nicht durch Datenschutz limitierte und nicht-vertrauliche Daten, die durch öffentliche Mittel produziert und verfügbar gemacht werden ohne jegliche Restriktionen hinsichtlich der Nutzung und Verbreitung“¹ verstanden.² Die Veröffentlichung von Verwaltungsdaten soll Transparenz, politische Teilhabe und wirtschaftliche sowie soziale Wertschöpfung für Bürgerinnen und Bürger, die Zivilgesellschaft, die Wirtschaft und für die Verwaltung selbst generieren.

Offene Verwaltungsdaten bereitzustellen, erfordert jedoch technische und finanzielle Ressourcen sowie Fachwissen und Expertise innerhalb der Behörden. Damit soll ein Datenangebot geschaffen werden, das für möglichst viele Nutzerinnen und Nutzer einen Mehrwert bietet. Ein kostenfreier Zugang, die Aktualität und Qualität der Daten (für bestimmte Nutzungszwecke), klare Richtlinien zur Verwendung der Daten sowie ein entsprechender technischer Standard der Daten (Maschinenlesbarkeit, Metadatenangebot etc.) sind dabei Voraussetzungen für ein niedrighwelliges Angebot, stellen aber auch Herausforderungen für die verantwortlichen bereitstellenden Behörden dar.³

Dem Aufwand und Ressourceneinsatz der Datenbereitstellung steht der durch die Bereitstellung erreichte Nutzen entgegen. Ein tatsächlicher Nutzen ist mit der Bereitstellung jedoch erst verbunden, wenn die Daten auch tatsächlich genutzt werden. Als zu erwartender Nutzen wird in der Literatur sowohl ein demokratischer Mehrwert der Transparenz und der politischen und sozialen Teilhabe beschrieben als auch ein hohes Potential für die wirtschaftliche und soziale Wertschöpfung.⁴ Dieser Argumentation folgend bieten offene Daten einer Vielzahl unterschiedlicher Akteure einen Nutzen: Nicht nur sollen Bürgerinnen und Bürger und die Zivilgesellschaft von der Transparenz der Verwaltungsdaten profitieren, auch Unternehmen sollen durch Innovationen oder Weiterverarbeitung der Verwaltungsdaten wirtschaftlichen Profit generieren können. Zudem kann die Verwaltung selbst durch einen effizienteren und effektiveren Zugang zu den Daten unterschiedlicher Behörden profitieren. Es ist jedoch davon auszugehen, dass nicht alle offenen Datenangebote den gleichen

1 Janssen et al. (2012), S. 258. (Eigene Übersetzung aus dem Englischen).

2 In der vorliegenden Untersuchung ist die Definition offener Daten abhängig von dem der Studie zugrunde liegenden Bottom-Up-Ansatz: So können Nutzerinnen und Nutzer, die an Untersuchungen der vorliegenden Studie teilhatten, ein unterschiedliches Verständnis von offenen Verwaltungsdaten haben. Aufgrund dessen versteht die vorliegende Studie offene Verwaltungsdaten in einem weitesten Sinn.

3 Janssen et al. (2012); Susha et al. (2015); Seibel (2016); Ksoll et al. (2017); Höchtl et al. (2011).

4 https://www.verwaltung-innovativ.de/SharedDocs/Publikationen/eGovernment/open_government_data_deutschland_langfassung.pdf?__blob=publicationFile&v=5 (Abrufdatum: 12. Juni 2021).

Nutzen erzielen, und welche Daten sich als besonders wertvoll erweisen werden, ist im Vorhinein bisher kaum feststellbar.

1.2 Forschungsstand und -lücken

Offene Verwaltungsdaten sind Gegenstand einer wachsenden Anzahl wissenschaftlicher Studien. Kosten und Herausforderungen bei der Bereitstellung und Weiterverwendung von offenen Verwaltungsdaten⁵, Analyse und Partizipationsmöglichkeiten von Nutzergruppen⁶, Modelle zum Vergleich und Bewertung von Datenportalen⁷ sowie die Analyse von Strukturen im Sinne eines Datenökosystems⁸ sind zentrale Untersuchungsgegenstände der theoretischen und empirischen Arbeiten.

Im Rahmen empirischer Studien wird immer wieder auf die Diskrepanz zwischen dem Angebot und den Bedarfen an offenen Verwaltungsdaten verwiesen⁹ – insbesondere das Unwissen darüber, welche Daten die Nutzerinnen und Nutzer für ihre Zwecke benötigen, schafft eine Situation, in der das Potential der Veröffentlichung von Verwaltungsdaten nicht in vollem Maße ausgeschöpft werden kann.¹⁰ Die Analyse der Nutzungsprozesse und Bedarfe sowie die kontinuierliche Kommunikation mit und das Feedback von Nutzerinnen und Nutzern sind elementar, um die Qualität der Daten für die Nutzerinnen und Nutzer zu sichern und zu optimieren.¹¹ Die Zusammenarbeit mit Nutzerinnen und Nutzern sowie die Abfrage ihrer Bedarfe erscheinen somit als eine zentrale Voraussetzung, um die Weiternutzung von Daten effektiv und effizient zu gestalten und einen möglichst hohen Wert aus der Veröffentlichung der Verwaltungsdaten zu ziehen.¹² Insbesondere in Deutschland gibt es bisher jedoch kaum empirische Studien, die Bedarfe, Nutzungsprozesse und Zwecke der Nutzung offener Verwaltungsdaten untersuchen. Offen bleiben demnach die folgenden Fragen:

- » Welche Nutzergruppen verwenden offene Verwaltungsdaten für welche Zwecke?
- » Welche Herausforderungen und Schwierigkeiten gibt es bei der Nutzung und Weiterverwendung des bestehenden Datenangebots?
- » Welche Anforderungen an die Qualität der Daten müssen gestellt werden?
- » Welche Bedarfe hinsichtlich des Datenangebots bestehen bei den verschiedenen Nutzerinnen und Nutzern?

5 Kucera & Chlapek (2014); Janssen et al. (2012); Conradie & Choenni (2014); Johnson et al. (2017); Zeleti & Ojo (2017).

6 Susha et al. (2015); Smith & Sandberg (2018); Welle Donker & van Loenen (2016); Zuiderwijk & Janssen (2015); Charalabidis et al. (2014); Lassinantti et al. (2019).

7 Harrison et al. (2012); Weber (2017); Welle & Donker (2016); Danneels et al. (2017); Sadiq & Indulska (2017).

8 Harrison et al. (2012); Abella et al. (2018); Zuiderwijk et al. (2014); Danneels et al. (2017).

9 Weber (2017), S.825; Smith & Sandberg (2018), S.255; Zuiderwijk et al. (2012), S.167.

10 Susha et al. (2015), S.194.

11 Vgl. Weber (2017), S.824f.; Smith & Sandberg (2018), S.225f.; Zuiderwijk & Janssen (2015), S.3; Martin et al. (2013), S.350f.

12 Vgl. Danneels et al. (2017), S.375.

Es bedarf empirischer Studien, die diese Zusammenhänge analysieren, um die Effekte und den Wert der Veröffentlichung von Verwaltungsdaten in Deutschland und Ansätze zur Optimierung der bestehenden Angebote aufzeigen.¹³

1.3 Umsetzungsstand in Deutschland

Bereits 2013 bekannte sich die Bundesrepublik Deutschland als G8-Staat mit der Open Data-Charta zu grundlegenden Prinzipien offener Verwaltungsdaten und verpflichtete sich durch die nationalen Aktionspläne im Rahmen der Teilnahme an der Open Government Partnership (2017-2019, 2019-2021, 2021-2023)¹⁴ zur Veröffentlichung von Datensätzen auf einer nationalen Plattform sowie zum Austausch mit Datennutzerinnen und -nutzern.

Die rechtlichen Grundlagen zur Bereitstellung von offenen Verwaltungsdaten sind in Deutschland insbesondere im E-Government-Gesetz (EGovG) definiert. Das Erste Gesetz zur Änderung des E-Government-Gesetzes vom 05. Juli 2017 hält zudem die Behörden der unmittelbaren Bundesverwaltung dazu an, erhobene Daten (ausgenommen weniger Daten) zu veröffentlichen (§ 12a EGovG) (Erstes Open-Data-Gesetz). Der Paragraph enthält darüber hinaus Kriterien zur Bereitstellung und Strukturierung der zu veröffentlichenden Daten. Am 24. Juni 2021 hat der Bundestag den Gesetzesentwurf „zur Änderung des E-Government-Gesetzes und zur Einführung des Gesetzes für die Nutzung von Daten des öffentlichen Sektors“ (Zweites Open-Data-Gesetz) beschlossen. In diesem Gesetz wird die Veröffentlichung von Daten nochmals ausgeweitet: Nunmehr sind die *Behörden des Bundes* dazu angehalten offene Daten (nun auch Daten, die zu Forschungszwecken erhoben wurden) bereitzustellen; zudem werden in zentralen Stellen der öffentlichen Verwaltung Open-Data-Koordinatorinnen bzw. Koordinatoren eingesetzt, die auf die Veröffentlichung von Daten in der jeweiligen Behörde hinwirken sollen. Ein Rechtsanspruch auf die Veröffentlichung von Daten wird jedoch weder im ersten noch im zweiten Open-Data-Gesetz begründet. Auch das Gesetz über die Nutzung von Daten des öffentlichen Sektors (Datennutzungsgesetz, DNG), das das Gesetz über die Weiterverwendung von Informationen öffentlicher Stellen (Informationsweiterverwendungsgesetz) ablöste und die EU-Richtlinie (EU) 2019/1024 (Public Sector Information-Richtlinie, PSI-Richtlinie) umsetzt, ist eine zentrale rechtliche Grundlage für die Struktur offener Daten in Deutschland. Es gilt nicht nur für die Behörden des Bundes, sondern auch für die Länder und Kommunen sowie einige öffentliche Unternehmen. Das Gesetz regelt und definiert die grundsätzlichen Kriterien sowie die Art und Weise der Veröffentlichung und Nutzung offener Daten. Neben den rechtlichen Grundlagen wird das Thema offene Verwaltungsdaten unter anderem durch die Open-Data-Strategie der Bundesregierung (2021) weiter vorangetrieben. Die Strategie listet eine Reihe von Handlungsfeldern und geplanter Maßnahmen auf, die den Komplex offener Verwaltungsdaten in Deutschland unterstützen sollen.

In Deutschland gibt es bereits ein Angebot offener Verwaltungsdaten auf allen föderalen Ebenen des politisch-administrativen Systems. Sowohl auf kommunaler wie auch auf Länder- und Bundesebene haben sich verschiedene Portale zur Bereitstellung offener Verwaltungsdaten etabliert. So stellen das

¹³ Vgl. Sadiq & Indulska (2017), S.153.

¹⁴ Bundesregierung (2017); Bundesregierung (2019a); Bundesregierung (2021a).

Transparenzportal Hamburg, Berlin Open Data oder das Portal Open Data Schleswig-Holstein beispielhaft Open-Data-Angebote auf Kommunal- bzw. Landesebene bereit. Auch auf Bundesebene bieten einige Bundesbehörden Verwaltungsdaten zur freien Weiterverwendung an. Die Bereitstellung erfolgt u. a. über die Portale Genesis-Online (Statistisches Bundesamt), mCloud (Bundesministerium für Digitales und Verkehr) und das Geoportal.de (Bundesamt für Kartographie und Geodäsie). Neben diesen separaten Angeboten verschiedener Bundesbehörden wurde mit GovData.de, einer Anwendung des IT-Planungsrates, eine einheitliche nationale Datenplattform geschaffen, die sowohl offene Verwaltungsdaten von Bundes- und Landesbehörden als auch von kommunalen Behörden zentral bereitstellt. GovData.de wurde durch die aktuelle Gesetzgebung und strategischen Überlegungen der Bundesregierung, insbesondere durch das DNG, in seiner Wichtigkeit als zentrale Plattform offener Daten gestärkt. Die Plattform GovData.de basiert auf einer Verwaltungsvereinbarung von Bund und dreizehn Bundesländern, welche den gemeinsamen technischen Betrieb sowie die Finanzierung der Plattform regelt. GovData.de schafft eine Plattform, auf der die Verwaltungsdaten zentral einsehbar und Metadaten zu den Datensätzen hinterlegt sind. Die Daten an sich verbleiben jedoch auf den Servern der bereitstellenden Behörden und sind durch einen externen Link von GovData.de aus erreichbar.

Die Bereitstellung offener Daten ist in Deutschland unter verschiedenen Lizenzen möglich. Neben den gängigen internationalen Lizenzen wurden auch deutsche Lizenzen zur Weiterverbreitung offener Daten entwickelt. Die Datenlizenz Deutschland, vorliegend in zwei Versionen, regelt die Lizenzbestimmungen von Datensätzen mit dem Zweck zur freien Weiterverwendung. Die Version „Datenlizenz Deutschland – Namensnennung – Version 2.0“ verpflichtet die Nutzerinnen und Nutzer bei der Weiterverwendung der Datensätze die Datenbereiterinnen und -bereiter zu nennen; die Version „Datenlizenz Deutschland – Zero – Version 2.0“ zieht keine Verpflichtungen bei der Weiterverwendung für die Nutzerinnen und Nutzer nach sich. Darüber hinaus wurde durch den IT-Planungsrat ein verbindlicher Standard zum Austausch von Metadaten geschaffen (DCAT-AP.de). Der deutsche Standard basiert auf dem Standard des Austausches von Metadaten auf europäischer Ebene und ist mit diesem kompatibel.

Das Europäische Datenportal dokumentiert in einem jährlichen Bericht die Fortschritte der EU-Mitgliedstaaten (EU27) sowie der Länder der Europäischen Freihandelszone und einiger weiterer Länder (Großbritannien, Ukraine, Georgien, Norwegen, Schweiz, Island und Montenegro) im Bereich Open Data¹⁵. Gemessen werden die Indikatoren Bereitschaft und Reife anhand von vier Dimensionen: Regulatorische Ausgestaltung zu offenen Daten, Erfassung der Nutzung von offenen Daten, Bewertung der Funktionen der zentralen Plattform für offene Daten sowie die Qualität der Daten. Im Open Data Maturity Report 2021 wird Deutschland im europäischen Vergleich als sogenannter *Fast-Tracker* kategorisiert (Fast Trackers erreichen 89-92 Prozent der möglichen Punktzahl der Messung) und befindet sich damit in dem zweihöchsten Cluster (von vier Clustern) und belegt damit insgesamt Platz 14 (von 34).¹⁶ Bei allen vier Dimensionen stellt sich Deutschland besser als der EU27-Durchschnitt dar; insbesondere bei der Policy- und Impact-Dimension schneidet Deutschland sehr gut ab – geringere Bewertungen erhielt Deutschland im Rahmen der Bewertung der zentralen Plattform (hier wurden im

¹⁵ Offene Verwaltungsdaten sind nur ein Teil von Open Data. Open Data können grundsätzlich auch Daten umfassen, die nicht mit öffentlichen Mitteln gesammelt wurden. Häufig werden diese Begriffe jedoch synonym füreinander verwendet. In den nachfolgenden Kapiteln werden die Begriffe synonym verwendet. Siehe dazu auch [Fußnote 2](#).

¹⁶ Vgl. European Data Portal (2021a).

Vergleich die Portaleigenschaften und die Bereitstellung der Daten schlechter bewertet) und der Qualität der Daten (hier wurden insbesondere die Entwicklung der Qualität sowie der Prozess der Aktualisierung von Daten im Vergleich schlechter bewertet).¹⁷

1.4 Ziele und Inhalte der Studie

Das Forschungsprojekt wurde in Zusammenarbeit mit dem Bundesministerium des Innern, für Bau und Heimat (BMI) durchgeführt. Ziel der vorliegenden Studie ist es, mit Hilfe mehrerer empirischer Erhebungen einen Beitrag zur Schließung der in Abschnitt 1.2 benannten Forschungslücken in Hinblick auf die Nutzung offener Daten in Deutschland zu leisten. Dementsprechend ist das zentrale Forschungsinteresse, zu einem besseren Verständnis der Nutzerinnen- und Nutzerperspektive auf offene Verwaltungsdaten in Deutschland beizutragen, sowie deren Nutzungsverhalten zu analysieren. Darüber hinaus sollen unterschiedliche Nutzergruppen identifiziert und deren spezifische Nutzungszusammenhänge dargestellt werden. Die so gewonnenen Erkenntnisse sollen auch eine Orientierungshilfe für eine bedarfsgerechtere Bereitstellung von offenen Verwaltungsdaten in Deutschland liefern.

Um diese Forschungsziele zu erreichen, wurden in zwei Projektphasen eine Vielzahl an unterschiedlichen Daten erhoben und mittels qualitativer und quantitativer Methoden ausgewertet. Die erste Projektphase konzentrierte sich auf die Nutzung der nationalen Metadatenplattform GovData.de. In der zweiten Projektphase wurde der Fokus erweitert auf Nutzungserfahrungen zu offenen Verwaltungsdaten in Deutschland, unabhängig von der genutzten Plattform. Hierzu wurde mit Nutzerinnen und Nutzern von offenen Datenangeboten in Fokusgruppen diskutiert sowie Statements zu diesem Thema ausgewertet.

Die nationale Metadatenplattform GovData.de wurde als Ausgangspunkt für die Betrachtungen zur Nutzung von offenen Datenangeboten gewählt, da es sich um eine Plattform handelt, auf der Informationen zu offenen Daten aus allen föderalen Ebenen der Bundesrepublik zu finden sind. Zudem ist die Plattform nicht thematisch gebunden, sondern enthält Informationen zu Datensätzen aus unterschiedlichen Bereichen. Aus diesen Gründen wurde die Plattform als das geeignetste Angebot ausgewählt, um einen möglichst breiten Eindruck von der Nutzung offener Daten in Deutschland zu erlangen. Hierzu wurde im ersten Schritt das Datenangebot von GovData.de mittels der dort hinterlegten Metadaten identifiziert. Zweitens wurden die Nutzung und die Nutzungsprozesse der auf GovData.de eingestellten offenen Verwaltungsdaten mittels Logfiles und eines Trackingtools (Matomo) analysiert. Ziel dieser Auswertung von Meta- und Nutzungsdaten ist es festzustellen, in welchem Umfang die nationale Metadatenplattform genutzt wird und welche Inhalte besonders häufig nachgefragt werden. Um zudem zu erfahren, mit welchen Motivationen die Seite genutzt wird und welchen Eindruck die Nutzerinnen und Nutzer von GovData.de haben, wurde zusätzlich eine Onlineumfrage geschaltet, in der die Teilnehmenden nach ihren Nutzungszwecken, ihrem Nutzungsverhalten sowie ihren Bedarfen und Wünschen an weiteren offenen Datenangeboten befragt wurden.

¹⁷ Vgl. European Data Portal (2021a); European Data Portal (2021b).

Ziel der zweiten Projektphase war es, den Blickwinkel der Untersuchung zu erweitern und die Vielschichtigkeit von Nutzungserfahrungen zu analysieren. Hierzu wurden Fokusgruppen als dialogorientiertes Verfahren gewählt, in denen Nutzerinnen und Nutzer aus drei zuvor identifizierten Nutzergruppen (öffentliche Verwaltung, Wirtschaft und Zivilgesellschaft) über ihre jeweiligen Erfahrungen und Positionen zu diesem Thema diskutierten. Auf diese Weise konnten zusätzliche Bedarfe und Herausforderungen bei der Nutzung von offenen Verwaltungsdaten identifiziert werden. Um so viele Positionen von Nutzerinnen und Nutzern wie möglich zu berücksichtigen, wurden zusätzlich veröffentlichte Dokumente, die Statements zu offenen Daten aus Nutzungssicht beinhalten, im Rahmen einer Onlinerecherche identifiziert. Die Ergebnisse der Fokusgruppen und die aufgefundenen Statements wurden mittels qualitativer Methoden ausgewertet.

Der vorliegende Abschlussbericht des Forschungsprojekts stellt die gesamten Ergebnisse der Untersuchungen zur Nutzung offener Verwaltungsdaten dar. Im Folgenden werden die Datenbasis sowie die Methoden der Datenerhebung und -auswertung kurz beschrieben ([Kapitel 2](#)). Anschließend werden die Ergebnisse der jeweiligen Analysen vorgestellt: [Kapitel 3.1](#) behandelt die Darstellung des Datenangebots mittels der Metadaten von GovData.de; [Kapitel 3.2](#) und [Kapitel 3.3](#) umfassen die Analyse der Nutzung des Datenportals GovData.de auf Basis Matomo Analytics-Daten und der Logfiles; [Kapitel 3.4](#) stellt die Ergebnisse der Onlineumfrage vor; in [Kapitel 3.5](#) werden die Ergebnisse der Fokusgruppen abgebildet sowie die Ergebnisse der Inhaltsanalyse der betrachteten Statements. Abschließend werden in [Kapitel 4](#) ein Fazit, basierend auf allen durchgeführten Untersuchungsschritten gezogen und mögliche Optimierungsansätze für eine nutzungsgerechtere Bereitstellung offener Daten in Deutschland formuliert.

2 Methoden der Datenerhebung und -auswertung

2.1 Phase 1: Betrachtung der nationalen Metadatenplattform GovData.de

Die Ergebnisse der vorliegenden Studie basieren auf einer breiten Datengrundlage, die im Folgenden detailliert vorgestellt wird. Zur Analyse der Nutzung von GovData.de wurden verschiedene Daten herangezogen, die unterschiedliche Einblicke sowohl in das Angebot offener Verwaltungsdaten als auch in die Nutzung auf GovData.de. Als technische Hilfsmittel wurden im Rahmen der Evaluation MAXQDA, die Programmiersprache R und die zugehörige Entwicklungsumgebung R-Studio, SoSci-Survey, ein Programm zur Erstellung und Durchführung von Onlineumfragen, sowie Microsoft Excel eingesetzt.

2.1.1 Metadaten

Zur Analyse des Angebots der offenen Daten auf GovData.de wurden die auf der Seite hinterlegten Metadaten ausgewertet. Bei GovData.de handelt es sich um eine Metadatenplattform. Der zentrale Mehrwert der Plattform ist somit der nach einem einheitlichen Standard angelegte Metadatenkatalog. Dieser ermöglicht einen vereinfachten Zugang zu Datensätzen verschiedener Bereitstellerinnen und Bereitsteller, da die Metadaten zu den verschiedenen Datensätzen standardisiert sind und somit leicht gesucht und gefiltert werden können. Die Metadaten beinhalten Informationen zu Inhalten, Datenlieferantinnen und -lieferanten, Eigentumsrechten, technischen und anderen Merkmalen der Daten und ermöglichen so detaillierte Aussagen über die Inhalte und Struktur der auf GovData.de

verlinkten Datensätze.¹⁸ Zur Auswertung wurden die Metadaten an einem Stichtag (01. Februar 2021) von GovData.de an das FÖV übermittelt. Die Metadaten wurden aufbereitet und mit deskriptiv-statistischen Methoden ausgewertet. Die Analyse der Metadaten erlaubt die Beantwortung der folgenden Forschungsfragen:

- » Wie viele Datensätze waren am Stichtag auf GovData.de verfügbar?
- » Welche Themenbereiche (zugeordnete Kategorie/vergebene Schlagwörter) decken die angebotenen Datensätze ab?
- » Welche Institutionen und Behörden stellen Verwaltungsdaten bereit?

Die Inhalte der Metadaten wurden grafisch aufbereitet und werden in [Kapitel 3.1](#) dargestellt und erläutert.

2.1.2 Logfiles und Matomo Analytics

Um die Nutzung der Plattform GovData.de zu untersuchen, wurden sowohl Matomo-Daten als auch Logfiles der Webseite GovData.de ausgewertet. Matomo ist die zur Analyse des Webtraffics der Seite eingesetzte Anwendung. Diese erstellt diverse Statistiken zur Nutzung der Website durch die Nutzerinnen und Nutzer. Die verschiedenen auf der Seite durchgeführten Aktionen werden dabei den jeweiligen Besucherinnen und Besuchern und deren Besuchen auf der Seite zugeordnet, sofern die Nutzerinnen und Nutzer dies nicht in ihren Browsereinstellungen ablehnen. Anhand der Matomo-Daten lassen sich u. a. die folgenden Fragen beantworten:

- » Wie viele Besuche gab es im beobachteten Zeitraum?
- » Wie viele Aktionen wurden (pro Besuch) durchgeführt? Wie viele Downloads wurden (pro Besuch) durchgeführt?
- » Wie viel Zeit verbrachten Besucherinnen und Besucher auf der Webseite?
- » Von welchen Webseiten griffen die Besucherinnen und Besucher auf GovData.de zu?

Matomo-Daten lagen dem FÖV für ein gesamtes Jahr (01. November 2020 bis 31. Oktober 2021) vor.

Darüber hinaus wurden die durch GovData.de zur Verfügung gestellten Logfiles genutzt, um detailliert die Nachfrage nach den unterschiedlichen Inhalten auf der Website nachzuvollziehen. Es handelte sich dabei um die durch den Web-Server der Seite aufgezeichneten Interaktionen. Ausgewertet wurden dabei die GET-Requests, d. h. die Abfragen von Inhalten auf GovData.de durch andere Systeme. Diese Logfiles wurden sowohl zu den Seitenabrufen als auch zu den Abrufen über die angebotene Schnittstelle (Application Programming Interface (API) durch GovData.de in anonymisierter Form an das FÖV übermittelt. Aus diesem Grund erfolgte keine Auswertung nach Besucherinnen und Besuchern bzw. Besuchen basierend auf den Logfiles. Stattdessen wurden die Logfiles genutzt um festzustellen, welche Inhalte auf der Website besonders häufig abgerufen wurden. Von besonderem Interesse sind dabei die durch Webbrowser erzeugten Abfragen von Suchergebnisseiten und Detailseiten zu den Datensätzen. Die Logfiles lagen dem FÖV für den Zeitraum 30. Juli 2020 bis 18. Januar 2021 vor.

¹⁸ Vgl. Riley (2017), S.6.

Anhand der Logfiles lasen sich u. a. die folgenden Fragen beantworten:

- » Welche Detailseiten werden abgerufen?
- » Welche Suchbegriffe werden verwendet? Welche Gruppen und Tags werden in der Suche verwendet?
- » Welche Datenformate werden gesucht?
- » Welche bereitstellenden Organisationen werden gesucht?

Logfiles und Matomo-Daten wurden entsprechend aufbereitet und deskriptiv statistisch ausgewertet. Die Ergebnisse werden in [Kapitel 3.2](#) und [Kapitel 3.3](#) vorgestellt.

2.1.3 Onlineumfrage

Logfiles und Matomo-Daten geben zwar Aufschluss über die durchgeführten Aktionen auf GovData.de, jedoch sagen sie wenig über die Nutzerinnen und Nutzer des Angebots offener Verwaltungsdaten aus. Um mehr über den Hintergrund, den Zweck der Nutzung und die Bewertung des Datenangebots durch Nutzerinnen und Nutzer zu erfahren, wurde zwischen dem 12. November 2020 und dem 31. März 2021 eine Onlinebefragung durchgeführt. Dazu wurde in Abstimmung mit dem BMI und der Geschäftsstelle von GovData.de ein standardisierter Fragebogen erstellt. Der Fragebogen wurde für den genannten Zeitraum sowohl auf der Startseite von GovData.de als auch als Blogbeitrag auf GovData.de verlinkt. Zusätzlich wurde von GovData.de ein Tweet mit dem Link zur Onlineumfrage abgesetzt.

Der Fragebogen deckte die folgenden Inhalte ab:

- » Häufigkeit der Nutzung,
- » Hintergrund Nutzerinnen und Nutzer,
- » Zweck der Nutzung,
- » Erfahrungen mit dem Suchen und Finden von Datensätzen,
- » Bewertung des Datenangebots und der Webseite.

Der Fragebogen zur Umfrage wurde von 94 Personen vollständig ausgefüllt.

2.2 Phase 2: Analyse der Perspektiven verschiedener Nutzergruppen

2.2.1 Fokusgruppen

Die Unterschiedlichkeit der Nutzerinnen und Nutzer sowie der vielfältige Einsatz und Gebrauch von offenen Verwaltungsdaten stellt für die Datenanbieter eine große Herausforderung bezüglich der Aufbereitung und dem konkreten Angebot von offenen Daten dar. Kontextspezifisches Wissen zu den Nutzerinnen und Nutzern sowie die Abfrage von Bedarfen, Herausforderungen und Wünschen hinsichtlich der technischen, inhaltlichen oder strukturellen Aufbereitung des Datenangebotes können

zu einer stärkeren Nutzung offener Daten beitragen und somit den Wert offener Verwaltungsdaten steigern.¹⁹

Aus diesem Grund wurden in der zweiten Projektphase stärker die Nutzung von offenen Verwaltungsdaten sowie die Bedarfe der Nutzerinnen und Nutzern in den Fokus der Betrachtung gerückt. Während in der ersten Erhebungsphase nichtreaktive Daten und die Ergebnisse einer teilstandardisierten Umfrage als Datengrundlage dienten, wurde in der zweiten Phase diese Datengrundlage mit Hilfe der Fokusgruppen als nichtstandardisierten, dialogorientierten Verfahren ergänzt, um die Vielschichtigkeit der Nutzungserfahrungen abbilden zu können, ohne von vorgefassten Thesen eingeschränkt zu werden.

Die empirische Methode der Fokusgruppe war für die zweite Phase des Forschungsprozess besonders geeignet, da die Möglichkeiten des offenen Austausches innerhalb einer Nutzergruppe prozessorientierte Diskussionen, Entwicklung von Ideen und den Austausch über Bedarfe der Teilnehmenden unterstützte.

Zudem hatte das Format der Fokusgruppe in diesem Fall auch einen partizipativen Anspruch. Den Nutzerinnen und Nutzern sollte so die Möglichkeit gegeben werden, ihre persönliche Perspektive zu schildern und sich mit anderen Nutzerinnen und Nutzern auszutauschen. Zu den Fokusgruppen wurde öffentlich eingeladen. Auch Nutzerinnen und Nutzer, die bisher nicht öffentlich in Erscheinung getreten waren bzw. nicht als solche durch das Forschungsteam identifiziert wurden, sollten die Chance erhalten, ihre Position einzubringen.

2.2.1.1 Bestimmung der Nutzergruppen

Die Auswahl der in den Fokusgruppen angesprochenen Nutzergruppen basierte auf einer Analyse bisheriger Studien zur Open-Data-Nutzung. In der aktuellen Literatur finden sich viele Modelle und Analysen zur Identifizierung und Einteilung von Nutzergruppen. Häufig werden Bürgerinnen und Bürgern, die Zivilgesellschaft und Unternehmen als zentrale Nutzergruppen genannt;²⁰ darüber hinaus finden sich noch eine Reihe weiterer Nutzungsfelder mit entsprechenden Nutzergruppen: Bildung und Forschung, Journalismus, die Verwaltung sowie die Politik selbst.²¹ Aus diesen Akteuren musste in der vorliegenden Evaluation für die Zusammenstellung der Fokusgruppen eine Auswahl im Hinblick auf die Nutzerlandschaft in Deutschland getroffen werden.

Da die Zivilgesellschaft und die Wirtschaft häufig als wichtige Nutzergruppen benannt werden (siehe Ausführungen im vorherigen Absatz) und auch die Bundesregierung im ersten Open-Data-Fortschrittsbericht die Intensivierung des Dialogs mit diesen beiden Nutzergruppen priorisiert,²² wurden diese für die Besetzung von zwei der drei vorgesehenen Fokusgruppen ausgewählt. Um die Übertragbarkeit der Ergebnisse in die Praxis bzw. den Wert der Untersuchung für die Open-Data-Strukturen in Deutschland zu steigern, war auch Perspektive der bereitstellenden Behörden

¹⁹ Vgl. Ruijter et al. (2017), S.470.

²⁰ U. a. Gásco-Hernández et al. 2018, S.234f.; Johnson et al. 2017, S.435; Zuiderwijk & Janssen 2014, S.257f.

²¹ U. a. Susha et al. 2015, S.184; Abella et al. 2019, S.2f.

²² Bundesregierung (2019b), S.35.

abzubilden. Der Austausch mit der Verwaltung sowohl als nutzender als auch als bereitstellender Akteur stellte für das Evaluationsteam darüber hinaus einen großen Mehrwert dar.

2.2.1.2 Einladung zur Teilnahme an den Fokusgruppen

Entsprechend der Zielsetzung der Fokusgruppe wurde zu den Veranstaltungen öffentlich eingeladen. Teilnahmebedingung war lediglich die Volljährigkeit der Teilnehmenden. Öffentlich eingeladen wurde über die Website des FÖV sowie über eine Terminankündigung auf der Seite GovData.de. Zusätzlich wurden Organisationen und Personen mittels einer Onlinerecherche ermittelt, bei denen anhand der veröffentlichten Selbstbeschreibungen eine eventuelle Nutzung von offenen Verwaltungsdaten angenommen wurde. Die so identifizierten Akteure wurden per Mail auf die Veranstaltung hingewiesen und gebeten, die Einladung an weitere potenzielle Nutzerinnen und Nutzer weiterzuleiten.

Für die Zielgruppe der Nutzerinnen und Nutzer aus der Wirtschaft wurden Unternehmensverbände ausgewählt und gebeten, die Einladung an ihre Mitgliedsunternehmen zu verteilen. Für die Nutzergruppe der Verwaltung wurden die kommunalen Spitzenverbände gebeten, die Einladung an die vertretenen Kommunen zu versenden. Zusätzlich wiesen auch das BMI und das Kompetenzzentrum Open Data (CCOD) über die eigenen Verteiler auf die Veranstaltung hin.

2.2.1.3 Ablauf der Fokusgruppen

Dem Ziel der Fokusgruppen entsprechend, möglichst unvoreingenommen die Erfahrungen und Bedürfnisse der Nutzerinnen und Nutzer einzuholen, wurden die durch das FÖV gesetzten Impulse minimal gehalten und nur eine sehr grobe Strukturierung der Diskussion in vier inhaltliche Blöcke vorgegeben.

Als anfänglicher Input durch das FÖV wurden der Ablauf des Projekts kurz vorgestellt und eine Definition von Open Data und offenen Verwaltungsdaten vorgestellt. Dies diente dazu, den Teilnehmenden zu vermitteln, welches Verständnis des zu diskutierenden Gegenstandes das Forschungsteam gewählt hat und wie sich die Erhebung in den Fokusgruppen in den größeren Projektzusammenhang einfügt.

Im Anschluss folgte eine Vorstellungsrunde, in der sich jeweils sowohl das Evaluationsteam als auch die Teilnehmenden vorstellten. Die Teilnehmenden wurden hierzu gebeten, neben Ihrem Namen und der Organisation, die sie vertreten, auch Auskunft darüber zu geben, in welchem Kontext sie bisher Open-Data-Angebote genutzt haben und über welche Aspekte der Nutzung sie gerne im Rahmen der Fokusgruppe diskutieren würden.

Die Antworten zu den zu diskutierenden Themen wurden noch während der Vorstellungsrunde von einer Mitarbeiterin des Projektteams dokumentiert und mit Hilfe der Online-Whiteboard- Lösung Conceptboard den vier vorgegebenen Blöcken zugeordnet. Das so entstandene Themen-Board wurde anschließend mit den Teilnehmenden geteilt und es wurde um Ergänzung gebeten. Diese Herangehensweise hat sich in den durchgeführten Fokusgruppen bewährt. Zum einen konnte durch das Board sichergestellt werden, dass alle Themenwünsche in der Diskussion berücksichtigt wurden, zum anderen wurde für die Teilnehmenden so leicht die Struktur der weiteren Diskussion ersichtlich. Zwischen den Blöcken und zum Abschluss wurde das Board wieder eingeblendet und durchgegangen, welche Punkte bereits diskutiert wurden und welche als nächstes folgen. Wurden in der Diskussion

Punkte angesprochen, die aus Sicht des Forschungsteams einem späteren Block zuzuordnen waren, so wurden diese auf dem Board entsprechend ergänzt.

2.2.1.4 Zusammensetzung der Teilnehmerinnen und Teilnehmer

In den veranstalteten Fokusgruppen wurden die folgenden Teilnehmendenzahlen erreicht:

- » Fokusgruppe für Nutzerinnen und Nutzer aus der öffentlichen Verwaltung (Fokusgruppe öffentliche Verwaltung): Acht Teilnehmende
- » Fokusgruppe für Nutzerinnen und Nutzer aus der privaten Wirtschaft (Fokusgruppe Wirtschaft): Sechs Teilnehmende
- » Fokusgruppe für Nutzerinnen und Nutzer aus der Zivilgesellschaft (Fokusgruppe Zivilgesellschaft): Fünf Teilnehmende

Aufgrund der Durchführung der Fokusgruppen als Onlineformat war die Teilnahme für Personen aus ganz Deutschland ohne entstandene Reisekosten möglich. Dementsprechend beteiligten sich Nutzerinnen und Nutzer aus verschiedenen Regionen. Besonders viele Teilnehmende vertraten Organisationen aus Nordrhein-Westfalen (sieben Personen) und aus Berlin (fünf Personen). Weiterhin waren Organisationseinheiten mit Sitz in Bayern, Niedersachsen, Hamburg und Schleswig-Holstein vertreten.

Insbesondere in der Fokusgruppe öffentliche Verwaltung hat sich gezeigt, dass viele der angesprochenen Nutzerinnen und Nutzer auch gleichzeitig mit der Bereitstellung von Daten befasst sind. Obwohl sich die Fokusgruppen explizit an Nutzerinnen und Nutzer der Angebote richtete, wurden in den Fokusgruppen auch Eindrücke aus Perspektive der Bereitstellerinnen und Bereitsteller eingebracht. Aus Sicht des Evaluationsteams ist dies jedoch nicht als problematisch zu bewerten, sondern wird lediglich als Hinweis auf die enge Verknüpfung der Perspektive von Bereitstellerinnen und Bereitsteller sowie von Nutzerinnen und Nutzern in dieser Nutzergruppe gewertet.

2.2.1.5 Beschreibung Auswertungsmethode

Während der Fokusgruppen wurde aus Datenschutzgründen auf eine Video- oder Tonaufnahme verzichtet. Die Ergebnissicherung erfolgte, neben der Sammlung von Themen auf dem geteilten Board, durch die parallele Anfertigung zweier Verlaufsprotokolle der jeweiligen Veranstaltungen.

Im Anschluss wurden diese zwei Protokolle durch eine Projektmitarbeiterin zu einem finalisierten Protokoll pro Fokusgruppe zusammengeführt. Der Abgleich ergab dabei bei allen drei Veranstaltungen nur minimale Abweichungen zwischen den beiden Versionen der Protokolle, welche sich leicht ausräumen ließen.

Die Auswertung der Protokolle erfolgte orientiert an den in den Gruppen vorgegebenen vier inhaltlichen Blöcken, auch die Darstellung der Ergebnisse wird diese Gliederung ([Kapitel 3.5](#)) aufgreifen. Das Codierungsschema für die weiteren Unterkategorien wurde induktiv aus den protokollierten Aussagen entwickelt. Anmerkungen, die sich explizit auf GovData.de bezogen, wurden gesondert erfasst. Ebenso wurde eine eigene Kategorie für die getätigten Äußerungen aus Perspektive der Bereitstellerinnen und Bereitsteller aufgenommen.

2.2.1.6 Limitationen

Aufgrund des diskursiven Formates der Fokusgruppen konnte im gewählten methodischen Design nur eine sehr geringe Zahl an Teilnehmerinnen und Teilnehmer berücksichtigt werden. Daher liefert diese Methode Schlaglichter auf die vielfältigen Erfahrungen der Nutzerinnen und Nutzer und ist nicht als repräsentativ für die Gesamtheit aller Nutzerinnen und Nutzer von Open-Data-Angeboten in Deutschland anzusehen.

Zudem hat das sehr offene Einladungsformat zu gewissen Überschneidungen in den Eigenschaften der Teilnehmerinnen und Teilnehmern geführt. So waren von drei Organisationen jeweils mehrere Personen in den Fokusgruppen anwesend, wodurch die Perspektiven dieser Organisationen in gewisser Weise überrepräsentiert wurden. Bei den Wortmeldungen kam es jedoch nur zu sehr geringen Doppelungen in den Äußerungen der Personen von den gleichen Organisationen, zumal die Personen jeweils unterschiedliche Arbeitsbereiche in den jeweiligen Organisationen vertraten.

Ebenfalls wurde vor Durchführung der Fokusgruppen keine Plausibilitätsprüfung der selbstgewählten Zuordnung der Teilnehmenden zu den verschiedenen Nutzergruppen vorgenommen. In Einzelfällen wurde von Teilnehmenden die inhaltliche Zuordnung der Fokusgruppen ignoriert, was die Diskussionen aus Sicht des Evaluationsteams jedoch nicht nachteilig beeinflusste.

2.2.2 Auswertung von Statements aus der Perspektive von Nutzerinnen und Nutzern

Die durchgeführten Fokusgruppen ermöglichten die direkte Interaktion mit den Nutzerinnen und Nutzern offener Verwaltungsdaten. Die Methode hat jedoch den Nachteil, dass sie nur die Positionen der anwesenden Personen erfasst und somit die Ergebnisse stark abhängig sind von der Zusammensetzung des Teilnehmendenkreises. Um diesem Problem entgegenzuwirken, wurde zusätzlich eine Recherche zu veröffentlichten Statements von Nutzerinnen und Nutzern offener Datenangebote in Deutschland durchgeführt. Dieser Schritt diente dazu, auch Meinungen von Nutzerinnen und Nutzer offener Daten zu berücksichtigen, die bei den Teilnehmenden der Fokusgruppen nicht vertreten waren.

Hierzu wurde eine Online-Recherche zu Äußerungen bezogen auf Open-Data-Angebote durchgeführt, die im deutschsprachigen Raum veröffentlicht wurden. Statements die dabei explizite offene Daten aus der Schweiz oder Österreich beschreiben, wurden ausgeschlossen, da sich die vorliegende Analyse auf die Bundesrepublik Deutschland bezieht. Als zeitliche Eingrenzung wurde der Zeitraum der Veröffentlichung von 2018-2021 gewählt. Damit fallen unter die Statements in diesem Zeitraum auch Stellungnahmen zu der Änderung der PSI-Richtlinie 2018 und dem Zweiten Open-Data-Gesetz. Nicht berücksichtigt wurden hingegen Social-Media-Beiträge und Kommentare unter medialen Beiträgen z. B. Blogposts, Zeitungsartikel etc.²³ Blogposts, Zeitungsartikel und Gutachten wurden nur insofern berücksichtigt, inwieweit sie erkennbar auf den Einschätzungen von Nutzerinnen und Nutzern offener Datenangebote beruhten. Viele der abgegebenen Statements wurden von Interessenverbänden und zivilgesellschaftlichen Organisationen formuliert, es fanden sich aber auch Äußerungen von Einzelpersonen. Ein Anspruch auf Vollständigkeit wird nicht erhoben. Dennoch ermöglicht die

²³ GitHub wurde dabei nicht als Social-Media-Plattform gewertet. Auf GitHub getätigte Äußerungen zu Open Data aus Perspektive der Nutzerinnen und Nutzer wurden also als Statement berücksichtigt.

zusätzliche Berücksichtigung eine Erweiterung der in den Fokusgruppen gewonnenen Eindrücke zu den Bedarfen und Erfahrungen der Nutzerinnen und Nutzer.

Die aufgefundenen Statements wurden anhand der vier für die Fokusgruppen gesetzten thematischen Blöcke codiert. Die vergebenen Sub-Codes wichen leicht von denen der Fokusgruppen ab, da sich in den ausgewerteten Statements auch einige thematische Aspekte fanden, die in den Fokusgruppen nicht angesprochen wurden und vice versa. Die jeweiligen Ergebnisse der Auswertung der in den Statements festgehaltenen Aussagen werden nach den vier inhaltlichen Blöcken unterteilt im Anschluss an die jeweiligen Ausführungen zu den Ergebnissen der Fokusgruppen dargestellt.

3 Ergebnisse der Studie

3.1 Analyse des Datenangebots: Metadaten

Die auf GovData.de hinterlegten Metadaten zu den jeweiligen Datensätzen folgen dem Standard DCAT-AP.de. Der Standard DCAT-AP.de wurde für deutsche Open Data-Portale entwickelt und orientiert sich am europäischen Standard DCAT-AP.de. Am 28. Juni 2018 wurde durch den IT-Planungsrat die verbindliche Nutzung von DCAT-AP.de für den Austausch von offenen Verwaltungsdaten in Deutschland entschieden und GovData.de mit dem Betrieb und der Pflege des Standards beauftragt.²⁴

Der DCAT-AP.de-Standard gibt vor, welche Metadaten für Datensätze hinterlegt und an GovData.de übermittelt werden müssen. Ziel des vorgegebenen Standards ist es, dass alle für die Datenrecherche erforderlichen Informationen verfügbar gemacht werden. Dazu gehört bspw. der Name der Organisation bzw. Plattform, die den Datensatz bereitstellt. Des Weiteren werden die Bereitstellerinnen und Bereitsteller u. a. dazu aufgefordert, den Datensätzen Kategorien und Schlagwörter zuzuordnen.²⁵

Durch die Analyse der Metadaten soll ein Eindruck davon vermittelt werden, wie viele Datensätze im Zeitverlauf mehrerer Jahre auf GovData.de verlinkt wurden, welche Plattformen besonders viele Datensätze eingestellt haben und wie diese inhaltlich ausgerichtet sind. Die Metadaten wurden am 01. Februar 2021 abgerufen und dem FÖV von GovData.de zur Verfügung gestellt.

Am Stichtag, dem 01. Februar 2021, waren auf GovData.de zu insgesamt 40.749 Datensätzen Metadaten hinterlegt.²⁶

3.1.1 Kategorien und Schlagwörter

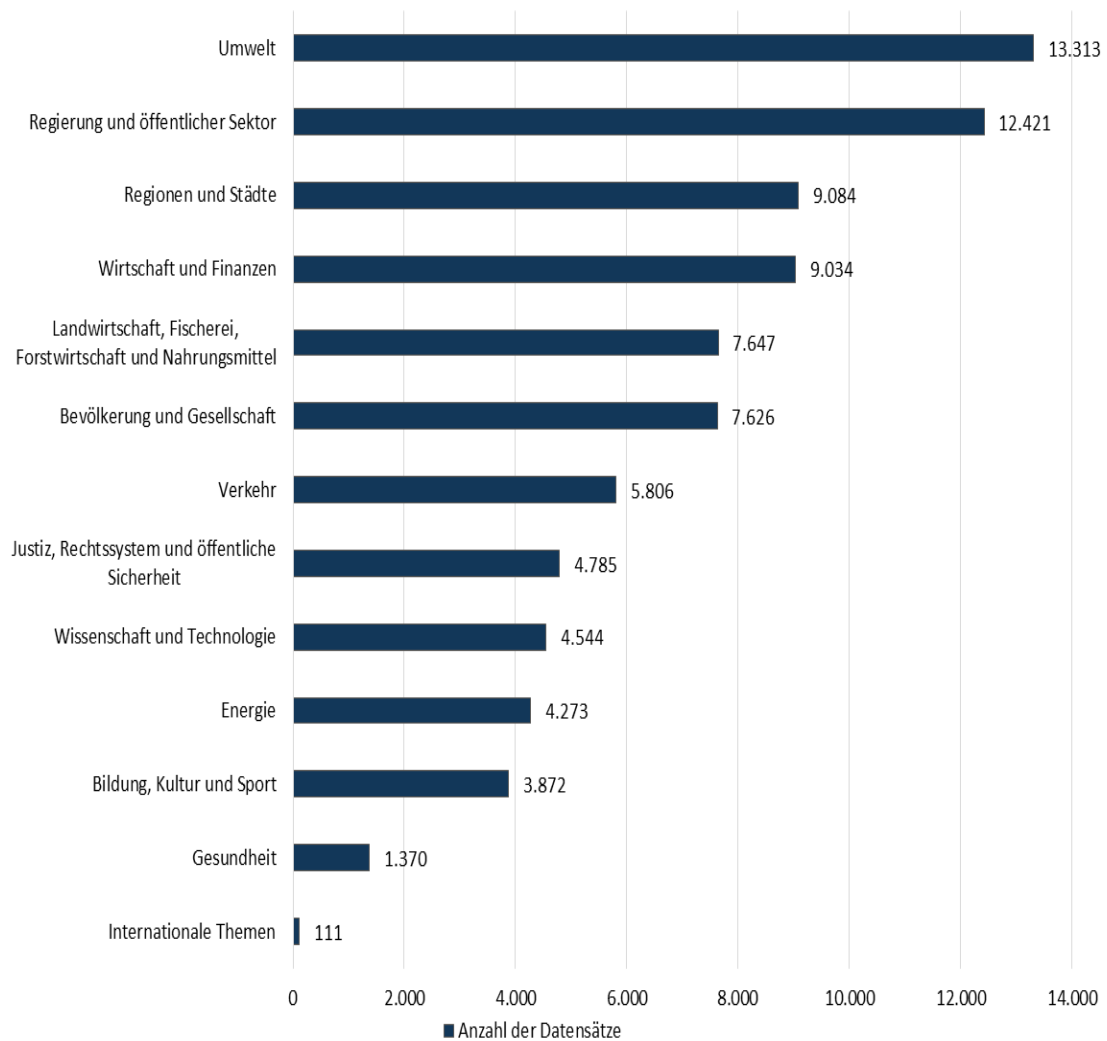
Um mehr über die inhaltliche Ausrichtung der Datensätze zu erfahren, wurden die Kategorien und Schlagwörter untersucht, die den Datensätzen zugeteilt wurden.²⁷ Abbildung 1 zeigt, dass es sich bei „Umwelt“ sowie „Regierung und öffentlicher Sektor“ um die Kategorien handelt, zu denen am Stichtag die meisten Datensätze auf GovData.de verlinkt waren. Auffällig wenige Datensätze finden sich zu den Kategorien „Gesundheit“ und „Internationale Themen“.

24 <https://www.dcat-ap.de/def/dcatde/1.1/spec/specification.pdf> (Abrufdatum 17. August 2021)

25 <https://www.govdata.de/metadaten/schema>; <https://www.dcat-ap.de/def/dcatde/1.1/spec/specification.pdf> (Abrufdatum 17. August 2021)

26 In [Kapitel 3.3.3.1](#) wird die Anzahl der Metadaten dieser noch einmal detaillierter dargestellt.

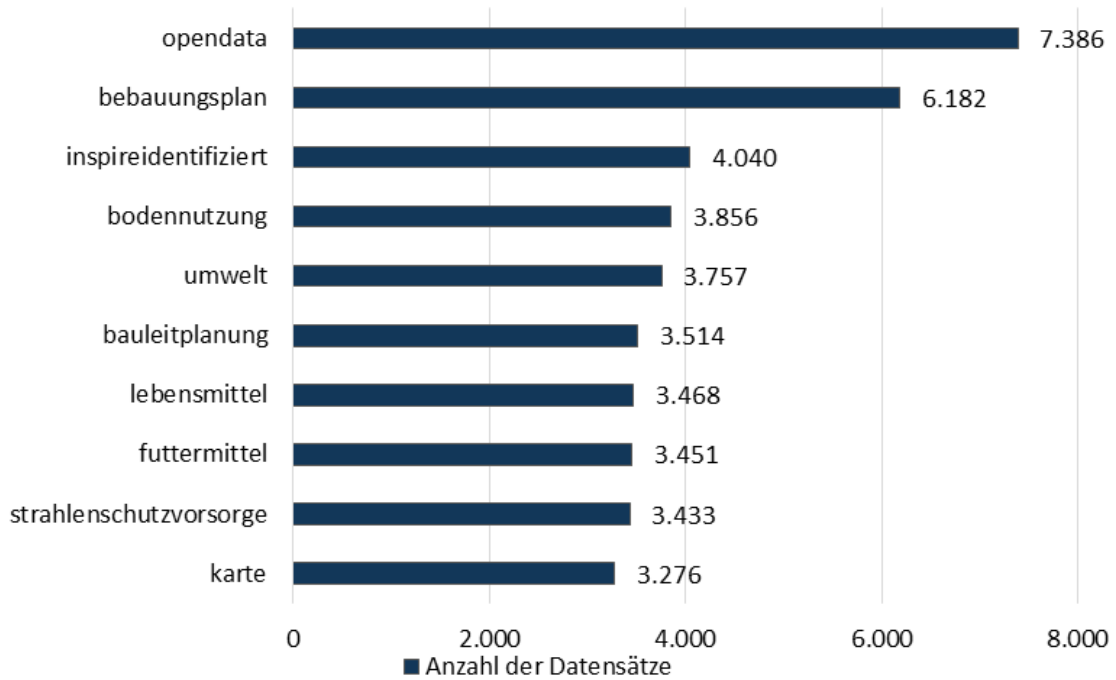
» **Abbildung 1: Kategorisierung der Datensätze**



Die Plattformen, die Datensätze bereitstellen, werden nicht nur dazu aufgefordert, die Datensätze zu kategorisieren, sondern auch ihnen Schlagwörter zuzuteilen. Diese Schlagwörter können von den Plattformbetreibern frei gewählt werden. Insgesamt wurden 33.715 unterschiedliche Schlagwörter vergeben.²⁸ Abbildung 2 zeigt die zehn häufigsten Schlagwörter. Besonders häufige Schlagwörter sind demnach neben „opendata“ bspw. „bebauungsplan“ und „inspireidentifiziert“ (in Bezug auf die Richtlinie 2007/2/EG Infrastructure for Spatial Information in the European Community, INSPIRE-Richtlinie, für die gesamteuropäische Geodateninfrastruktur).

28 Mehrere Schlagwörter pro Datensatz sind möglich.

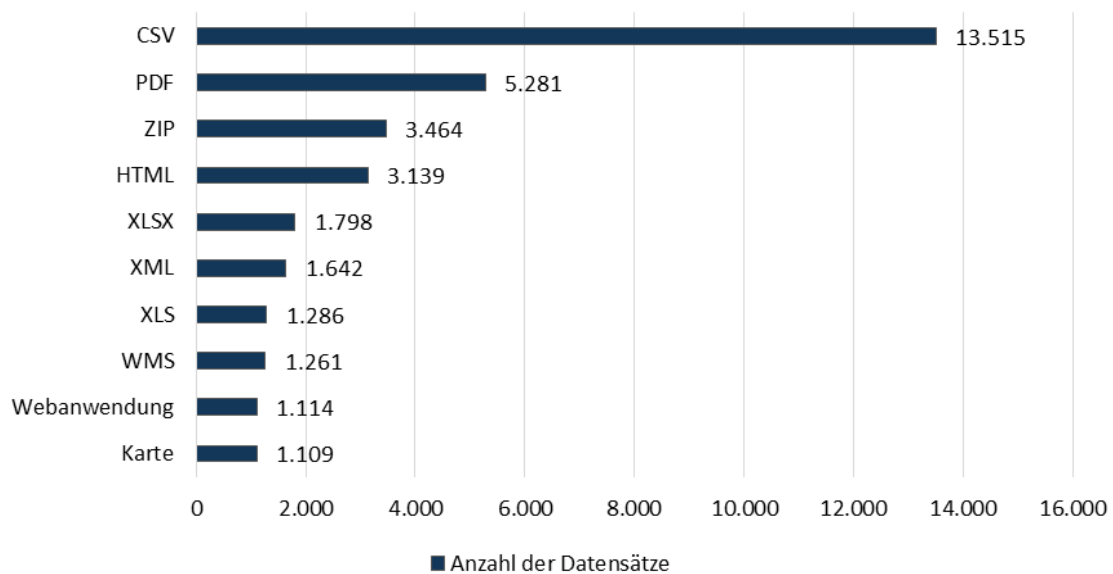
» **Abbildung 2: Die zehn häufigsten Schlagwörter**



3.1.2 Dateiformate und Lizenzen

Der Abbildung 3 sind die zehn häufigsten Dateiformate der verlinkten Datensätze zu entnehmen. Dabei zeigt sich, dass Datensätze mit Abstand am häufigsten im CSV-Format bereitgestellt wurden, gefolgt von Datensätzen im PDF-Format. Welche Formate in ZIP-Dateien zur Verfügung gestellt wurden, lässt sich aus den Metadaten nicht erschließen. Zudem sei darauf hingewiesen, dass manche Bereitstellerinnen und Bereitsteller für ihre Datensätze mehrere Formate angaben, eine stichprobenartige Prüfung jedoch ergab, dass die jeweiligen Datensätze nicht in all diesen Formaten bereitstehen. Aus diesem Grund sind in diese Analyse nur Datensätze eingegangen, denen jeweils ein Format zugeordnet wurde.

» **Abbildung 3: Dateiformate**



Des Weiteren wurde untersucht, unter welchen Lizenzen die Datensätze laut der Metadaten zur Verfügung stehen. Wie Abbildung 4 zu entnehmen ist, ist die mit Abstand häufigste Lizenz die „Datenlizenz Deutschland Namensnennung 2.0“. Es handelt sich hierbei um „eine von Bund, Ländern und kommunalen Spitzenverbänden entwickelte Empfehlung für eine einheitliche Nutzungsbestimmung für Verwaltungsdaten in Deutschland“²⁹. Die Lizenz erlaubt die kommerzielle und nicht-kommerzielle Nutzung der Daten; als Quellenvermerk muss jedoch der Name der Urheberin oder des Urhebers, der Vermerk „Datenlizenz Deutschland – Namensnennung – Version 2.0“ oder „dl-de/by-2-0“ mit Verweis auf den Lizenztext, sowie einen Verweis auf die Uniform Resource Identifier (URL) des Datensatzes angegeben werden.³⁰ Datensätze mit der Lizenz „Datenlizenz Deutschland – Zero – Version 2.0“ dürfen dagegen uneingeschränkt genutzt werden.³¹

Die Lizenz „Creative Commons Namensnennung“ erlaubt eine Weiterverbreitung und Weiterverwendung von Daten unter der Bedingung, dass angemessen auf die Urheberin oder den Urheber hingewiesen wird.³² Datensätze mit der Lizenz „Creative Commons Namensnennung – 4.0 International“ dürfen ebenso weiterverbreitet und weiterverwendet werden, solange die Urheberin oder der Urheber angemessen genannt ist, ein Link zur Lizenz bereitgestellt wird und Änderungen am Datensatz vermerkt werden.³³ Weitere Informationen zu den jeweiligen Lizenzen sind auf GovData.de zu finden.³⁴

29 <https://transparenz.hamburg.de/open-data/13460318/datenlizenz-deutschland/> (Abrufdatum: 19. Juli 2021).

30 <https://www.govdata.de/dl-de/by-2-0> (Abrufdatum: 19. Juli 2021).

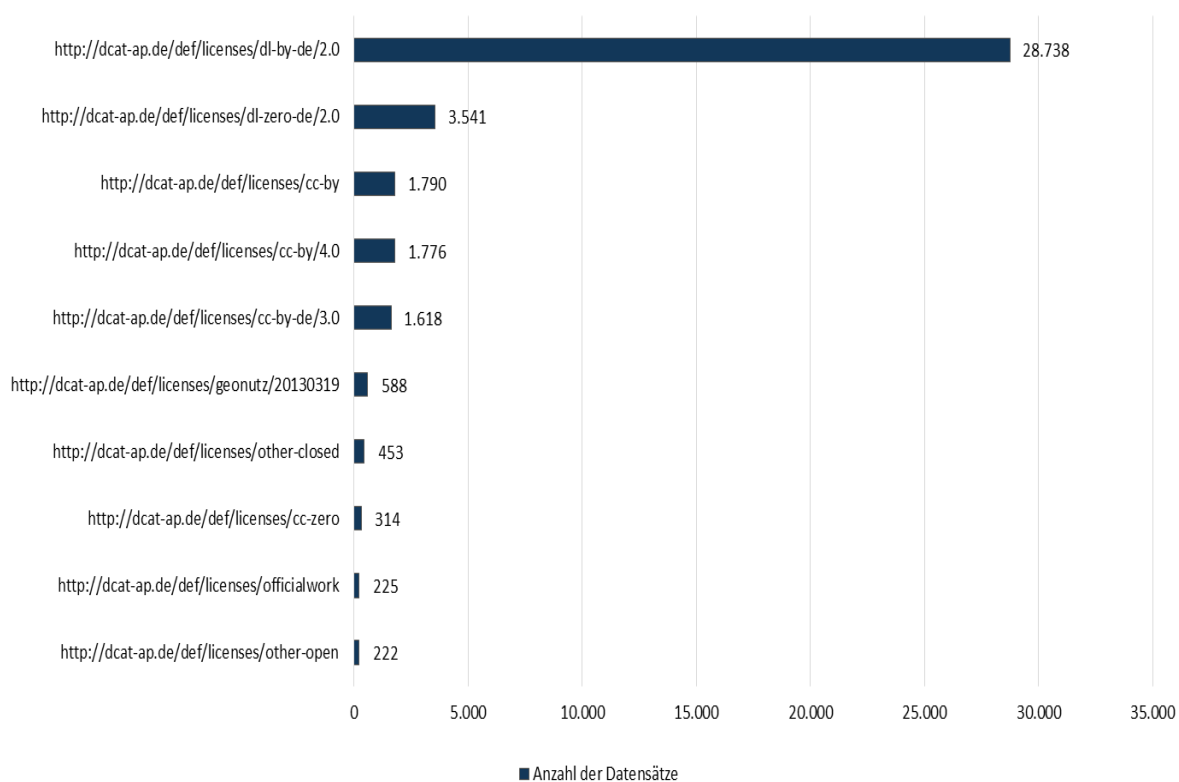
31 <https://www.govdata.de/dl-de/zero-2-0> (Abrufdatum: 19. Juli 2021).

32 <http://opendefinition.org/licenses/cc-by/> (Abrufdatum: 19. Juli 2021).

33 <https://creativecommons.org/licenses/by/4.0/> (Abrufdatum: 19. Juli 2021).

34 <https://www.dcat-ap.de/def/licenses/> (Abrufdatum: 19. Juli 2021).

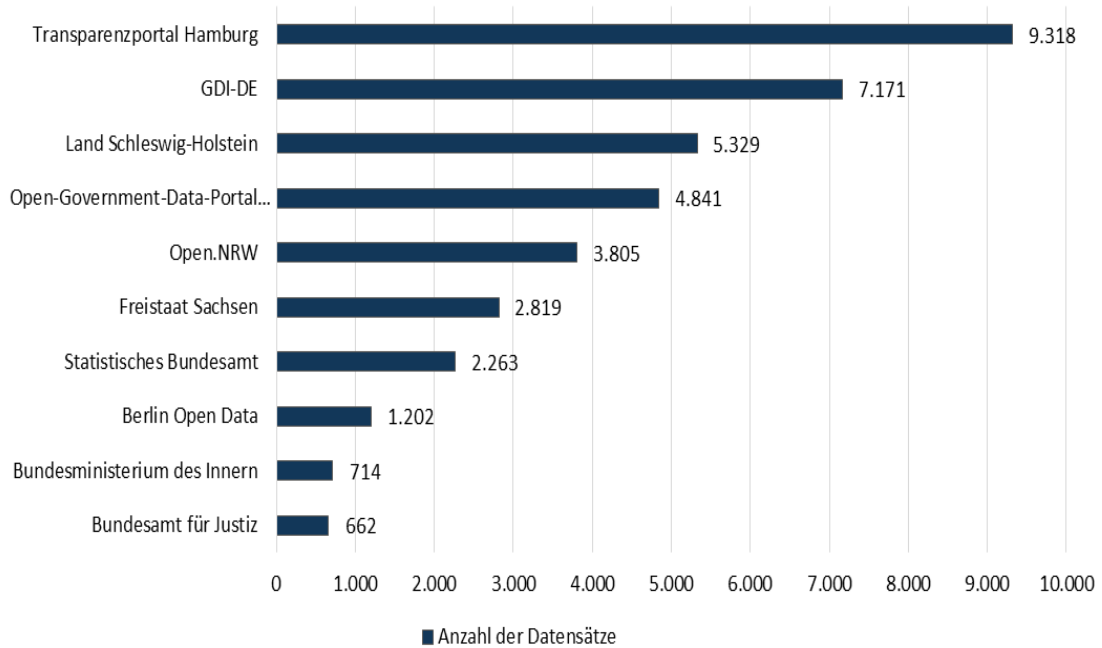
» **Abbildung 4: Datenlizenzen**



3.1.3 Zulieferung und Bereitstellung von Datensätzen

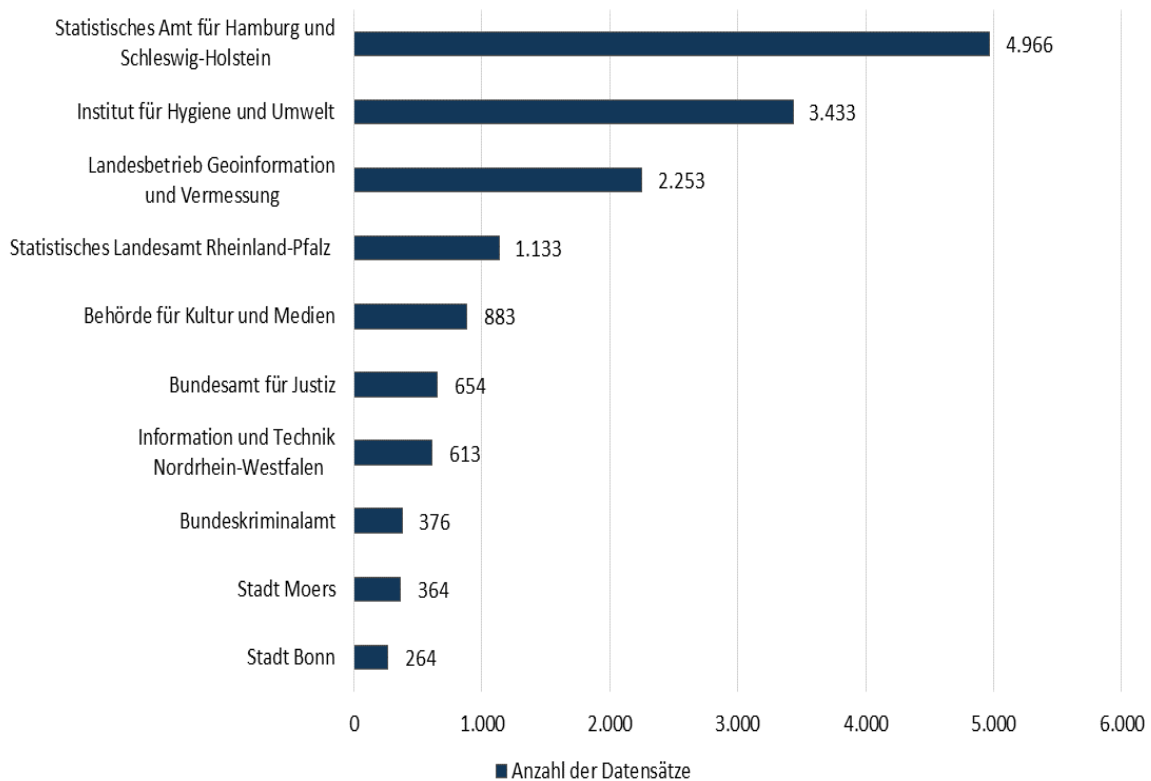
Zuletzt wurden die Metadaten in Bezug auf die Plattformen, auf denen die Datensätze bereitgestellt wurden, sowie die Datenlieferantinnen und -lieferanten analysiert. Die Ergebnisse zeigen, dass insgesamt 44 Plattformen offene Daten bereitstellten. Besonders viele Datensätze standen auf den Plattformen Transparenzportal Hamburg, Geodateninfrastruktur Deutschland (GDI-DE), Open Data Schleswig-Holstein, Open Government-Data-Portal Rheinland-Pfalz, Open.NRW sowie dem Open Data Portal des Freistaates Sachsen zur Verfügung (siehe Abbildung 5).

» **Abbildung 5: Plattformen**



Die Plattformen, die die Datensätze bereitstellten, wurden aus unterschiedlichen Quellen mit Datensätzen beliefert. Die Analyse der Metadaten ergab eine Anzahl von insgesamt 474 Quellen. Unter denen, die am meisten Datensätze lieferten, waren das Statistische Amt für Hamburg und Schleswig-Holstein, das Institut für Hygiene und Umwelt, der Landesbetrieb Geoinformation und Vermessung, das statistische Landesamt Rheinland-Pfalz und die Behörde für Kultur und Medien (siehe Abbildung 6).

» Abbildung 6: Datenlieferantinnen und -lieferanten



3.1.4 Zusammenfassung

Die Analyse der Metadaten zeigt zunächst, dass in Bezug auf die 40.749 Datensätze, die am Stichtag (01. Februar 2021) auf GovData.de verlinkt waren.

Die Kategorien, zu denen am meisten Datensätze auf GovData.de verlinkt waren, waren Umwelt sowie Regierung und öffentlicher Sektor. Sehr wenige Datensätze waren dagegen zu den Kategorien Gesundheit sowie Internationale Themen zu finden. Schlagwörter, die neben „opendata“ besonders vielen Datensätzen zugeteilt wurden, waren u. a. „bebauungsplan“ und „inspireidentifiziert“. Diese Schlagwörter deuten darauf hin, dass auf GovData.de eine große Anzahl an räumlichen Daten verlinkt sind.

Insgesamt wurden auf GovData.de Datensätze von 44 Plattformen verlinkt, auf denen offene Daten bereitgestellt waren. Hierbei war das Transparenzportal Hamburg mit 9.318 verlinkten Datensätzen das Portal mit den meisten Datensätzen. Unter den 474 Datenlieferantinnen und -lieferanten stach an erster Stelle das Statistische Amt für Hamburg und Schleswig-Holstein hervor, von dem auf GovData.de Metadaten zu insgesamt 4.966 Datensätzen zur Verfügung standen.

3.2 Analyse der Datennutzung: Matomo Analytics

Die Web-Analyseanwendung Matomo liefert strukturierte Kennzahlen über unterschiedliche Aspekte der Nutzung von GovData.de. Im Rahmen dieses Forschungsvorhabens wurde der Zeitraum zwischen dem 01. November 2020 und dem 31. Oktober 2021 beobachtet. Die Matomo-Daten wurden von

GovData.de erhoben und den Autorinnen der vorliegenden Studie übermittelt. Im Folgenden werden die Ergebnisse der deskriptiven Analyse der Matomo-Daten vorgestellt.

3.2.1 Besuche

Die Anwendung Matomo ermöglicht es, die Zahlen von Besucherinnen und Besuchern des Portals GovData.de differenziert zu erfassen.³⁵ Darüber hinaus wird die durchschnittliche Zeit der Besuche aufgezeichnet. Hierbei muss jedoch beachtet werden, dass GovData.de die Privatsphäreneinstellung der Nutzerinnen und Nutzer respektiert. Ist im Browser ein Do-Not-Track Header aktiviert, werden die entsprechenden Interaktionen nicht von Matomo registriert. Bots werden von Matomo erkannt und nicht in der Analyse berücksichtigt. Den Matomo-Daten liegt eine andere Methode der Bot-Erkennung zu Grunde als der später in diesem Bericht beschriebenen Logfiles-Analyse.³⁶ Im folgenden Kapitel finden sich die Darstellungen der Besuchszahlen sowie der Durchschnittszeiten der Besuche und der Absprungrate der Besucherinnen und Besucher.

3.2.1.1 (Neue und wiederkehrende) Besuche

Eine der zentralen Metriken der Nutzungsanalyse sind die Zahlen der Besuche auf einer entsprechenden Webseite. Matomo definiert Besuche als das Aufsuchen der analysierten Webseite durch Besuchende, welche die Webseite zum ersten Mal oder mehr als 30 Minuten nach dem ersten Besuch erneut aufsuchen.³⁷ Liegen also zwischen einem ersten und zweiten Besuch mehr als 30 Minuten, wird das wiederholte Aufrufen der Seite als zwei getrennte Besuche gezählt. Ferner kann über Matomo zwischen neuen und wiederkehrenden Besuche erfasst werden. Als „neue“ Besuche gelten Seitenaufrufe von GovData.de, die zum ersten Mal durchgeführt wurden. Jeder weitere Seitenaufruf von GovData.de am selben Tag oder zu einem späteren Zeitpunkt wird als „wiederkehrender“ Besuch gezählt.³⁸ Die Anzahl an Besuchen umfasst also die Summe aus neuen und wiederkehrenden Besuchen.

Im gesamten Erhebungszeitraum vom 01. November 2020 bis zum 31. Oktober 2021 wurden insgesamt 233.635 Besuche auf GovData.de gezählt. Für den gesamten Zeitraum sind dies durchschnittlich circa 640 Besuche am Tag. Abbildung 7 zeigt, dass mit Blick auf die Gesamtsumme der neuen und wiederkehrenden Besuche für den Zeitraum November 2020 bis Oktober 2021, dass 81 Prozent neue und lediglich 19 Prozent wiederkehrende Besuche sind.

³⁵ Matomo zeichnet lediglich Besucherinnen und Besucher auf, die Javascript einsetzen. Da dies für die meisten Bots nicht gilt, kann davon ausgegangen werden, dass Matomo kaum Bots aufzeichnet und die dargestellten Metriken auf tatsächliche Besucherinnen und Besucher und nicht auf Bots zurückzuführen sind.

³⁶ Die durch Matomo angewendete Methode wurde an den Logfiles getestet, wurde jedoch als weniger zielführend eingestuft als die schlussendlich verwendete Methode. Zur entsprechenden Matomo-Methode siehe auch: <https://github.com/matomo-org/device-detector>.

³⁷ Matomo (2020a).

³⁸ Vgl. Matomo (2020a).

» **Abbildung 7: Anteile neuer und wiederkehrender Besuche**

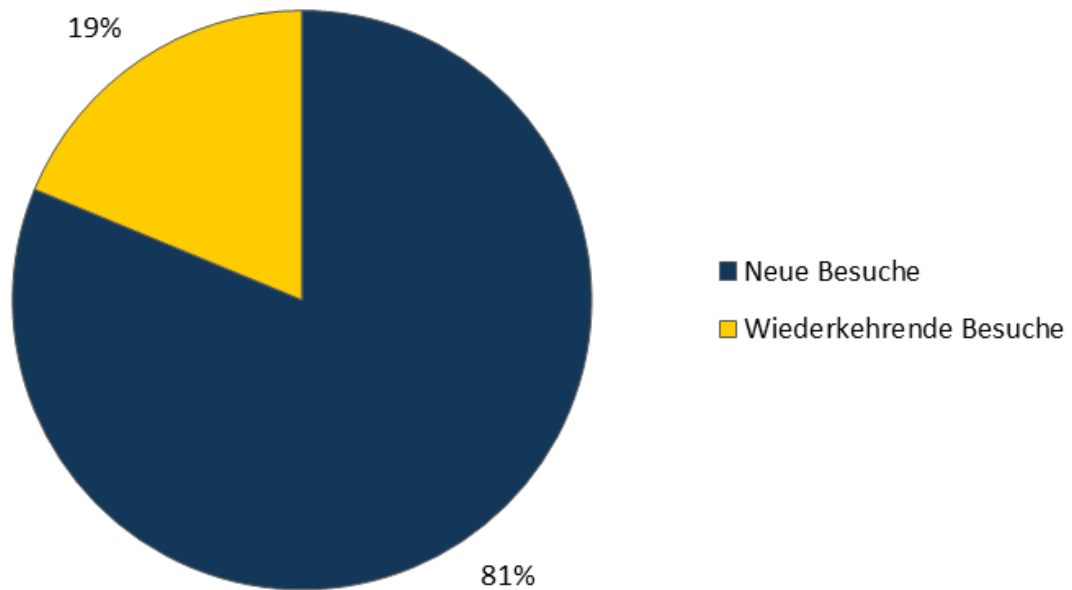
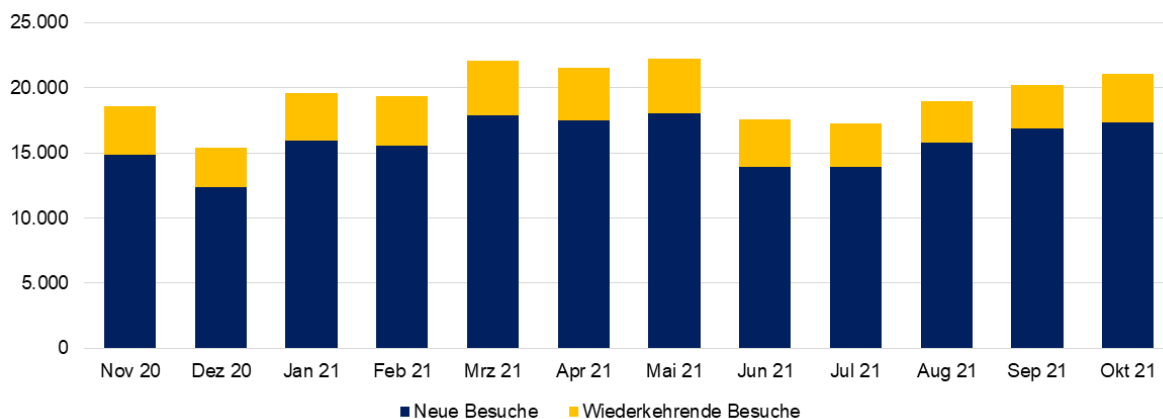


Abbildung 8 zeigt die monatlichen Summen der Besuche im Zeitraum November 2020 bis Oktober 2021 bestehend aus neuen und wiederkehrenden Besuchen. Monatlich schwankt die Gesamtanzahl der Besuche zwischen circa 15.000 und 22.000 Besuchen im Monat. Die besuchsstärksten Monate sind März 2021 bis Mai 2021; am besuchtschwächsten im Vergleich ist der Dezember 2020 sowie die Monate Juni und Juli 2021. Damit fällt auf, dass die besuchsschwächeren Monate insbesondere in Ferienzeiten liegen – dies kann eine Erklärung für die Abnahme der Besuchszahlen in diesen Zeiträumen sein. Ein klarer zunehmender oder abnehmender Trend lässt sich weder für die Gesamtsumme der Besuche noch für die Anteile der neuen und wiederkehrenden Besuche erkennen.

» **Abbildung 8: Monatliche Summe der Besuche**



Die Anwendung Matomo bietet darüber hinaus die Möglichkeit, Besucherinnen und Besucher, die mehrmals an einem Tag das Portal GovData.de besuchen, wieder zu erkennen und aus der Zählung auszuschließen.³⁹ Die Zählung „eindeutige Besucherinnen und Besucher“ stellt somit diejenige Anzahl von Besucherinnen und Besucher dar, in der jede Besucherin bzw. jeder Besucher nur einmal gezählt wird, auch wenn diese GovData.de mehrmals am Tag besuchen.⁴⁰ Jedoch unterscheidet sich die Gesamtsumme der Besuche nicht erheblich von der Gesamtsumme der eindeutigen Besucherinnen und Besucher im Zeitraum November 2020 bis Oktober 2021: 219.734 eindeutige Besucherinnen und Besucher wurden im Zeitraum gezählt (im Vergleich zu 233.635 Besuchen). Dabei sind es, der Logik der Definition eindeutiger Besucherinnen und Besuchern folgend, vor allem wiederkehrende Besucherinnen und Besucher, die aus der Zählung ausgeschlossen werden. Aufgrund des geringen Unterschieds wird nicht näher auf die eindeutigen Besucherinnen und Besucher eingegangen.

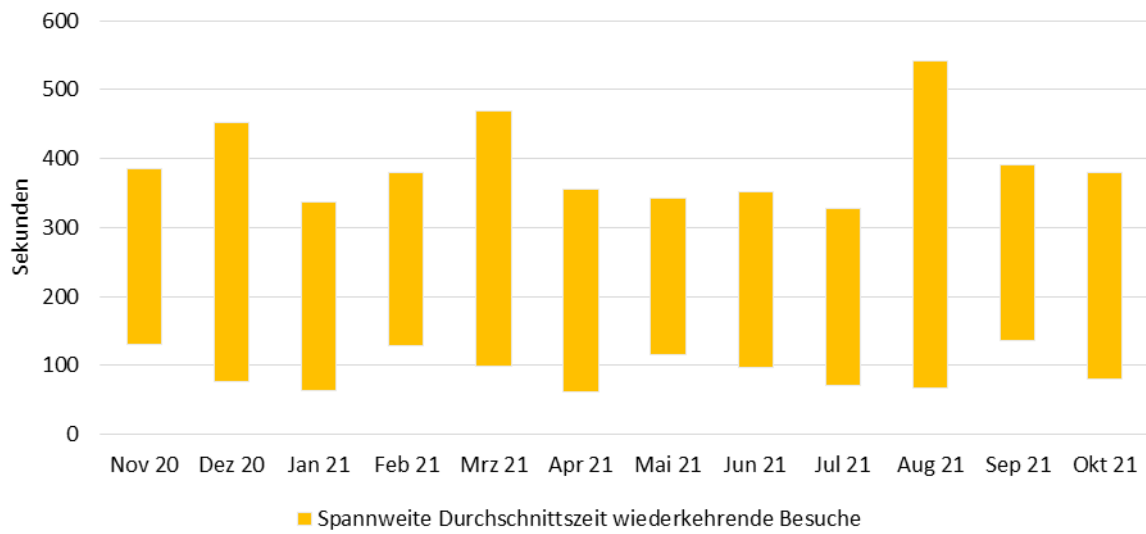
3.2.1.2 Durchschnittliche Besuchszeit

Matomo zeichnet darüber hinaus die durchschnittliche Zeit der Besuche eines Tages auf. Die kürzeste durchschnittliche Besuchszeit eines Tages im Beobachtungszeitraum betrug 58 Sekunden, die längste durchschnittliche Besuchszeit 191 Sekunden. Abbildung 9 und Abbildung 10 zeigen jeweils die Spannbreite zwischen den minimalen und maximalen Zeiten bei wiederkehrenden und neuen Besuchen in den Monaten des Beobachtungszeitraums. Bei neuen Besuchen liegen Minimal- und Maximalwert (deutlich) niedriger als bei den wiederkehrenden Besuchen. Bei neuen Besuchen beträgt die geringste durchschnittliche Besuchszeit eines Monats 51 Sekunden, die maximale Besuchsdauer bei 161 Sekunden. Bei wiederkehrenden Besuchen war die kürzeste Besuchszeit eines Monats 61 Sekunden, die längste durchschnittliche Besuchszeit eines Monats 541 Sekunden – und ist damit deutlich länger als bei neuen Besuchen. Da es sich bei diesen Werten um arithmetische Mittel handelt, kann hierbei nicht auf einzelne Besucherinnen und Besucher geschlossen werden. Möglich ist sowohl, dass alle wiederkehrenden bzw. neuen Besuche eine entsprechende Dauer aufweisen, als auch, dass einzelne Ausreißer hinsichtlich Kürze oder Länge des Besuchs die Durchschnittswerte prägen. Es kann nicht komplett ausgeschlossen werden, dass in die Zählung auch Zugriffe durch Bots einfließen, wenn deren Verhalten ein natürliches Nutzungsverhalten glaubhaft imitiert. Dies kann sich dahingehend auswirken, dass die Besuchszeit bei wiederkehrenden Besuchen höher liegt.

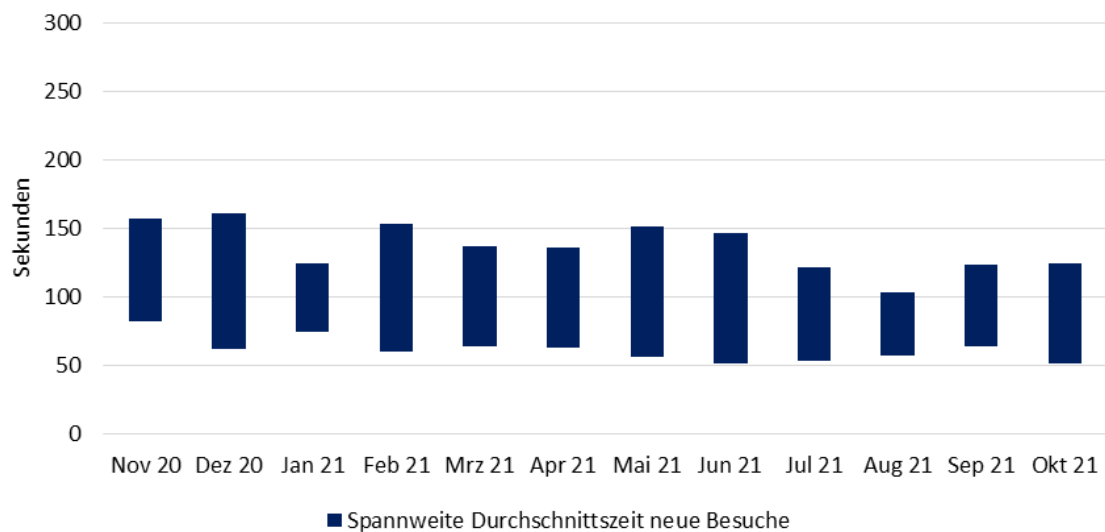
³⁹ Besucherinnen und Besucher werden durch Matomo vorrangig über Cookies einander zugeordnet, um eindeutige Besucherinnen und Besucher wiederzuerkennen. Werden keine Cookies verwendet bzw. werden diese durch die Nutzerinnen und Nutzer geblockt, ordnet Matomo eindeutige Besucherinnen und Besucher durch wiedererkennbare Eigenschaften (IP, Browser, Plug-Ins etc.) zu. Damit versucht die Software, Besucherinnen und Besucher zu erkennen, die mehr als einmal am Tag eine Webseite besucht haben. Vgl. Matomo (2020b).

⁴⁰ Vgl. Matomo (2020b). Siehe im Vergleich dazu die Definition der Besuche.

» **Abbildung 9: Durchschnittliche Besuchszeit bei wiederkehrenden Besuchen**



» **Abbildung 10: Durchschnittliche Besuchszeit bei neuen Besuchen**



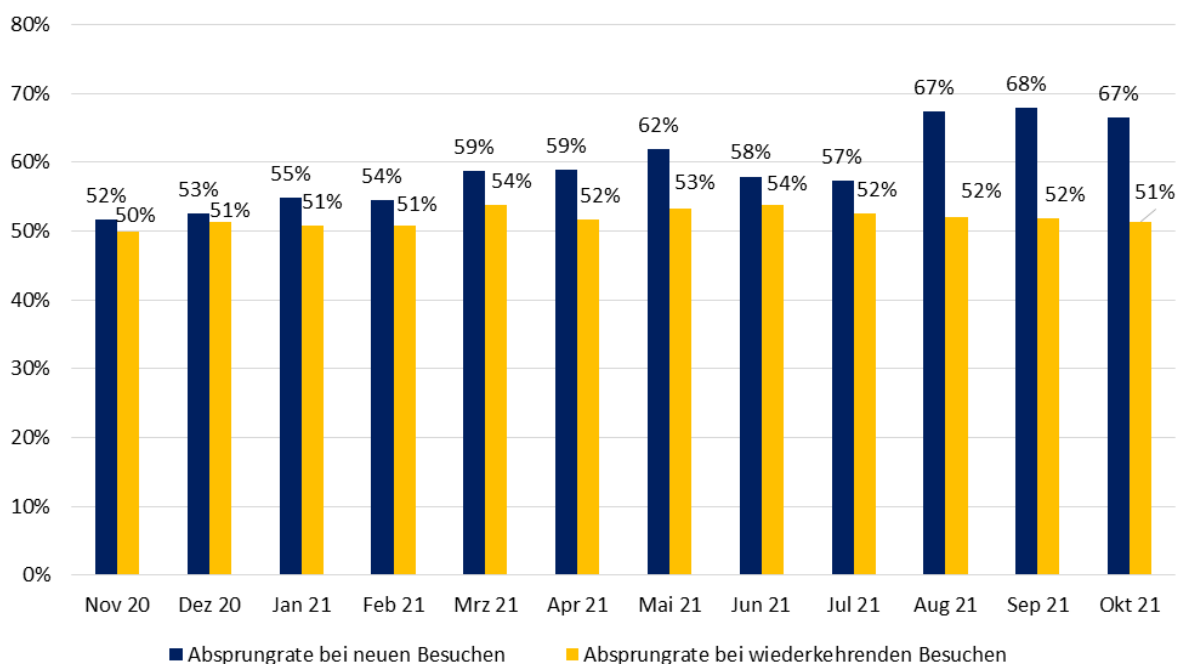
3.2.1.3 Absprungrate

Die Absprungrate sagt aus, wie viel Prozent der Besucherinnen und Besucher GovData.de nach dem Aufrufen der ersten aufgerufenen Unterseite direkt wieder verlassen und keine weiteren Aktionen durchführen.⁴¹ Die Absprungrate liegt im Durchschnitt des gesamten Beobachtungszeitraums bei 58 Prozent der Besuche; ähnlich verhält es sich bei der Rate der Absprünge bei Besuchen von neuen

⁴¹ Vgl. Matomo (2020c).

Besucherinnen und Besuchern (59 Prozent), die Rate bei wiederkehrenden Besucherinnen und Besuchern ist mit 52 Prozent etwas geringer. Abbildung 11 zeigt die Absprungraten bei neuen und wiederkehrenden Besuchen im Verlauf des Beobachtungszeitraums. Die Raten unterscheiden sich dabei im Verlauf nur geringfügig, ein Trend für die unterschiedlichen Besuchsmetriken ist im Verlauf nicht zu erkennen.

» **Abbildung 11: Absprungrate bei (neuen und wiederkehrenden) Besuchen**



Die Absprungraten zeigen, dass weniger als die Hälfte der (neuen und wiederkehrenden) Besucherinnen und Besucher GovData.de über die Einstiegsseite hinaus nutzt. Warum die Besucherinnen und Besucher direkt wieder von der ersten aufgerufenen Unterseite abspringen, können die Daten nicht erklären.

3.2.2 Durchgeführte Aktionen der Besucherinnen und Besucher

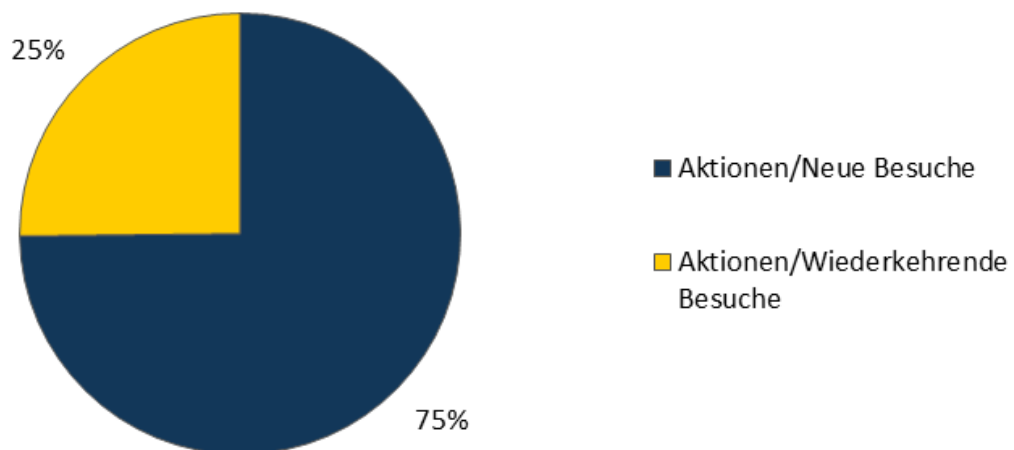
Neben den Besuchs-Kennzahlen wird auch die Anzahl (unterschiedlicher) Aktionen durch Matomo gespeichert. Die Analyseanwendung erlaubt eine differenzierte Darstellung der Anzahl unterschiedlicher Aktionen. Aktionen, die in diesem Fall von Matomo gezählt werden, umfassen sowohl Seitenansichten als auch Downloads und das Anklicken von ausgehenden Verweisen. In [Kapitel 3.3](#) werden auf Grundlage der Logfiles weitere Aktionen der Besucherinnen und Besucher, die über die möglichen Aussagen auf Grundlage der Informationen der Matomo-Daten hinausgehen, dargestellt.

3.2.2.1 Anzahl der Aktionen

Auch in Bezug auf die Aktionen besteht die Gesamtsumme aus Aktionen von neuen und wiederkehrenden Besucherinnen und Besuchern. Die Anzahl der Aktionen können in Relation zur Anzahl der (neuen und wiederkehrenden) Besuche gesetzt werden und geben somit Aufschluss darüber, wie viele Aktionen pro (neuem oder wiederkehrendem) Besuch von den Nutzerinnen und Nutzern durchgeführt wurden.

Die Gesamtzahl der Aktionen im gesamten Beobachtungszeitraum wird dabei von Aktionen neuer Besucherinnen und Besucher dominiert (siehe Abbildung 12). Diese Zahl ist jedoch wenig aussagekräftig, da auch, wie oben bereits erläutert, 81 Prozent der Besuche von neuen Besucherinnen und Besuchern durchgeführt wird.

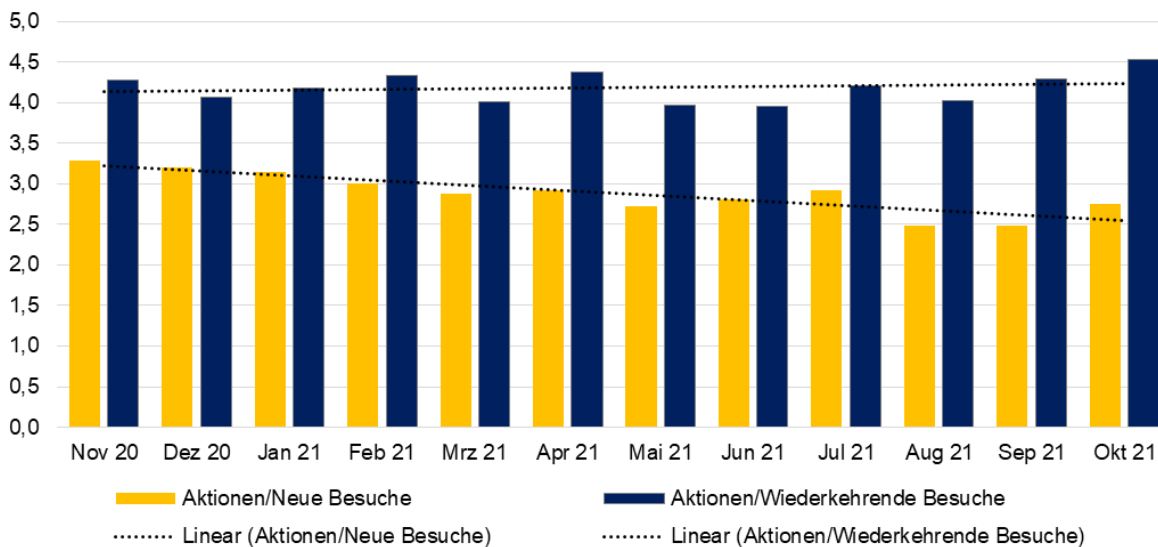
» **Abbildung 12: Anteile der Aktionen von neuen und wiederkehrenden Besuchen**



Aussagekräftiger ist daher die Relation zwischen der Anzahl der Besuche und den dabei durchgeführten Aktionen. Im gesamten Zeitraum liegt die durchschnittliche Anzahl pro Besuch bei (circa) drei Aktionen je Besuch. Anhand des Vergleichs zwischen neuen und wiederkehrenden Besucherinnen und Besuchern zeigt sich, dass wiederkehrende Besucherinnen und Besucher mehr Aktionen durchführen (im Gesamtdurchschnitt circa vier Aktionen) als neue Besucherinnen und Besucher (im Gesamtdurchschnitt drei Aktionen). Abbildung 13 zeigt die durchschnittlichen Aktionen pro Besuch von neuen und wiederkehrenden Besucherinnen und Besuchern zwischen November 2020 und Oktober 2021. Während die Aktionen je Besuch von wiederkehrenden Besucherinnen und Besuchern im Jahresverlauf gleichbleibend ist, nimmt die Anzahl der Besuche von neuen Besucherinnen und Besuchern im Jahresverlauf ein wenig ab.

Somit lässt sich auch die längere Durchschnittszeit pro Besuch (siehe [Kapitel 3.2.1.2](#)) anhand der mehr durchgeführten Aktionen pro Besuch von wiederkehrenden Besucherinnen und Besuchern besser einordnen.

» **Abbildung 13: Aktionen pro neuen und wiederkehrenden Besuchen**



Wie die Mehraktivitäten der wiederkehrenden Besucherinnen und Besucher zu erklären sind, geht aus den Daten nicht hervor. Eine Möglichkeit wäre, dass wiederkehrende Besucherinnen und Besucher bereits häufiger und regelmäßiger mit der Plattform GovData.de gearbeitet haben bzw. die Plattform mit einem klaren Ziel nutzen. Neue Besucherinnen und Besucher hingegen könnten häufiger lediglich aus grundlegendem Interesse die Webseite aufrufen und führen über das Aufrufen der Einstiegsseite weniger Aktionen durch. Auch versehentliche Aufrufe der Webseite ohne Intention des Besuchs von GovData.de sind möglich und hier zu berücksichtigen.

Auch wenn diese Zahlen einen Eindruck des Nutzungsverhaltens der Besucherinnen und Besucher vermitteln, sind die Aussagen, die aufgrund dieser Ergebnisse getroffen werden können, in ihrer Aussagekraft begrenzt. So können zwar die Durchschnittswerte verschiedener Metriken ausgewiesen werden, wie sich jedoch die gesamte Anzahl der Aktionen auf die Besucherinnen und Besucher verteilt, bleibt dabei unklar. Möglich ist bspw., dass einige wenige Besucherinnen und Besucher (regelmäßig) viele Aktionen durchführen, während die meisten Besucherinnen und Besucher nur wenige Aktionen durchführen.

3.2.2.2 Aktionen: Seitenaufrufe, Downloads und ausgehende Verweise

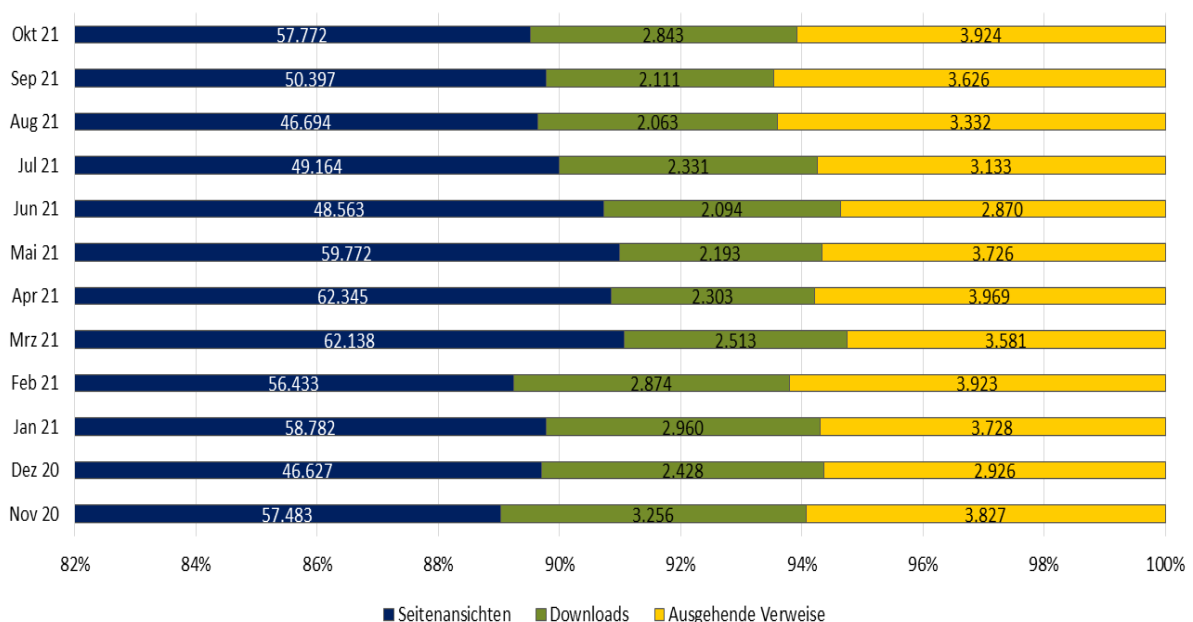
Es lagen ebenfalls Daten darüber vor, welche Aktionen auf GovData.de durchgeführt wurden. Matomo zeichnet drei verschiedene Aktionen auf: Seitenaufrufe, Downloads und ausgehende Verweise.⁴² Da die Datensätze nicht auf den Servern von GovData.de selbst gespeichert sind, werden die Besuchenden zum Herunterladen der Datensätze auf die Webseiten der Bereitstellerinnen und Bereitsteller geleitet und in der Metrik als „ausgehende Verweise“ widergespiegelt. Ein Vergleich der jeweiligen Aktionen

⁴² Laut Matomo-Webseite können auch Suchen als Aktionen aufgezeichnet werden. Die Anzahl der Suchaktionen lag jedoch für GovData.de nicht vor. Die Gesamtanzahl der Aktionen verteilte sich lediglich auf die Anteile der Seitenaufrufe, Downloads und ausgehenden Verweise.

zwischen neuen und wiederkehrenden Besucherinnen und Besuchern ist hier nicht möglich, da die entsprechenden Daten nicht erhoben wurden.

Abbildung 14 zeigt die prozentualen Anteile der Aktionen in Seitenaufrufen, Downloads und ausgehenden Verweisen. Seitenaufrufe machen mit circa 90 Prozent in allen Monaten des Beobachtungszeitraums den Großteil der Aktionen aus. Circa 15 Prozent der Seitenaufrufe (in den jeweiligen Monaten sowie im Gesamtzeitraum) beziehen sich dabei auf die Startseite. Welche Seiten von Besucherinnen und Besuchern aufgerufen wurden, kann anhand der Logfiles näher bestimmt werden (siehe [Kapitel 3.3.3](#)). Downloads und ausgehende Verweise spielen gleichermaßen eine geringere Rolle. Im Überblick über alle Monate des Beobachtungszeitraums zeigen sich hier nur marginale Verschiebungen zwischen den Anteilen der Aktionen.

» **Abbildung 14: Durchgeführte Aktionen im Gesamtzeitraum**



Diese Zahlen zeigen deutlich, dass nur der geringere Anteil der Aktionen auf den Download von Datensätzen abzielt. Auch hier können die Matomo-Daten keine Erklärung geben. Möglich ist, dass die dargebotenen Datensätze hinsichtlich des Inhalts oder der Aktualität nicht den gesuchten Datensätzen entsprachen. Möglich ist auch, dass die Besucherinnen und Besucher kein Interesse an den Datensätzen an sich, sondern nur an Informationen über die Datensätze oder offenen Daten im Allgemeinen hatten. Eine abschließende Beurteilung und Einschätzung der Werte kann aufgrund fehlender Vergleiche mit ähnlichen Webseiten nicht vorgenommen werden. Mit einer Umfrage unter Besucherinnen und Besuchern von GovData.de wurde versucht, einen Einblick in mögliche Erklärungen für die Hintergründe der Aktionen von Besucherinnen und Besuchern zu gewinnen (siehe [Kapitel 3.4](#)).

3.2.3 Übergänge der Besuche

Matomo ermöglicht es, die Ursprünge der Besuche bzw. bestimmter Seitenaufrufe zu identifizieren. Im Folgenden werden die Ursprungsseiten der Besuche auf allen Unterseiten von GovData.de dargestellt.

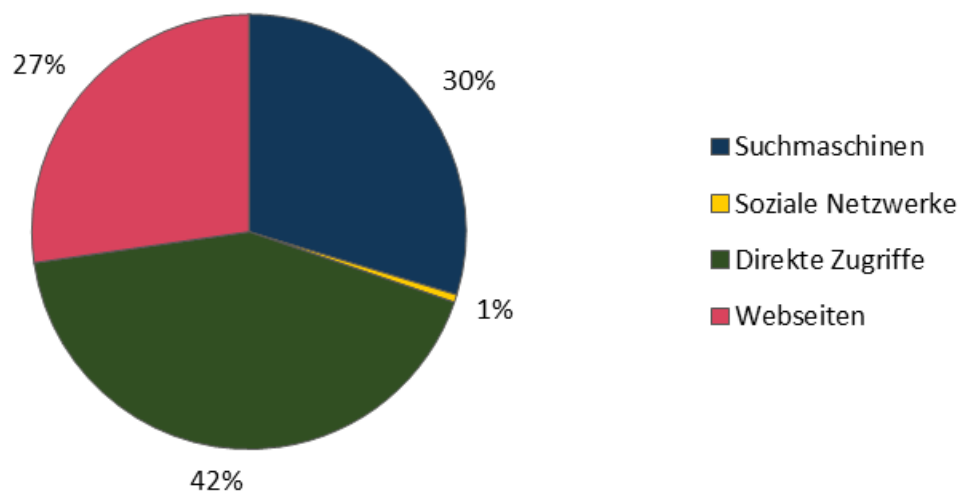
3.2.3.1 Einstiegsrouten aller (Unter-)Seiten

Ausgehend von allen GovData.de-Seiten kann die Gesamtzahl der Besuche im Gesamtzeitraum auf die folgenden Einstiegsrouten zurückgeführt werden:

- » Suchmaschinen
- » Soziale Netzwerke
- » Webseiten
- » Direkte Zugriffe

Abbildung 15 stellt die Anteile der jeweiligen Einstiegsrouten gemessen an der Zahl der Besuche im gesamten Beobachtungszeitraum dar.

» **Abbildung 15: Einstiegsrouten der Besuche im Gesamtzeitraum**



Die meisten Einstiege erfolgen als direkte Zugriffe (42 Prozent) auf die verschiedenen (Unter-) Seiten von GovData.de. Ungefähr gleich häufig wurden andere Webseiten (27 Prozent) und Suchmaschinen (30 Prozent) als Einstiegsrouten identifiziert. Soziale Netzwerke (ein Prozent) hingegen spielen nur eine geringe Rolle als Ursprung der Besuche.

3.2.4 Zusammenfassung

Von November 2020 bis Oktober 2021 wurden insgesamt 233.635 Besuche auf GovData.de registriert, der Großteil davon sind neue Besuche. Im Beobachtungszeitraum sind die Besuchszahlen neuer und wiederkehrender Besuche weder merklich gestiegen noch gesunken. Über die Hälfte der Besucherinnen und Besucher verlässt die erste aufgerufene Unterseite von GovData.de direkt wieder und führt keinen weiteren Aktionen durch.

Der Vergleich zwischen neuen und wiederkehrenden Besuchen zeigte, dass wiederkehrende Besuche zum einen eine längere Dauer haben und zum anderen mehr Aktionen pro Besuch durchgeführt werden. Diese Beobachtung zwischen wiederkehrenden und neuen Besuchen bleibt über den Beobachtungszeitraum konstant. Dies lässt jedoch auch nur geringfügige Schlüsse auf die Verwendung der Daten zu: Möglich ist es bspw., dass regelmäßig mit offenen Daten gearbeitet wird, die im Rahmen eines Besuches gezielt gesucht und heruntergeladen werden. Die Aktionen wiederkehrender Besuche nehmen in der Tendenz leicht zu; die Anzahl der Aktionen bei neuen Besuchen ist im beobachteten Zeitraum wiederum gesunken. Die Aktionen, die auf GovData.de durchgeführt werden, sind zu über 90 Prozent Seitenaufrufe. Das zeigt, dass nur ein geringer Teil der Nutzung der Plattform auf das Herunterladen eines Datensatzes hinausläuft. Möglich ist, dass die Besucherinnen und Besucher von GovData.de die gesuchten Daten finden und somit der Such- und Aufrufprozess die Vielzahl der Seitenaufrufe verursacht. Alternativ kann die geringe Anzahl von Downloads und ausgehenden Verweisen auf unzureichende und fehlende Datensätze hinweisen. Das könnte bedeuten, dass Besucherinnen und Besucher die gesuchten Daten nicht finden oder die vorhandenen Datensätze als für ihre Zwecke unbrauchbar einschätzen. Hinsichtlich der Einstiegsrouten wurde deutlich, dass die Besucherinnen und Besucher hauptsächlich direkt auf GovData.de zugreifen oder durch Suchmaschinen auf die Seiten von GovData.de gelangen. Soziale Netzwerke sind im Gegensatz dazu nur wenig relevant.

Die deskriptive Analyse der Matomo-Daten gibt Aufschluss über einige Metriken der Nutzung von GovData.de. Eine Einordnung der Ergebnisse bleibt jedoch aufgrund fehlender Vergleichsstudien ähnlicher Plattformen bzw. Vergleichszeiträume schwierig. Eine kontinuierliche Beobachtung und Auswertung der Matomo-Daten kann hilfreich sein, um längerfristige Trends in den unterschiedlichen Metriken der Nutzung zu erkennen und die vorliegenden Ergebnisse besser einordnen zu können, insbesondere bezüglich der Anteile der unterschiedlichen Aktionen, die bei Besuchen durchgeführt werden.

3.3 Analyse der Datennutzung: Logfiles

Um einen möglichst vollständigen Eindruck von der bisherigen Nutzung der Webseite GovData.de zu erlangen, wurde ergänzend zu der Auswertung der Matomo-Daten auch eine Analyse der in den Logfiles der Seite festgehaltenen sogenannten GET-Requests durchgeführt. Diese Analyse diente in erster Linie einer differenzierteren Betrachtung der Nachfrage nach den verschiedenen Inhalten der Seite.

Hierzu wurden Logfiles aus dem Zeitraum vom 30. Juli 2020 bis zum 18. Januar 2021 in einem anonymisierten Format an das FÖV übermittelt. Dementsprechend erfolgt keine Auswertung der Logfiles nach Nutzerinnen und Nutzern, sondern lediglich eine Betrachtung der aufgezeichneten Interaktionen. Im gesamten Zeitraum wurden über 27 Millionen Interaktionen mit der Seite festgestellt. An drei Tagen zwischen dem 18. Dezember 2020 und dem 20. Dezember 2020 konnten aufgrund eines Problems bei der Erhebung der Logfiles keine Daten übermittelt werden. Unter die in

den Logfiles aufgezeichneten Interaktionen fallen weitaus mehr Interaktionen als von Matomo erfasst werden, da automatisierte Abfragen etc. in den Logfiles protokolliert, von Matomo hingegen ignoriert werden.

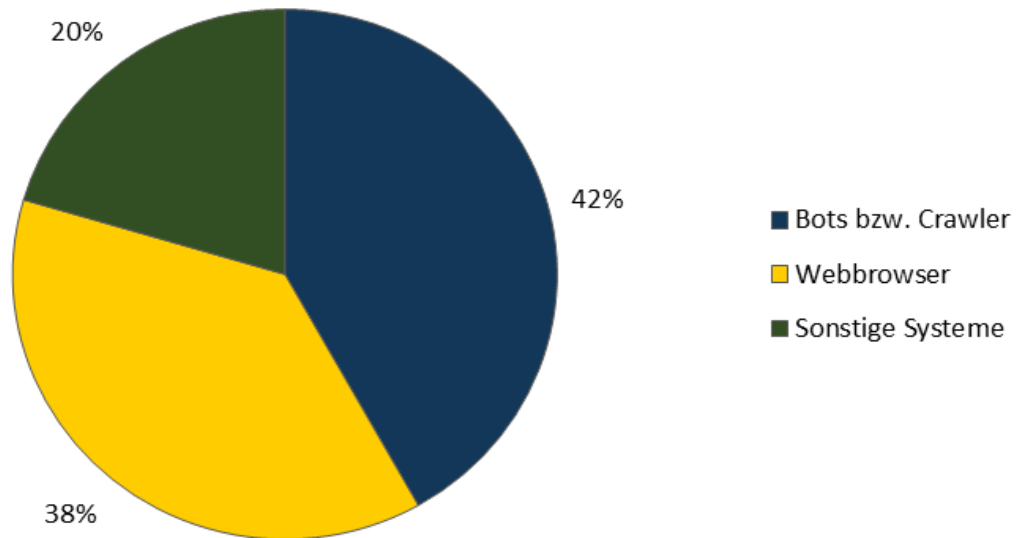
3.3.1 Zusammensetzung der Anfragen nach User-Client

Die Logfiles geben u. a. Auskunft darüber, welches System zur Abfrage der Seite genutzt wurde. Anhand dieser Variable ist im Regelfall erkennbar, ob es sich um eine Abfrage über einen Webbrowser oder um eine automatisierte Abfrage handelt. Hierzu ist eine Kategorisierung der aufgefundenen User-Agents in Webbrowser und automatisierten Abfragen notwendig. Neben der Unterscheidung zwischen Webbrowsern und identifizierten Bots wurden zur Analyse noch die Kategorie „sonstige Systeme“ eingefügt. Hierunter wurden Abfragen aus Programmen gefasst, die üblicherweise für automatisierte Abfragen genutzt werden.

Es muss bei der Einstufung der genutzten Systeme nach User-Agent davon ausgegangen werden, dass diese Unterscheidung nicht zu 100 Prozent genau ist. Im Wesentlichen basiert die Einstufung auf der freiwilligen Kennzeichnung der Bots in den User-Agents der abfragenden Systeme durch die Betreiberinnen und Betreiber, zudem wurden die vorkommenden User-Agents mit einer Liste von User-Agents bekannter Bots abgeglichen. Allerdings bestehen verschiedene Möglichkeiten, Bot-Aktivitäten zu tarnen, sodass wahrscheinlich nicht alle automatisierten Abfragen als solche erkannt wurden. In einigen Fällen wurden Interaktionen als Bot-Aktivität klassifiziert, obwohl der User-Agent einem Webbrowser entsprach.⁴³ Abbildung 16 zeigt die Zusammensetzung aller beobachteten Interaktionen mit der Seite, außer den Abfragen über die API, nach den drei unterschiedenen Typen von abfragenden Systemen.

⁴³ Die entsprechende Einstufung erfolgte basierend auf der Verhaltensweise: Sehr viele Abfragen mit gleichen oder ähnlichen Einstellungen durch das gleiche System im Abstand von wenigen Millisekunden.

» **Abbildung 16: Zusammensetzung der GET-Requests nach abfragendem System**



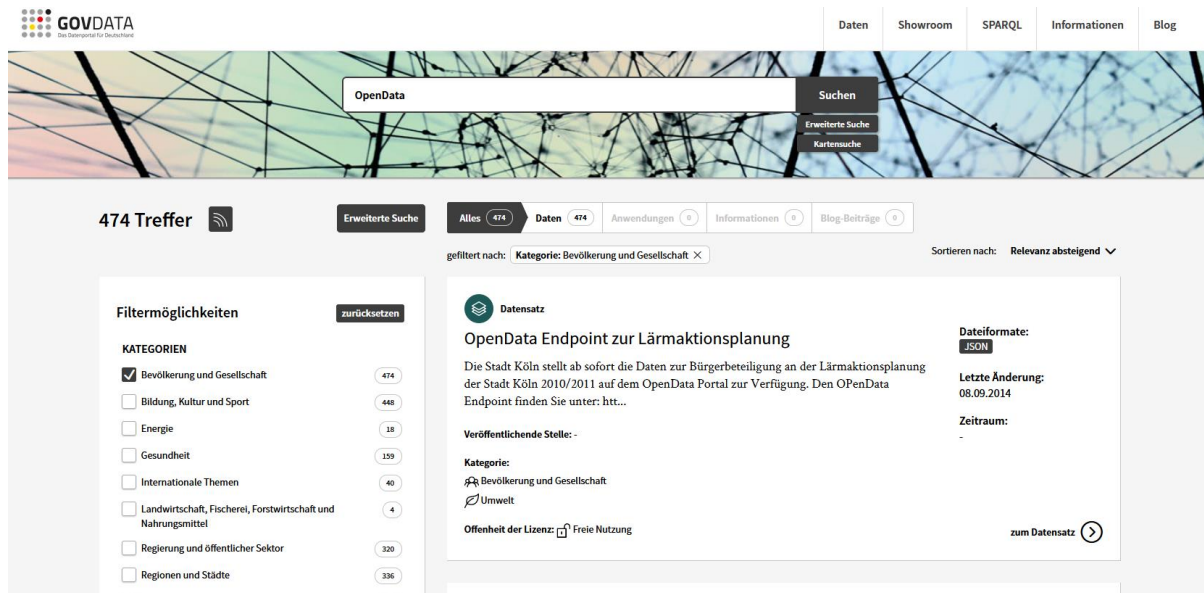
Dementsprechend wurden nur 38 Prozent der GET-Requests von Webbrowsern verursacht, der überwiegende Anteil der Abfragen erfolgte automatisiert. Dies muss jedoch nicht bedeuten, dass damit keine zielgerichtete Nutzung der bereitgestellten Informationen verbunden ist. Es ist jedoch zu beachten, dass mit einer automatisierten Abfrage in kurzer Zeit sehr viele einzelne Abfragen abgesetzt werden können, dementsprechend fallen diese Abfragen proportional mehr ins Gewicht als die aufgezeichneten manuellen Interaktionen menschlicher Nutzerinnen und Nutzer mit der Seite.

Die Anfragen durch Bots und sonstige Systeme werden im letzten Abschnitt dieses Kapitels überblicksartig dargestellt. Alle weiteren Abschnitte zu den in den Logfiles festgehaltenen Interaktionen mit der Website konzentrieren sich auf die durch Webbrowser erzeugten Anfragen, basierend auf der Annahme, dass diese im Regelfall durch menschliche Nutzerinnen und Nutzer erzeugt wurden. Die Aktivitäten von Bots und sonstigen Systemen werden sonst nur im Kapitel zu den Abfragen über die API berücksichtigt, da die API auf automatisierte Abfragen ausgelegt ist.

3.3.2 Aufgerufene Suchergebnisseiten

Die eingegebenen Suchbegriffe und die Betrachtung der verwendeten Suchvariablen können Auskunft darüber geben, an welchen Inhalten die Nutzerinnen und Nutzer Interesse zeigen, selbst wenn hierzu keine Inhalte gefunden wurden. Insgesamt wurden im gesamten Zeitverlauf 253.201 sogenannte Such-Requests abgesetzt. Damit sind Anfragen über die Suchfunktion der Seite gemeint, die Abfragen über die API wurden hier nicht miteingerechnet; diese werden in [Kapitel 3.3.5](#) näher betrachtet. Abbildung 17 zeigt eine beispielhafte Suchergebnisseite auf Govdata.de, welche durch einen Such-Request abgefragt wurde.

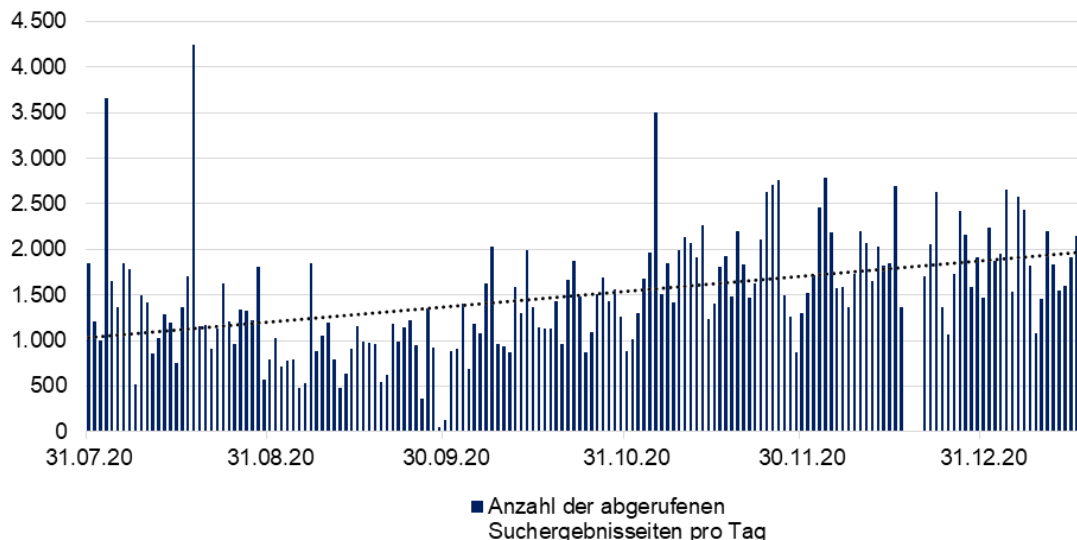
» Abbildung 17: Ausschnitt aus einer Suchergebnisseite



Der zur Abbildung gehörende Such-Request lautet: „GET /web/guest/suchen/-/searchresult/q/OpenData/f/groups%3Asoცი%2C/s/relevance_desc HTTP/1.1“. Die gewählten Filtereinstellungen und verwendeten Suchbegriffe sind also aus dem GET-Request ablesbar.

Abbildung 18 zeigt die Anzahl an Such-Requests pro Tag. Im Zeitverlauf ist eine leicht zunehmende Tendenz an Suchanfragen zu beobachten.

» Abbildung 18: Anzahl an Such-Requests pro Tag



Die verwendeten Suchbegriffe bei der Nutzung der Suchfunktion der Seite können Auskunft über die generellen Interessensbereiche der Nutzerinnen und Nutzer liefern. Im gesamten Zeitraum wurden bei 64.926 Suchanfragen einer oder mehrere Suchbegriffe eingegeben. Dies entspricht einem Anteil von 26 Prozent aller Suchanfragen. In den restlichen Fällen wurde die Suche lediglich anhand von sonstigen

Filtereinstellungen konkretisiert oder es wurde überhaupt keine Einschränkung vorgenommen. Insgesamt konnten 20.269 unterschiedliche Suchbegriffe identifiziert werden. Tabelle 1 zeigt die am häufigsten verwendeten Suchbegriffe.

» **Tabelle 1: Die häufigsten gewählten Suchbegriffe bei Nutzung der Suchfunktion**

	Suchbegriff	Verwendungshäufigkeit
1	COVID-19	862
2	corona	848
3	Corona	662
4	Baden-Württemberg	373
5	Brandenburg	369
6	covid	358
7	Niedersachsen	356
8	Kriminalstatistik	354
9	Bayern	346
10	Hessen	344
11	4	313
12	Berlin	281
13	Schleswig-Holstein	263
14	berlin	253
15	münster	250

Bei der Nutzung der Suchfunktion kann auch eine Kombination aus verschiedenen Suchbegriffen verwendet werden z. B. „Corona in Niedersachsen“. Während in Tabelle 1 diese Kombination aus Begriffen als ein Suchbegriff gewertet wurde, zeigt Tabelle 2 die am häufigsten verwendeten Bestandteile von Suchbegriffen.

» **Tabelle 2: Die häufigsten gewählten Bestandteile von Suchbegriffen bei Nutzung der Suchfunktion**

	Als Bestandteil von Suchbegriffen verwendete Begriffe	Verwendungshäufigkeit
1	corona	1.029
2	COVID-19	947
3	Corona	827
4	berlin	688
5	Kriminalstatistik	639
6	Niedersachsen	599
7	Brandenburg	594
8	Berlin	582
9	2019	536
10	Deutschland	504
11	covid	484
12	Baden-Württemberg	481
13	hamburg	470
14	Bayern	454
15	und	447

Die Betrachtung der verwendeten Suchbegriffe und deren Bestandteile weist eindeutig auf ein großes Interesse der Nutzerinnen und Nutzer an dem Themenbereich Covid-19 hin. Darüber hinaus finden sich viele Suchbegriffe mit regionalem Bezug, es wurde also speziell nach Daten zu einer bestimmten Region gesucht. Die Seite bietet als Filter auch eine Eingrenzung des regionalen Bezuges an, dennoch haben sich viele Nutzerinnen und Nutzer stattdessen oder zusätzlich dazu entschieden, die entsprechende Region als Suchbegriff zu verwenden.

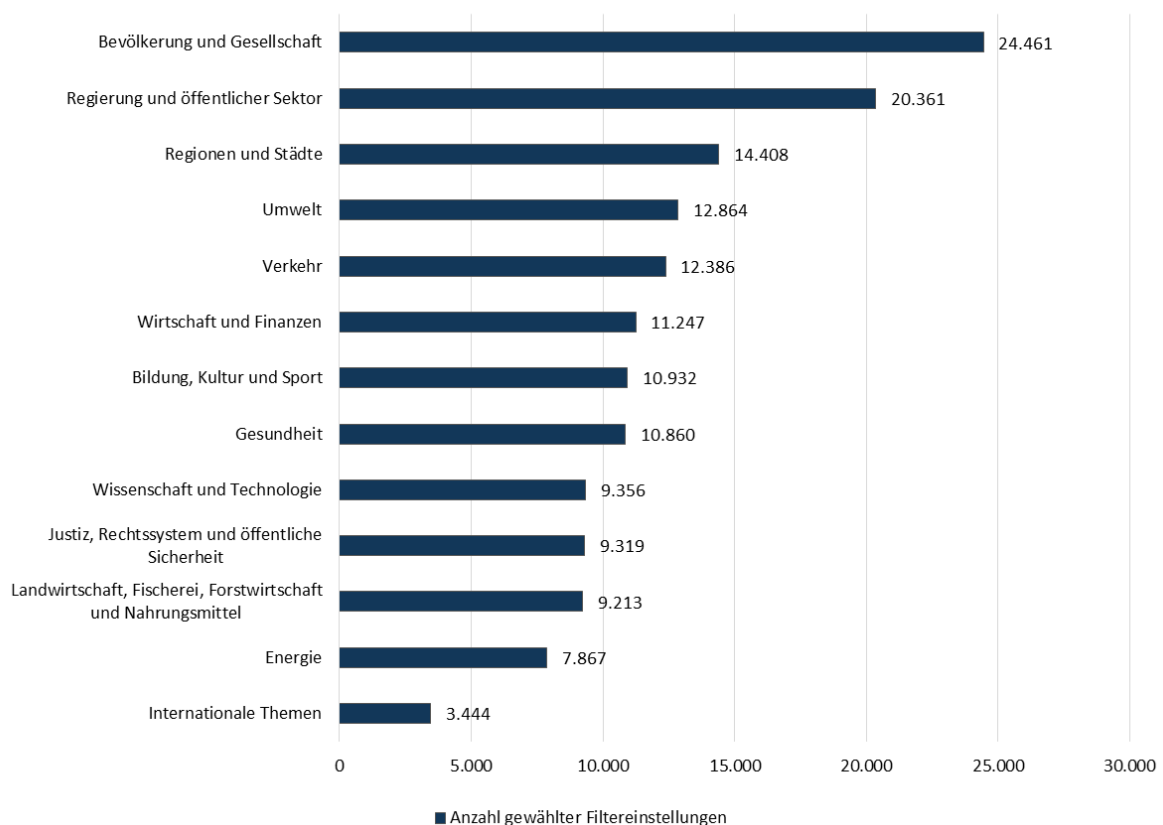
Neben den Suchbegriffen können auch die gewählten Filtereinstellungen Rückschlüsse auf die Interessen der Nutzerinnen und Nutzer der Seite ermöglichen. Govdata.de bietet derzeit die folgenden Filtereinstellungen zur Verfeinerung der Suche an: Kategorie, Datenbereitstellende Behörde, Schlagwörter, Dateiformat, Offenheit der Lizenz, Art der Lizenz sowie Raumbezug (Kartensuche) und einen Zeitbezug. Über die erweiterte Suche ist zusätzlich eine Filterung nach datenveröffentlichender

Stelle, datenverantwortlicher Stelle, Beschreibung der Daten und Typ der gesuchten Information (Daten, Informationen, Blogbeiträge) möglich.⁴⁴

Die Seite ermöglicht eine kombinierte Anwendung der Filter und eine Mehrfachauswahl innerhalb der einzelnen Filterkategorien. Es kann also bspw. nach Daten gesucht werden, die entweder der Kategorie „Verkehr“ oder der Kategorie „Energie“ zugeordnet wurden. Im Folgenden werden die häufig gewählten Filtereinstellungen näher betrachtet. Hierbei ist jedoch zu beachten, dass die Anzahl der vorhandenen Datensätze einen Einfluss auf die gewählten Filter hat, da die entsprechenden Filtereinstellungen nach der Anzahl von passenden Datensätzen angeordnet sind. Somit können Datenkategorien, welche viele Datensätze enthalten, einfacher gefunden werden, weil die entsprechenden Filter an sichtbarer Stelle platziert sind.

Die bereitstellenden Plattformen können den Datensätzen sowohl eine (oder mehrere) von 13 festgelegten Kategorien zur inhaltlichen Orientierung zuordnen, als auch die Datensätze mit selbstgewählten Schlagwörtern versehen. Im gesamten Zeitraum wurde bei 129.775 Suchanfragen eine Kategorie angegeben; das entspricht 51 Prozent aller Suchanfragen. Abbildung 19 zeigt die in den Suchen am häufigsten verwendeten Filterungen nach Kategorien.

» **Abbildung 19: Anzahl der Requests mit gewählter Filterung nach Kategorie**



44 Stand 08. Juli 2021.

Die Abbildung zeigt, dass am häufigsten nach Informationen aus der Kategorie Bevölkerung und Gesellschaft und am zweithäufigsten nach Informationen aus der Kategorie Regierung und öffentlicher Sektor gesucht wurde.

Tabelle 3 zeigt die in den Suchen am häufigsten verwendeten Schlagwörter. Dazu ist anzumerken, dass zwar grundsätzlich jedes beliebige Wort als Schlagwort innerhalb der Suche verwendet werden kann, die Suchfunktion jedoch so gestaltet ist, dass in der einfachen Suche eine Reihe von häufig vergebenen Schlagwörtern als Auswahlmenü angeboten wird. Eine freie Eingabe von Schlagwörtern ist nur über die „erweiterte Suche“ möglich. Es ist davon auszugehen, dass die vorgeschlagenen Schlagwörter häufiger gewählt werden. Insgesamt wurde in 65.614 Fällen eine Filterung nach Schlagwörtern vorgenommen (26 Prozent aller Such-Requests).

» **Tabelle 3: Am häufigsten verwendete Schlagworte bei Nutzung der Suchfunktion**

	Verwendete Schlagworte	Verwendungshäufigkeit
1	inspireidentifiziert	5.274
2	bauleitplanung	3.632
3	opendata	3.350
4	bebauungsplan	3.265
5	bodennutzung	2.431
6	deutsche-bucht	1.835
7	umwelt-und-klima	1.378
8	easygsh-db	1.368
9	raumbezogene-information	1.367
10	mfund-projekt-easygsh-db	1.345
11	meer	1.286
12	bauleitplan	1.244
13	deutschland	1.217
14	geoinformation	1.153
15	deutschland-insgesamt	952

Die Schlagwörter „opendata“, „Bebauungsplan“, „Umwelt“, „inspireidentifiziert“ und „bodennutzung“ finden sich derzeit auf der Website unter den vorgeschlagenen Schlagwörtern in der allgemeinen

Suche.⁴⁵ Die vorgeschlagenen Schlagwörter werden anhand der weiteren gesetzten Filter angepasst. Wird bspw. das Schlagwort „inspireidentifiziert“ ausgewählt, so werden als weitere mögliche Schlagwörter u. a. die Begriffe „bebauungsplan“, „bebauungsplanung“ und „bauleitplan“ vorgeschlagen.

Insgesamt legen die häufig verwendeten Schlagwörter bei der Suche ein verbreitetes Interesse unter den Nutzerinnen und Nutzern an Geodaten nahe, welche der INSPIRE-Richtlinie entsprechend gestaltet sind. Die häufige Verwendung der Begriffe „deutsche bucht“, „meer“ und „easygsh-db“ weisen auf ein Interesse an Gewässerdaten hin, wobei in diesen Fällen eine verdeckte Bot-Aktivität nicht ausgeschlossen werden konnte. Bei easygsh-db handelt es sich um eine Datenquelle, welche Daten zur typographischen Vermessung von Gewässersohlen bereitstellt.

Tabelle 4 zeigt die zehn am häufigsten ausgewählten datenbereitstellende Plattformen zur Filterung der Daten. Am häufigsten wurden demnach nach Daten des Transparenzportals Hamburg und des Open-Data-Portals des Landes Nordrhein-Westfalen gefiltert. Beide Plattformen finden sich auch unter den Plattformen die am meisten Daten auf GovData.de bereitstellen.⁴⁶

» **Tabelle 4: Am häufigsten als Filter verwendete datenbereitstellende Plattformen**

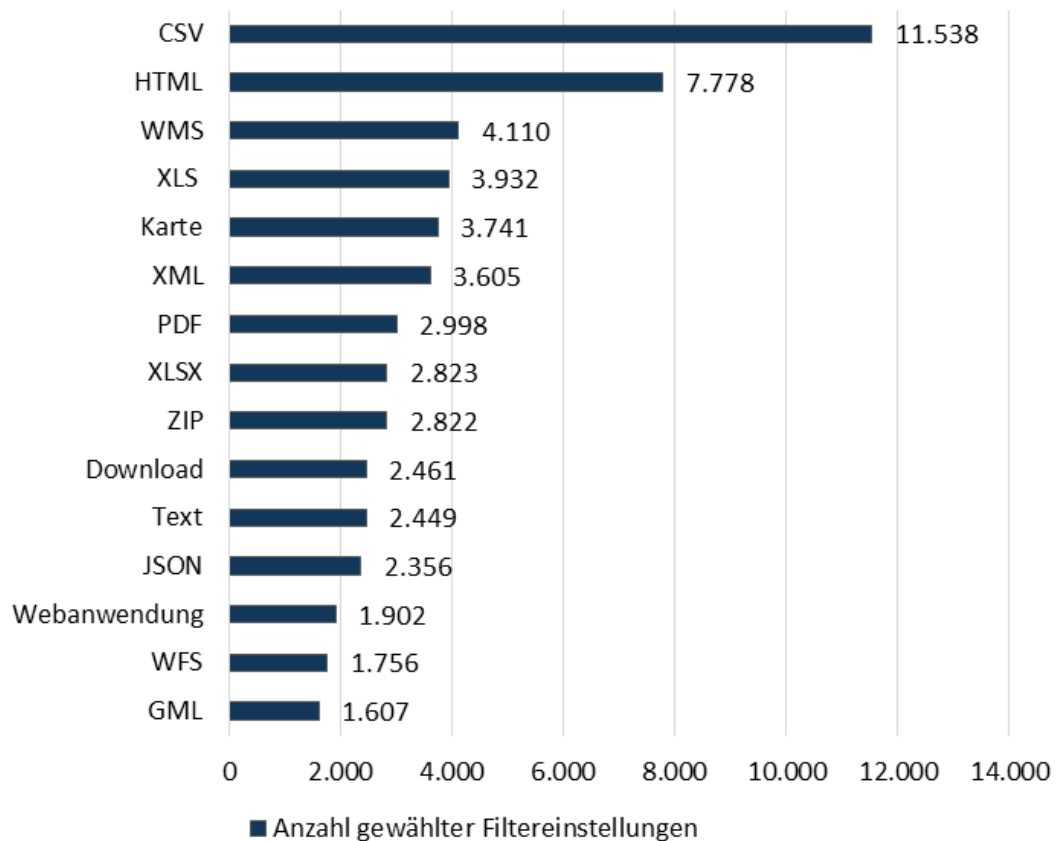
Datenbereitstellende Plattformen	Verwendungshäufigkeit
Transparenzportal Hamburg	3.920
Open.NRW	3.671
OpenGovernment-Data-Portal Rheinland-Pfalz	3.648
GDI-DE	3.562
Statistisches Bundesamt	3.269
Berlin Open Data	3.119
Freistaat Sachsen	2.127
OpenData.HRO	2.049
Bundesministerium des Innern	1.741
Land Schleswig-Holstein	1.685

45 Stand 08. Juli 2021.

46 Vgl. [Kapitel 3.1.3](#)

Abbildung 20 zeigt die am häufigsten gewählten Datenformate als Suchfilter. Insgesamt wurde in 56.992 Fällen eine Filterung nach Datenformat vorgenommen, das entspricht 23 Prozent aller Suchen.

» **Abbildung 20: Am häufigsten als Filter verwendetes Datenformat in der Suchfunktion**



Dementsprechend war die mit großem Abstand am häufigsten gewählte Formatart das CSV-Format, gefolgt von HTML-Dokumenten und dem WMS-Kartenformat. Bei den meisten auf der Seite bereitgestellten Datensätzen handelt es sich um CSV-Dateien.⁴⁷

Tabelle 5 zeigt die am häufigsten als Filterkriterium genutzten Lizenzbezeichnungen. Insgesamt wurden bei (25.579 Suchanfragen) eine oder mehrere Datenlizenzen als Filter ausgewählt, dies entspricht zehn Prozent aller Suchanfragen im betrachteten Zeitraum.

47 Siehe auch [Kapitel 3.1.2.](#)

» **Tabelle 5: Am häufigsten zur Filterung gewählte Lizenzen**

Bezeichnung der Lizenz	Verwendungshäufigkeit
Datenlizenz Deutschland Namensnennung 2.0	6.199
Creative Commons Namensnennung – 4.0 International (CC BY 4.0)	3.143
Andere geschlossene Lizenz	2.942
Creative Commons Namensnennung 3.0 Deutschland (CC BY 3.0 DE)	2.302
Creative Commons Namensnennung (CC-BY)	2.177
Datenlizenz Deutschland – Zero – Version 2.0	1.783
Nutzungsbestimmungen für die Bereitstellung von Geodaten des Bundes	1.048
Creative Commons CC Zero License (cc-zero)	980
Lizenz: Amtliches Werk, lizenzfrei nach §5 Abs. 1 UrhG	830
Andere offene Lizenz	785

Die „Datenlizenz Deutschland Namensnennung 2.0“ und die „Creative Commons Namensnennung – 4.0 International“ sind die am häufigsten gewählten Lizenzinstellungen und werden derzeit auch an oberster Stelle im entsprechenden Filtermenü angezeigt.⁴⁸ Beide Lizenztypen gewähren umfangreiche Nachnutzungsrechte. Etwas überraschend ist die häufig gewählte Einstellung „Andere geschlossene Lizenz“. Unter diese Sammelkategorie für Datensätze mit sonstigen Lizenztypen fallen auf GovData.de u. a. die easygsh-db Datensätze, deren Bezeichnung auch unter den am häufigsten verwendeten Schlagwörter vorkommen. Dies könnte die häufige Filterung nach dieser Lizenz erklären.

Weiterhin ist auf der Seite die Filterung nach einem Zeit- und einem Ortsbezug möglich. Im betrachteten Zeitraum wurde die angebotene Funktion der Kartensuche (Ortsbezug) 10.410 Mal angewendet, dies entspricht vier Prozent aller Suchanfragen. Eine Filterung der Suchergebnisse nach Datum wurde nur bei 795 Suchanfragen angewendet (unter einem Prozent aller Suchen).

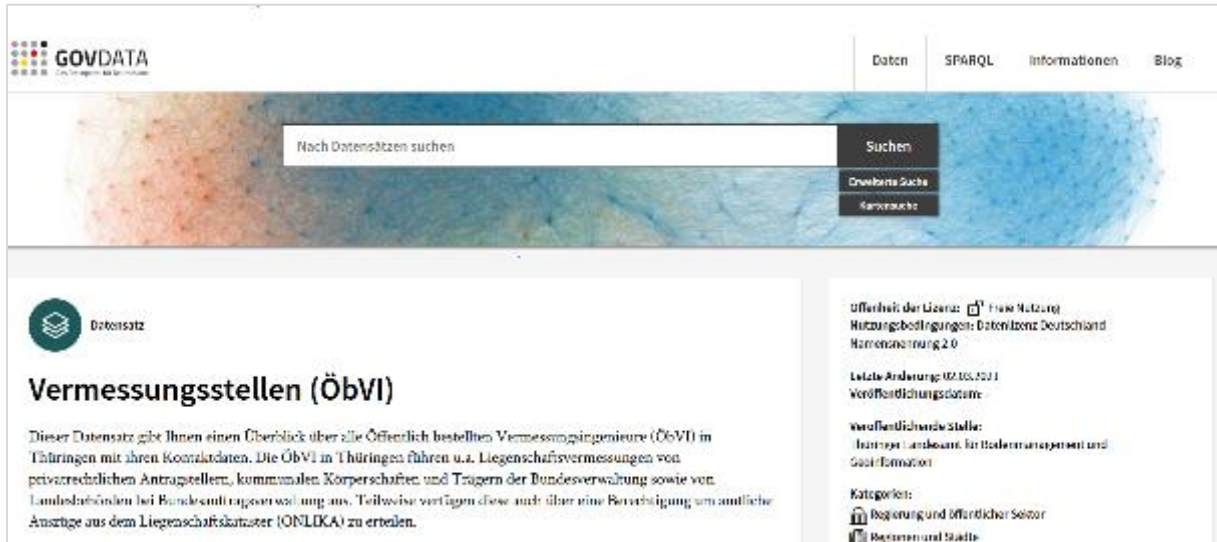
3.3.3 Aufgerufene Detailseiten

Von zentralem Interesse für die Abschätzung der Nachfrage nach den auf der Website enthaltenen Datensätzen ist die Anzahl der Aufrufe der jeweiligen Detailseiten zu den verschiedenen Datensätzen. Die Detailseiten bilden einen zentralen Bestandteil der Seite GovData.de, da sie die Informationen zu

⁴⁸ Stand 08. Juli 2021.

den verzeichneten Datensätzen enthalten. Abbildung 21 zeigt einen beispielhaften Ausschnitt einer Detailseite.

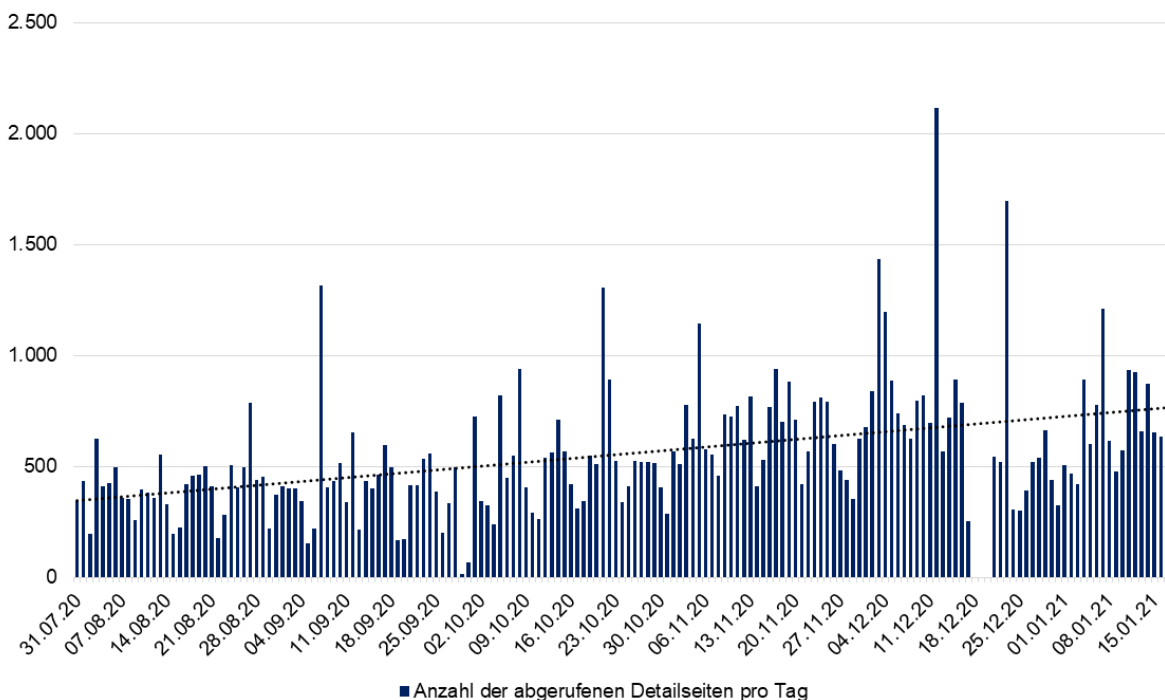
» **Abbildung 21: Beispiel einer Detailseite auf GovData.de**



Der zugehörige Request zur abgebildeten Detailseite lautet „GET /web/guest/suchen/-/details/vermessungsstellen-obvi HTTP/1.1“.

Im gesamten Zeitraum wurden 93.775 Detailseitenabrufe von Webbrowsern verzeichnet. Abbildung 22 zeigt die Anzahl der Abfragen aller Detailseiten pro Tag.

» **Abbildung 22: Abgerufene Detailseiten pro Tag**



Im betrachteten Zeitverlauf ist eine leichte Zunahme an Abfragen von Detailseiten im Zeitverlauf erkennbar.

Tabelle 6 zeigt die Namen der Datensätze zu den jeweils am häufigsten abgerufenen Detailseiten.

» **Tabelle 6: Bezeichnung der Datensätze zu denen am häufigsten Detailseiten aufgerufen wurden**

Name des Dokuments der jeweiligen Detailseite	Abrufe im betrachteten Zeitraum
Anschriftenverzeichnis des Bundes	2.056
Abkürzungsverzeichnis des Bundes	1.460
GovData Metadatenkatalog	895
Wichtige Adressen Moers	846
Bebauungspläne der Stadt Leipzig, Übersicht der Geltungsbereiche	717
schutzgebiete-deutschlands-naturschutz	493
3d-stadtmodell-2	455
preussische-kartenaufnahme-1-25-000-uraufnahme- urmesstischblätter-1843-1879	376
wetterdaten-jahreszahlen	351

Mit Abstand am häufigsten wurden dementsprechend die Detailseiten zum Anschriftenverzeichnis des Bundes und zum Abkürzungsverzeichnis des Bundes angefragt. Beide Datensätze werden durch das Bundesverwaltungsamt zur Verfügung gestellt und auf der Seite service.bund.de veröffentlicht. Auf den jeweiligen Detailseiten auf GovData.de besteht für die Nutzerinnen und Nutzer auch die Möglichkeit, einen Kommentar zum jeweiligen Datensatz zu hinterlassen. Diese Möglichkeit wurde auf den am häufigsten angefragten Detailseiten genutzt, um Fragen zum Datensatz zu stellen oder kritische Anmerkungen zur Datenqualität zu hinterlassen. Diese wurden durch die Redaktion des Portals service.bund.de teilweise beantwortet. Am dritthäufigsten wurde die Detailseite zum Metadatenkatalog der Seite aufgerufen. Die dort beschriebenen Informationen sind für das grundsätzliche Verständnis der Beschaffenheit der auf GovData.de vorhandenen Inhalte relevant.

3.3.3.1 Kombinierte Auswertung der Detailseiten anhand der Logfiles und Metadaten

Da im Rahmen des Projekts auch die Metadaten zu den jeweiligen Datensätzen vorlagen, ist durch eine kombinierte Auswertung von Metadaten und Logfiles auch eine Betrachtung der Nachfrage von Detailseiten zu Datensätzen anhand der verschiedenen zugeordneten Metadaten möglich. Hierzu

wurden den in den Logfiles aufgefundenen Abrufen von Detailseiten anhand des Titels der beschriebenen Datensätze die zugehörigen Metadaten zugeordnet. Hierbei muss jedoch einschränkend angemerkt werden, dass zu knapp 3.600 der abgerufenen Datensätze keine Metadaten aufgefunden werden konnten, da die entsprechenden Metadaten in der Zwischenzeit angepasst oder gelöscht wurden.

Um eine bessere Orientierung bezüglich des Verhältnisses der abgerufenen Datensätze zu den vorhandenen Datensätzen zu ermöglichen, wird in den folgenden Darstellungen auch immer die Gesamtzahl an Datensätzen angegeben, auf die die verschiedenen Kriterien zutreffen, und zu denen zu Beginn und Ende des Erhebungszeitraums auf GovData.de Detailseiten vorlagen ⁴⁹. Die Anzahl an Detailseiten zu Beginn und Ende der Betrachtung ist jedoch nur als Momentaufnahme zu verstehen. So wurden beispielsweise die Detailseiten zu den Datensätzen des Transparenzportals Hamburg zum Zeitpunkt des Beginns der Untersuchungsperiode (30. Juli 2020) zeitweise auf GovData.de entfernt, da es technische Probleme auf Seiten des Transparenzportals gab.

Tabelle 7 zeigt die Abrufe von Detailseiten im gesamten Zeitraum, aufgeschlüsselt nach der Kategorienzuordnung der beschriebenen Datensätze.

» **Tabelle 7: Abrufe von Detailseiten nach vergebener Kategorie**

Zugeordnete Kategorie	Abrufe	Anzahl vorhandene Datensätze Beginn (30. Juli 2020)	Anzahl vorhandene Datensätze Ende (18. Januar 2021)
Regierung und öffentlicher Sektor	26.351	11.230	11.751
Umwelt	20.583	9.108	12.528
Regionen und Städte	17.958	5.608	8.640
Wirtschaft und Finanzen	13.671	7.785	8.242
Verkehr	13.379	3.668	5.192
Bevölkerung und Gesellschaft	13.320	7.439	7.107
Bildung, Kultur und Sport	10.623	2.662	3.615
Wissenschaft und Technologie	9.283	4.247	4.284
Gesundheit	8.367	1.170	1.324

⁴⁹ Die dargestellten Zahlen zu den vorhandenen Metadaten beruhen auf den zu zwei Zeitpunkten übermittelten Metadaten. Die Angaben zum Beginn basieren auf einer Datenlieferung vom 21. August 2020 und die Datensätze zum Ende auf einer Datenlieferung vom 01. Februar 2021.

Landwirtschaft, Fischerei, Forstwirtschaft und Nahrungsmittel	7.941	6.273	6.948
Justiz, Rechtssystem und öffentliche Sicherheit	7.235	4.143	4.530
Energie	4.467	469	4.228
Internationale Themen	911	101	111

Insbesondere aus den Kategorien „Regierung und öffentlicher Sektor“ und „Umwelt“ wurden besonders viele Detailseiten zu zugeordneten Datensätzen aufgerufen. Im Verhältnis zur Anzahl der Datensätze ist besonders auffällig, dass zugehörige Detailseiten in den Kategorien „Gesundheit“ und „Internationale Themen“ häufig besucht wurden, was auf ein besonderes Interesse an Daten in diesen Bereichen hinweist.

Tabelle 8 zeigt die Schlagwörter, welche den Datensätzen zu den abgerufenen Detailseiten am häufigsten zugeordnet wurden.

» **Tabelle 8: Schlagwörter zu denen am häufigsten Detailseiten abgerufen wurden**

Zugeordnete Schlagwörter	Abrufe	Anzahl vorhandene Datensätze Beginn (30.07.2020)	Anzahl vorhandene Datensätze Ende (18.01.2021)
Opendata	13.647	5.334	6.806
Inspireidentifiziert	9.397	3.161	3.691
Bebauungsplan	6.915	2.444	5.878
Bodennutzung	6.820	282	3.815
Berlin	6.369	2.890	1.969
Raumbezogene-Information	5.076	25	3.080
Bauleitplanung	4.876	53	3.499
Jahr	4.724	2.094	2.103
Deutschland	4.637	1.545	1.637
Hamburg	3.826	1.191	1.223

Die Betrachtung der Schlagwörter ergibt, dass häufig Detailseiten zu Datensätzen mit lokalem Bezug zu Hamburg und Berlin aufgerufen wurden. Des Weiteren weisen die themenspezifischen Schlagwörter auf ein erhöhtes Interesse an den Bereichen Bebauungsplanung und Geodaten hin.

Tabelle 9 zeigt die datenbereitstellenden Plattformen, zu deren Datensätze die meisten Detailseiten abgerufen wurden.

» **Tabelle 9: Abgerufene Detailseite nach datenbereitstellenden Plattformen**

Datenbereitstellende Plattformen	Abrufe	Anzahl vorhandene Datensätze Beginn (30. Juli 2020)	Anzahl vorhandene Datensätze Ende (18. Januar 2021)
Open.NRW	12.141	3.173	3.299
GDI-DE	10.183	4.776	6.593
Transparenzportal Hamburg	7.449	1	8.641
Open Government-Data-Portal Rheinland-Pfalz	6.349	4.487	4.521
Freistaat Sachsen	6.261	2.762	2.111
Berlin OpenData	5.716	2.269	1.199
Statistisches Bundesamt	5.364	2.178	2.093
Bundesministerium des Innern, für Bau und Heimat	4.297	584	714
Land-Schleswig-Holstein	4.185	8.323	5.043
Bundesverwaltungsamt	3.516	2	2

Am häufigsten wurden Detailseiten zu Datensätzen der Plattformen Open.NRW und GDI-DE abgerufen. An dritter Stelle folgt das Transparenzportal Hamburg. Wie schon beschrieben, war dieses zu Beginn des Untersuchungszeitraums von einem technischen Problem betroffen, wodurch zwischenzeitlich keine Daten bereitgestellt wurden. Das Problem bestand jedoch nur kurzzeitig. Das Bundesverwaltungsamt stellt, wie bereits erwähnt, die beiden am häufigsten abgerufenen Datensätze bereit, was die hohe Anzahl an Abrufen bei nur zwei Datensätzen erklärt.

Tabelle 10 zeigt die am häufigsten zugeordneten Datensatz-Formate zu den abgerufenen Detailseiten.

» **Tabelle 10: Abgerufene Detailseiten nach Datensatz-Format**

Format	Abrufe	Anzahl vorhandene Datensätze Beginn (30. Juli 2020)	Datensätze vorhandene Datensätze Ende (18. Januar 2021)
CSV	20.771	7.714	13.089
HTML	7.745	3.836	2.848
PDF	5.783	6.584	5.022
XML	5.009	1.510	1.608
JSON	4.970	322	274
XLSX	4.342	1.132	1.643
XLS	3.642	997	1.203

Mit Abstand am häufigsten wurden Detailseiten zu Datensätzen im CSV-Format abgerufen, welches auch während des gesamten betrachteten Zeitraumes das am häufigsten vergebene Datenformat darstellt. Im Verhältnis zu den bereitgestellten Datensätzen wurden jedoch Detailseiten zu Datensätzen in den anderen der sieben am häufigsten vergebenen Formate häufiger abgerufen, mit Ausnahme von Daten im PDF-Format. Besonders häufig im Verhältnis zu den vorhandenen Datensätzen wurden Detailseiten aufgerufen, die eine als JSON-Datei vorhandenen Datensatz beschreiben.

3.3.4 Sonstige häufig aufgerufene Teilbereiche der Seite

Neben den bisher betrachteten Detailseiten zu den einzelnen Datensätzen gibt es auf GovData.de einen Blog mit Artikeln zu neuen Entwicklungen im Bereich offene Verwaltungsdaten und einen Seitenbereich mit verschiedenen Informationen zur Nutzung der Seite. Diese gesonderten Funktionen werden im Folgenden kurz dargestellt, wobei wie bei der Betrachtung der Detailseiten und der Suchen nur Abfragen berücksichtigt wurden, welche einem klassischen Webbrowser zugeordnet wurden. Bot-Aktivitäten und Abfragen aus anderen Systemen wurden auch hier, wie bei der Betrachtung der Detail- und Suchergebnisseiten, nicht berücksichtigt.

Tabelle 11 zeigt die Titel der am häufigsten abgerufenen Informationsseiten.

» **Tabelle 11: Häufig abgerufene Informationsseiten**

Seitentitel	Link	Abrufe
Open Government	https://www.govdata.de/web/guest/open-government	6.249
Datenlizenz Deutschland	https://www.govdata.de/web/guest/lizenzen	2.706
Fragen und Antworten zu GovData.de	https://www.govdata.de/web/guest/faq	2.169
Das Portal	https://www.govdata.de/web/guest/hilfe	1.644
Metadatenstruktur für Daten in Deutschland	https://www.govdata.de/web/guest/metadaten-schema	1.624
Die Bereitstellung von Daten über GovData.de	https://www.govdata.de/web/guest/datenbereitsteller	1.472
OGD-DACHLI Kooperation	https://www.govdata.de/web/guest/ogd-dachli	687
Termine	https://www.govdata.de/web/guest/termine	663

Auf der Informationsseite „Open Government“ wird eine umfangreiche Sammlung an Materialien zum Themengebiet Open Government von Gesetzestexten bis zu wissenschaftlichen Studien zur Verfügung gestellt. Auf der Seite „Datenlizenz Deutschland“ werden die Feinheiten besagter Datenlizenzen näher dargestellt. In den Bereichen „Fragen und Antworten zu GovData“ und „Das Portal“ wird die Gestaltung des Portals GovData.de beschrieben. Die Seite zur Bereitstellung von Daten über GovData.de richtet sich primär an Institutionen, welche Ihre Daten über die Plattform bereitstellen möchten und die Informationsseite „OGD-DACHLI Kooperation“ schildert eine Kooperation zwischen Deutschland, Österreich, der Schweiz und Liechtenstein zum Informationsaustausch im Bereich offene Verwaltungsdaten. Im Bereich „Termine“ wurden im betrachteten Zeitraum aufgrund der Covid-19-Pandemie nur wenige Informationen bereitgestellt.

Im Gegensatz zu diesen recht allgemein gehaltenen Informationen im Bereich der Infoseiten wird im Bereich der Blogbeiträge auf Neuigkeiten im Bereich offene Verwaltungsdaten eingegangen. Tabelle 12 zeigt die am häufigsten im betrachteten Zeitraum abgerufenen Blogbeiträge. Auch hierbei ist zu beachten, dass nicht alle Blogbeiträge den gesamten Zeitraum über auf der Website vorhanden waren, was die Darstellung zugunsten der älteren Blogbeiträge verzerrt.

» **Tabelle 12: Häufig abgerufene Blogbeiträge**

Titel des Blogbeitrages	Link zum Blogbeitrag	Anzahl der Abrufe
DCAT-AP.de - neue Version verfügbar	https://www.govdata.de/web/guest/neues/-/blogs/dcat-ap-de-neue-version-verfuegbar	750
Gastbeitrag: Musterdatenkatalog für Open Data: Welche Daten veröffentlichen Kommunen?	https://www.govdata.de/web/guest/neues/-/blogs/gastbeitrag-musterdatenkatalog-fur-open-data-welche-daten-veroeffentlichen-kommunen-	731
Covid-19 - Zahlen zur Pandemie	https://www.govdata.de/web/guest/neues/-/blogs/covid-19	559
#WirVsVirusHack - Linksammlung aus dem Hackathon	https://www.govdata.de/web/guest/neues/-/blogs/-wirsvirushack-linksammlung-aus-dem-hackathon	524
Onlineumfrage zur Nutzung und Bewertung von GovData.de	https://www.govdata.de/web/guest/neues/-/blogs/onlineumfrage-zur-nutzung-und-bewertung-von-govdata-de	501
Auswertung Suchbegriffe	https://www.govdata.de/web/guest/neues/-/blogs/auswertung-suchbegriffe	475
Online-Konsultation zum 2. Nationalen Aktionsplan im Rahmen der Open Government Partnership gestartet	https://www.govdata.de/web/guest/neues/-/blogs/online-konsultation-zum-2-nationalen-aktionsplan-im-rahmen-der-open-government-partnership-ogp-gestartet	472
DCAT-AP.de vom IT-Planungsrat als formaler Metadatenstandard für offene Verwaltungsdaten bestätigt	https://www.govdata.de/web/guest/neues/-/blogs/dcat-ap-de-vom-it-planungsrat-als-formaler-metadatenstandard-fur-offene-verwaltungsdaten-bestaetigt	447
Video veröffentlicht: Was ist DCAT-AP.de und wofür braucht man das?	https://www.govdata.de/web/guest/neues/-/blogs/video-veroeffentlicht-was-ist-dcat-ap-de-und-wofur-braucht-man-das-	444
DCAT-AP.de in der Version 1.0.2 veröffentlicht - Entwicklung zur Version 1.1 gestartet!	https://www.govdata.de/web/guest/neues/-/blogs/dcat-ap-de-in-der-version-1-0-2-veroeffentlicht	442

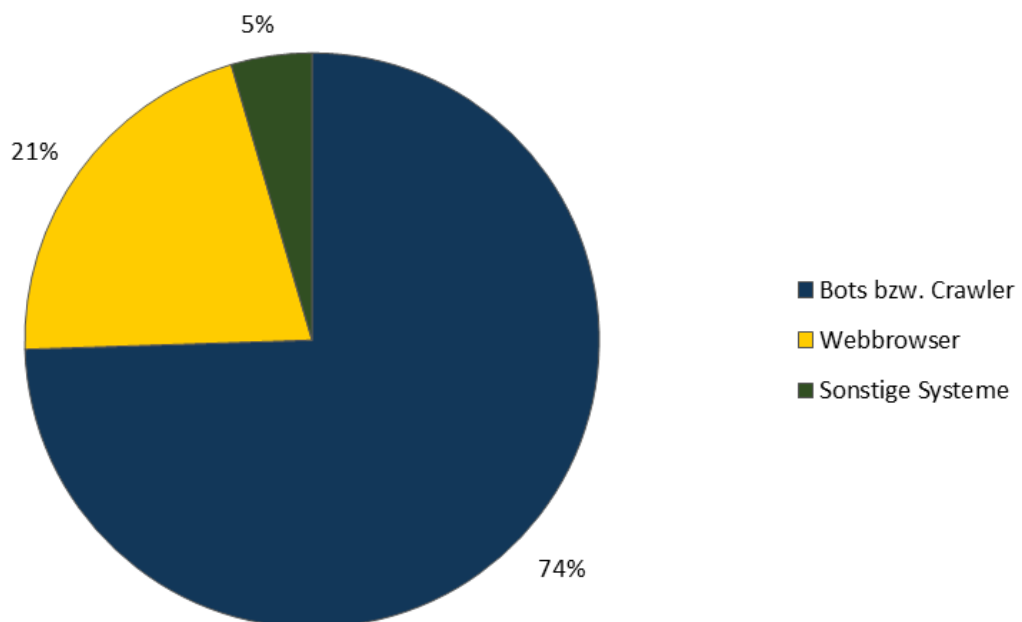
Insgesamt beschäftigen sich also vier der zehn am häufigsten abgerufenen Blogbeiträgen mit dem DCAT-AP.de Standard, was auf einen entsprechenden Informationsbedarf zum Metadatenstandard hinweist. Zudem zeigt sich auch anhand der häufig abgerufenen Blogbeiträgen ein hohes Interesse am Themenbereich Corona.

3.3.5 Abfragen über die API

GovData.de bietet auch eine API zur Abfrage der Metadaten an, welche es den Nutzerinnen und Nutzern ermöglicht, direkt auf die Metadaten der Datensätze zuzugreifen. Hierbei können die Metadaten zu konkreten Datensätzen anhand der Bezeichnung des Datensatzes ausgewählt werden. Alternativ besteht die Möglichkeit, vergleichbar mit den Filteroptionen auf der Suchergebnisseite, Metadaten zu Datensätzen abzurufen, die eine gewählte Kombination von Metadaten aufweisen. Auch zu den API-Abrufen wurden dem FÖV durch GovData.de Logfiles übermittelt.

Da die API auf den automatisierten Abruf von Daten ausgelegt ist, wurden in der Auswertung auch Bot-Abfragen und Abfragen durch andere Systeme als Webbrowser berücksichtigt. Abbildung 23 zeigt die Zusammensetzung der Abfrage über die API nach Webbrowsern, identifizierten Bots und sonstigen Systemen. Insgesamt wurden im gesamten Zeitraum über 14 Millionen Abfragen über die API verzeichnet.

» **Abbildung 23: Zusammensetzung der API-Abfragen nach abfragendem System**



Der überwiegende Teil der Abfragen über die API wird demnach durch Bots verursacht.

Von primärem Interesse bei der Betrachtung der API-Abfragen ist die Frage, welche Datensätze auf diesem Wege besonders häufig abgerufen wurden. Tabelle 13 zeigt die Bezeichnungen der Datensätze, zu denen am häufigsten Metadaten anhand der Datensatzbezeichnung abgerufen wurden.

» **Tabelle 13: Am häufigsten in den API-Abfragen verwendete Datensatzbezeichnungen**

Bezeichnung des Datensatzes	Abrufe	Kurzbeschreibung
c31_lebensmittelmarkt_haardtstrasse_2	2.217	WMS Karte Landau (RLP)
brandenburger-strassen-und-volksfeste-2017	2.067	REST-API zu Website über Volksfeste
f1_neuaufstellung_ii	2.067	WMS Karte Landau (RLP)
4-aenderung-beim-weisen-stein-4-aenderung-des-bebbauungsplanes-beim-weisen-stein	1.839	WMS Karte (RLP)
kfz-bestand-kfz-bestand-kfz-arten-grossenkl-kfz-anhanger-mit-amtl-kz-pkw-kfz-je-1000-einw-veran	1.824	CSV Kfz-Daten Sachsen
auslandische-bevolkerung-veranderung-gegenuber-2005-in-kreisebene	1.770	WMS Karte (RLP)
mh3	1.731	WMS Karte (RLP)
g zr-daten-zur-schwarzarbeit-2014-8-abs-1-nr-2-i-v-m-8-abs-1-nr-1e-schwarzarbg-nat-gew	1.710	CSV Statistik Schwarzarbeit
g1	1.704	WMS Karte (RLP)
abkürzungsverzeichnis-des-bundes	1.698	CSV Abkürzungsverzeichnis

Bei vielen der zehn am häufigsten abgerufenen Metadaten nach Datensatzbezeichnung handelt es sich demnach um Kartendaten zu Gemeinden in Rheinland-Pfalz. Auf Platz zehn ist das Abkürzungsverzeichnis des Bundes zu finden, welches auf Platz zwei der am häufigsten abgerufenen Detailseiten auf der Website zu finden ist. Da Bots erhebliche Mengen von Anfragen erzeugen können, kann jedoch keine Aussage darüber gemacht werden, wie viele Personen hinter diesen Anfragen stehen. Es ist durchaus möglich, dass es sich bei dem Abfragen in Tabelle 14 mit Bezug zu Kartendaten in Rheinland-Pfalz überwiegend um Abfragen von Bots einer einzigen Person oder Organisation handelt.

Einen etwas breiteren Eindruck vom Inhalt der Datensätze, zu denen über die API häufig Informationen abgerufen wurden, können die am häufigsten in den Titeln der abgefragten Datensätze verwendeten Wörter geben. Hierzu wurden Füllwörter in den Titeln nicht berücksichtigt z. B. „in“, „mit“ „und“; ebenso wurden Zahlenwerte entfernt, die häufig in den Datensatzbezeichnungen vorkommen.

» **Tabelle 14: Häufig in den abgefragten Datensatzbezeichnungen verwendete Begriffe**

Einzelne Begriffe aus den verwendeten Titeln der Datensätze	Verwendungshäufigkeit
Hamburg	448.785
Schleswig	365.325
Holstein	348.630
Deutschland	264.030
Karte	262.095
Topographische	244.680
Bebauungsplan	212.439
Berlin	208.206
Messergebnisse	183.027
Radioaktivität	180.396

Die häufig verwendeten Begriffe weisen auf ein starkes Interesse an Kartendaten bei den Abfragen über die API hin, sowohl in Bezug auf topografische Karten als auch in Bezug auf Bebauungspläne.

Auch die in den API-Abfragen genutzten Metadaten zur Auswahl von Datensätzen können Anhand der Logfiles analysiert werden. Tabelle 15 zeigt die in den API-Abfragen am häufigsten verwendeten Schlagwörter.

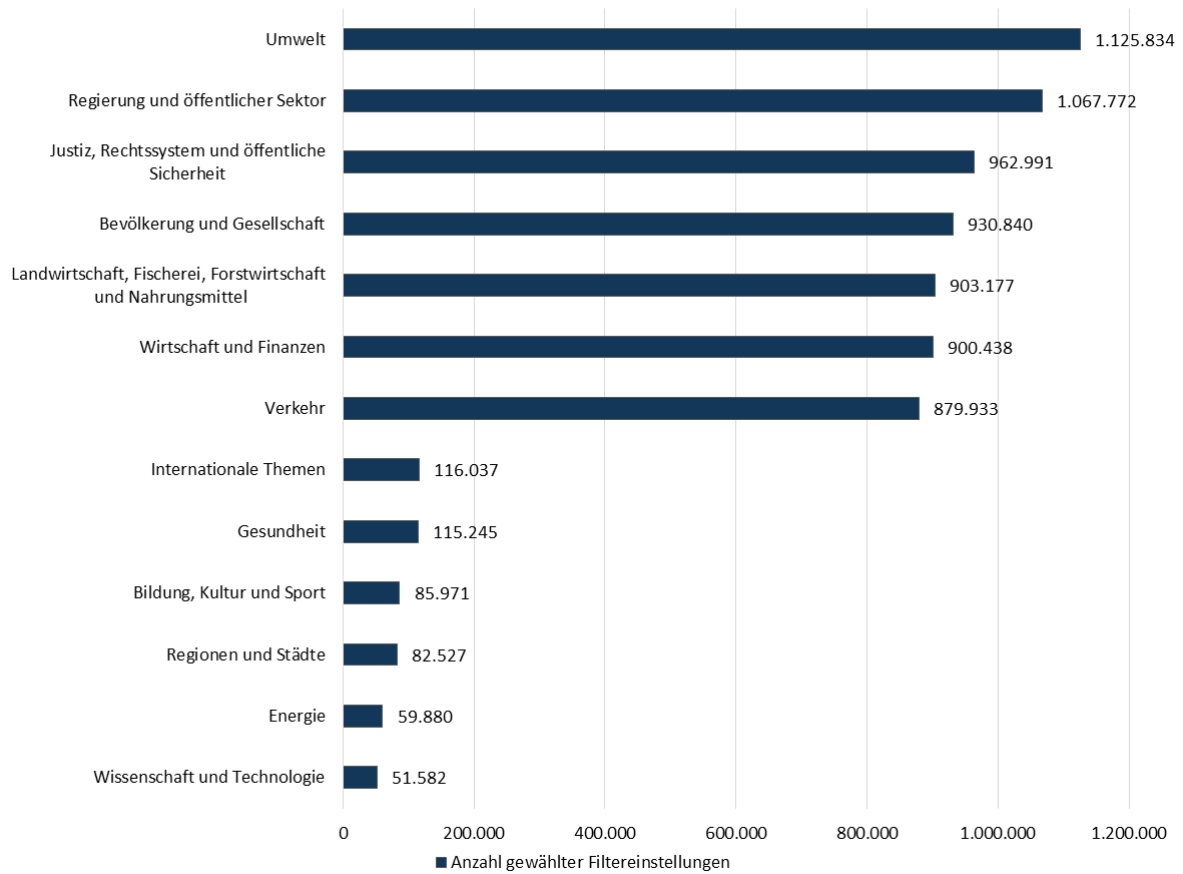
» **Tabelle 15: Am häufigsten in den API-Abfragen verwendete Schlagwörter**

	Verwendete Schlagwörter	Verwendungshäufigkeit
1	Regionalplan	2.081.127
2	Planungsgemeinschaften	1.573.536
3	Regionalplanung	1.364.844
4	Inspireidentifiziert	774.258
5	Denkmal	489.120
6	Planungsregion	417.198
7	IATI	140.901
8	Entwicklungszusammenarbeit	135.678
9	Transparenz	134.583
10	Zusammenarbeit	124.122
11	BMZ	115.824
12	International	114.324
13	Windenergie	89.286
14	Opendata	58.986
15	Kodierung	50.829

Im Gegensatz zur Nutzung der Schlagwörter als Filter im Rahmen der Suchfunktion auf der Seite GovData.de wurde über die API auffällig häufig nach Metadaten zu den Schlagwörtern Regionalplanung bzw. Regionalplan gefiltert.

Abbildung 24 zeigt die Verteilung der Abfragen über die API nach gewählter Kategorie.

» **Abbildung 24: Anzahl der in den API-Abfragen gewählten Kategorie zur Auswahl der Daten**



Über die API wurde am häufigsten die Kategorie „Umwelt“ zur Selektion der Daten angewendet. In dieser Kategorie finden sich viele Geo- bzw. Kartendaten. Ebenfalls auffällig ist die häufige Filterung nach Daten aus den Bereichen „Justiz, Rechtssystem und öffentlicher Sektor“ und „Landwirtschaft, Fischerei, Forstwirtschaft und Nahrungsmittel“.

Tabelle 16 zeigt die zehn am häufigsten gewählten Formate zur Selektion der Daten bei API-Abfragen.

» **Tabelle 16: Am häufigsten in den API-Abfragen ausgewählte Dateiformate**

Formatbezeichnungen	Abfragen
WMS	1.002.165
Webanwendung	950.787
HTML	868.530
Karte	841.491
REST	750.762
Diverse	658.269
XML	199.662
CSV	130.935
PDF	88.881
WFS	44.511

Anhand der gewählten Formate lässt sich ein deutliches Interesse der Nutzerinnen und Nutzer der API an Kartendaten erkennen. Metadaten zu CSV-Dateien sind hingegen im Vergleich zu den gewählten Filtereinstellungen auf der Website weniger gefragt.

3.3.6 Betrachtung der Abfragen durch Bots und sonstige Systeme

Wie eingangs bereits dargestellt, wird die Mehrheit der Abrufe auf GovData.de, soweit ersichtlich, automatisiert erzeugt. Diese Abfragen wurden in der bisherigen Auswertung, außer bei der Betrachtung der API-Abrufe, weitgehend außer Acht gelassen. Abschließend soll daher ein kurzer Überblick über derartige Aktivitäten erfolgen. Dies dient eher dem besseren Verständnis der verschiedenen Nutzungsweisen der Seite als einer tiefergehenden Analyse von abgefragten Inhalten.

Die abrufenden Systeme wurden für die vorliegende Auswertung in drei Kategorien anhand des in den Logfiles enthaltenen User-Agent eingeteilt: in Bots bzw. Crawler, Webbrowser und sonstige Systeme. Tabelle 17 zeigt die fünf User-Agents, welche in der Kategorie Bots bzw. Crawler die meisten Abfragen erzeugt haben.

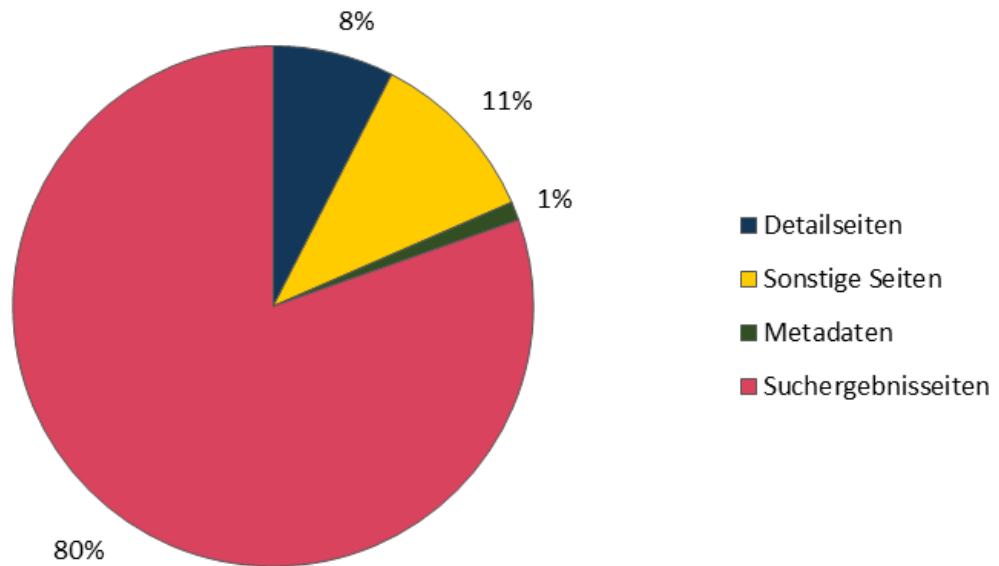
» **Tabelle 17: Bots bzw. Crawler mit den meisten erzeugten Abrufen**

	Bots	Generierte Abrufe
1	Mozilla/5.0 (compatible; bingbot/2.0; +http://www.bing.com/bingbot.htm)	4.985.666
2	Mozilla/5.0 (compatible; SemrushBot/6~bl; +http://www.semrush.com/bot.html)	1.187.031
3	Mozilla/5.0 (Linux; Android 7.0;) AppleWebKit/537.36 (KHTML, like Gecko) Mobile Safari/537.36 (compatible; PetalBot;+https://aspiegel.com/petalbot)	948.620
4	Mozilla/5.0 (compatible; MegaIndex.ru/2.0; +http://megaindex.com/crawler)	947.543
5	Mozilla/5.0 (compatible; SemrushBot/7~bl; +http://www.semrush.com/bot.html)	683.790

Über 50 Prozent der identifizierten Bot-Anfragen wurden durch zwei Bots verursacht; dabei handelte es sich um den Bot der Microsoft-Suchmaschine Bing (Mozilla/5.0 (compatible; bingbot/2.0; +http://www.bing.com/bingbot.htm)) und um einen Bot der Firma Semrush (Mozilla/5.0 (compatible; SemrushBot/6~bl; +http://www.semrush.com/bot.html)). Bei Semrush handelt es sich um ein Marketing-Unternehmen, welches für die Erstellung eines eigenen Such-Tools einen Bot einsetzt. Der Petalbot gehört zu einer Suchmaschine von Huawei. Megaindex ist ebenfalls ein Marketingunternehmen mit Fokus auf Suchmaschinenoptimierung. Insgesamt verursachen die fünf in Tabelle 17 dargestellten Bots 32 Prozent aller festgestellten GET-Requests im betrachteten Zeitraum.

Abbildung 25 zeigt überblicksartig die Zusammensetzung der Bot-Abfragen in Bezug auf Suchergebnisseiten, Detailseiten, Metadaten und sonstige Seiten. In dieser Zusammensetzung wurden die Abfragen durch Bots über die API nicht berücksichtigt. Dementsprechend fasst die Kategorie „Metadaten“ nicht die Abfragen der Metadaten über die API, sondern die direkte Abfragen von Metadaten über die Detailseiten der Website zusammen.

» **Abbildung 25: Zusammensetzung der Bot- bzw. Crawler-Abfragen**



Es werden durch Bots also hauptsächlich Suchergebnisseiten abgefragt, dies entspricht dem gängigen Vorgehen von Crawlern für Suchmaschinen und ähnlichen Suchanwendungen. Diese gehen regelmäßig die verzeichneten Seiten ab, um aktuelle Änderungen zu registrieren und so die jeweilige Suchanwendung aktuell zu halten.

Interessant ist es auch, sich die in den Logfiles registrierten Abfragen von sonstigen Systemen anzusehen. Tabelle 18 zeigt die fünf User-Agents zu den Systemen, welche in dieser Kategorie die meisten Abfragen erzeugt haben. 89 Prozent der durch andere Systeme generierten Abfragen stammen aus der Python Requests Library 2.22.

» **Tabelle 18: Sonstige Systeme mit den meisten erzeugten Abrufen**

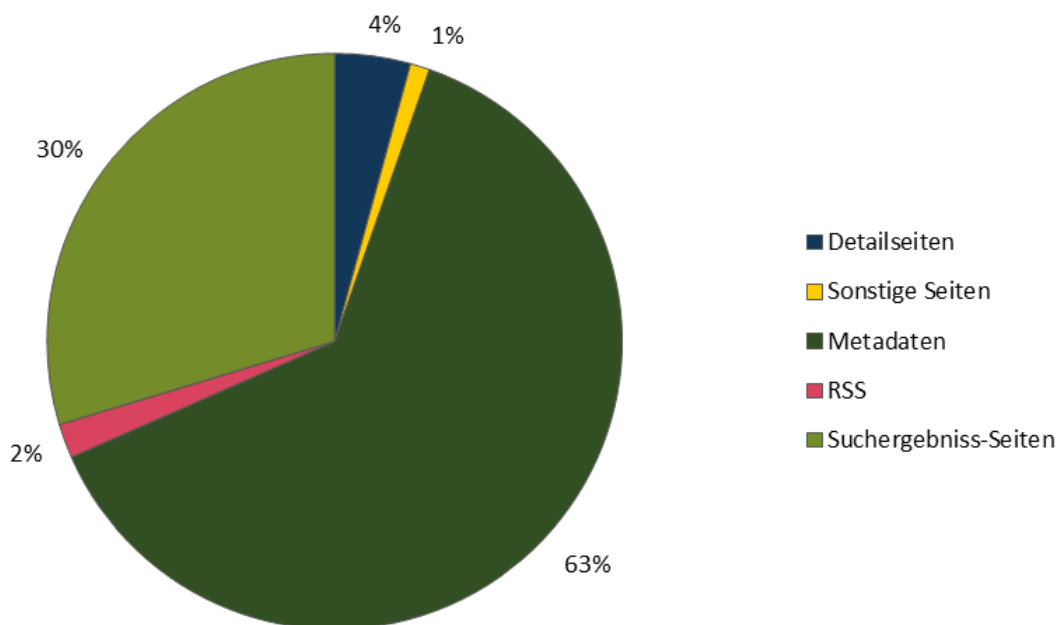
	Sonstiges System	Generierte Abrufe
1	python-requests/2.22.0	4.897.107
2	PHP/5.5	245.232
3	python-requests/2.24.0	163.611
4	Java/1.8.0_202	99.910
5	FME/2019.7.38.19813 libcurl/7.65.2 Schannel zlib/1.2.11 WinIDN libssh2/1.9.0 nghttp2/1.38.0	55.177

Zu den Abfragen aus sonstigen Systemen zeigt Abbildung 26 einen Überblick nach abgerufener Seite. Abweichend zu der Darstellung der Abfragen durch die Bots bzw. Crawler wurden hierbei die Abrufe

des RSS-Feed als ein separater Punkt dargestellt, da diese einen relevanten Anteil der Abfragen ausmachten.

GovData.de bietet einen RSS-Feed an. Dieser kann genutzt werden, um sich regelmäßig über aktuelle Änderungen auf der Seite zu informieren, ohne diese jedes Mal aufrufen zu müssen. Die Abrufe dieses RSS-Feed machten zwei Prozent aller Abrufe der Seite durch sonstige Systeme aus.

» **Abbildung 26: Zusammensetzung der Abfragen durch sonstige Systeme**



Am häufigsten wurden sonstige Systeme jedoch zur Abfrage von Metadaten genutzt. Damit ergibt sich eine eindeutig abweichende Nutzung zu der durch die identifizierten Bots bzw. Crawler. Warum in diesen Fällen die Metadaten direkt von der Website und nicht über die API abgerufen wurden und wie viele Personen hinter diesen Anfragen aus anderen Systemen stehen, ließ sich im Rahmen der vorliegenden Untersuchung nicht feststellen.

3.3.7 Zusammenfassung

Die Auswertung der Logfiles ermöglicht eine detaillierte Betrachtung der abgerufenen Inhalte. Hierzu wurden sowohl Seitenaufrufe auf der Seite GovData.de als auch Abrufe über die API der Seite ausgewertet. Betrachtet wurde jeweils Abrufe im Zeitraum vom 30. Juli 2020 bis zum 18. Januar 2021.

Generell lässt sich bei dieser Betrachtung ein hoher Anteil an Bot-Aktivitäten sowohl bei den Seitenabrufen der Website als auch bei den Abfragen über die API feststellen. Die abfragenden Systeme wurden für die Analyse in die Kategorien identifizierte Bots/Crawler, Webbrowser und sonstige Systeme unterteilt. Die Abfragen durch Bots bzw. Crawler und sonstige Systeme wurden separat betrachtet. Die Betrachtung der Aktivitäten durch diese Systeme ergab für die Bots und Crawler ein häufiges Aufrufen von Suchergebnisseiten, was dem Verhalten von Bots zur Generierung

von Suchmaschinen und Ähnlichem entspricht. Über sonstige Systeme wurden hingegen überwiegend Metadaten zu Datensätzen abgerufen.

Zu den durch Webbrowser ausgelösten Requests wurden im Besonderen die Abrufe der Suchergebnisseiten und der Detailseiten zu den unterschiedlichen Datensätzen näher betrachtet. Sowohl anhand der in den Logfiles festgehaltenen Abrufe von Suchergebnisseiten als auch bei den Detailseiten war eine leichte Zunahme an Abrufen im Zeitverlauf erkennbar.

Die Betrachtung der häufig verwendeten Suchbegriffe weist vor allem auf ein großes Interesse am Themenbereich Covid-19 hin. Weiterhin wurden häufig regionale Einheiten als Suchbegriffe verwendet. Mit Abstand am häufigsten wurde nach Datensätzen im CSV-Format gesucht, welches auch das am häufigsten vergebene Format unter den vorhandenen Datensätzen auf der Website ist. Die verwendeten Schlagwörter weisen auf ein besonderes Interesse an Kartendaten hin, welche entsprechend der INSPIRE-Richtlinie gestaltet sind.

Die am häufigsten im betrachteten Zeitraum aufgerufenen Detailseiten sind die Detailseiten zu den Datensätzen des Abkürzungsverzeichnisses des Bundes und des Anschriftenverzeichnisses des Bundes. Beide Datensätze werden durch das Bundesverwaltungsamt bereitgestellt.

Die kombinierte Auswertung von Logfiles und Metadaten ermöglichte eine Betrachtung der aufgerufenen Detailseiten nach den Metadaten der zugehörigen Datensätze. Unter anderem ergab diese Betrachtung anhand der zugeordneten Schlagwörter, dass besonders häufig Detailseiten zu Datensätzen im Bereich Bebauungsplanung abgerufen wurden.

Weiterhin wurde die Nachfrage in den Teilbereichen der Website „Blogeinträge“ und „Informationsseiten“ näher betrachtet. Die häufig abgerufenen Blog-Beiträge weisen auf einen Informationsbedarf an dem Metadatenstandard DKAT-AP.de und ebenfalls am Themenbereich Covid-19 hin.

Im Gegensatz zu den auf der Website GovDat.de betrachteten Detailseiten und verwendeten Sucheinstellungen wurde bei den Abfragen über die API auffällig häufig nach Daten im WMS-Kartenformat gesucht, Metadaten zu Daten im CSV-Format waren hingegen weniger gefragt. Auch die verwendeten Schlagwörter und Titel der am häufigsten abgerufenen Datensätze weisen auf einen entsprechenden Bedarf nach Informationen zu Kartendaten über die API hin.

3.4 Analyse der Nutzerinnen und Nutzer: Onlineumfrage

Während Logfiles und Matomo-Daten über die durchgeführten Aktionen auf GovData.de Aufschluss geben, liefern sie wenige Informationen über die Eigenschaften der Nutzerinnen und Nutzer offener Verwaltungsdaten. Um mehr darüber zu erfahren, zu welchem Zweck und in welchem Kontext Nutzerinnen und Nutzer auf das Angebot von GovData.de zurückgreifen und wie sie die Plattform bewerten, wurde daher eine standardisierte Onlinebefragung durchgeführt. Der standardisierte Fragebogen wurde in Abstimmung mit dem Bundesministerium des Innern, für Bau und Heimat und der Geschäftsstelle GovData.de erstellt und adressiert folgende Aspekte:

- » Demographische Daten der Nutzerinnen und Nutzer
- » Häufigkeit der Nutzung
- » Zweck und Kontext der Nutzung
- » Erfahrungen mit dem Suchen und Finden von Datensätzen
- » Bewertung des Datenangebots und der Webseite

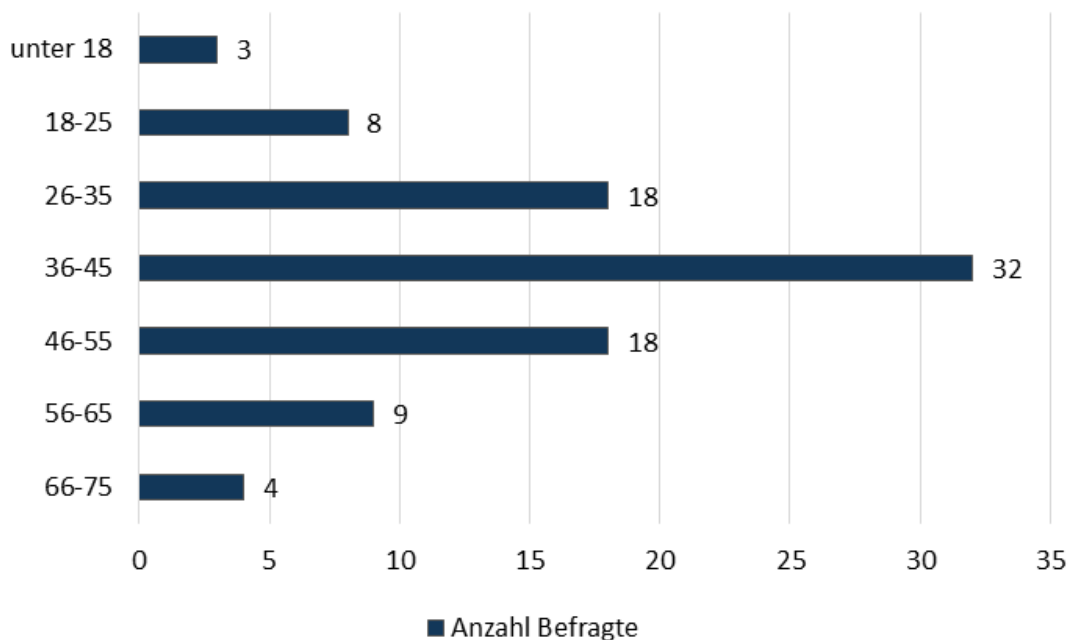
Nach einem internen Pretest wurde die Befragung zwischen dem 12. November 2020 und dem 31. März 2021 auf der Startseite von GovData.de und in einem Blogbeitrag auf GovData.de verlinkt. Zusätzlich wurde von GovData.de ein Tweet mit dem Link zur Onlineumfrage abgesetzt, um auf die Befragung aufmerksam zu machen.

Insgesamt haben 94 Personen den Fragebogen vollständig beantwortet. Hier sei darauf hingewiesen, dass es sich nicht um einen repräsentativen Datensatz handelt und die Ergebnisse somit keine Rückschlüsse auf die Gesamtheit der Nutzerinnen und Nutzer von GovData.de zulassen. Da die Teilnahme an der Befragung aufgrund der Freiwilligkeit auf einer Selbstselektion beruht, sind Personen mit bestimmten Eigenschaften womöglich überrepräsentiert. Vor diesem Hintergrund liefern die Ergebnisse der Datenanalyse ausschließlich Informationen über die Nutzerinnen und Nutzer von GovData.de, die motiviert waren, an der Befragung teilzunehmen.

3.4.1 Demographische Daten der Teilnehmerinnen und Teilnehmer

In Bezug auf die demographischen Daten der Befragten zeigt Abbildung 27, dass vor allem die Altersgruppe zwischen 36 und 45 Jahren in der Stichprobe vertreten war. Ein recht großer Anteil gehörte zudem der Altersgruppe zwischen 26 und 35 Jahren sowie zwischen 46 und 55 Jahren an. Deutlich weniger Teilnehmende waren jünger als 26 oder älter als 55.

» **Abbildung 27: Alter der Befragten**

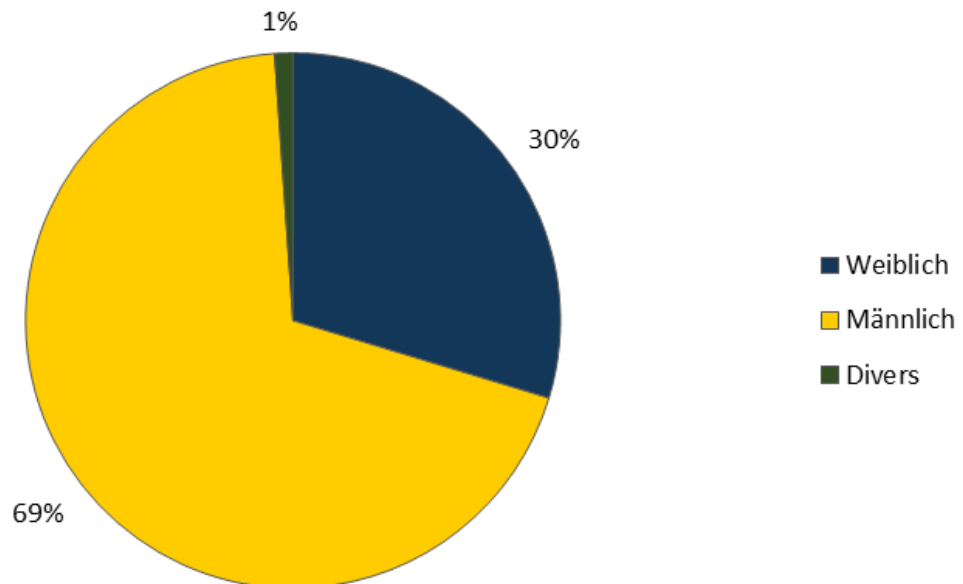


» Anmerkung: Berechnung auf Basis gültiger Werte (N = 92). Kleinere Abweichungen in den Fallzahlen gehen auf fehlende Antworten zurück.

Darüber hinaus fällt auf, dass der Großteil der Befragten männlich war (siehe Abbildung 28). Aufgrund der Limitation statistischer Erhebungsmethoden ist eine Einordnung dieses Ergebnisses schwierig: Obwohl es sich nicht um repräsentative Daten handelt, deutet das unausgeglichene

Geschlechterverhältnis der Befragten zwar darauf hin, dass das GovData.de-Angebot häufiger von Männern als von Frauen in Anspruch genommen wird. Ebenso besteht jedoch die Möglichkeit, dass dieses Ergebnis der Selbstselektion geschuldet ist und Männer im Vergleich zu Frauen motivierter waren, an der Befragung teilzunehmen.

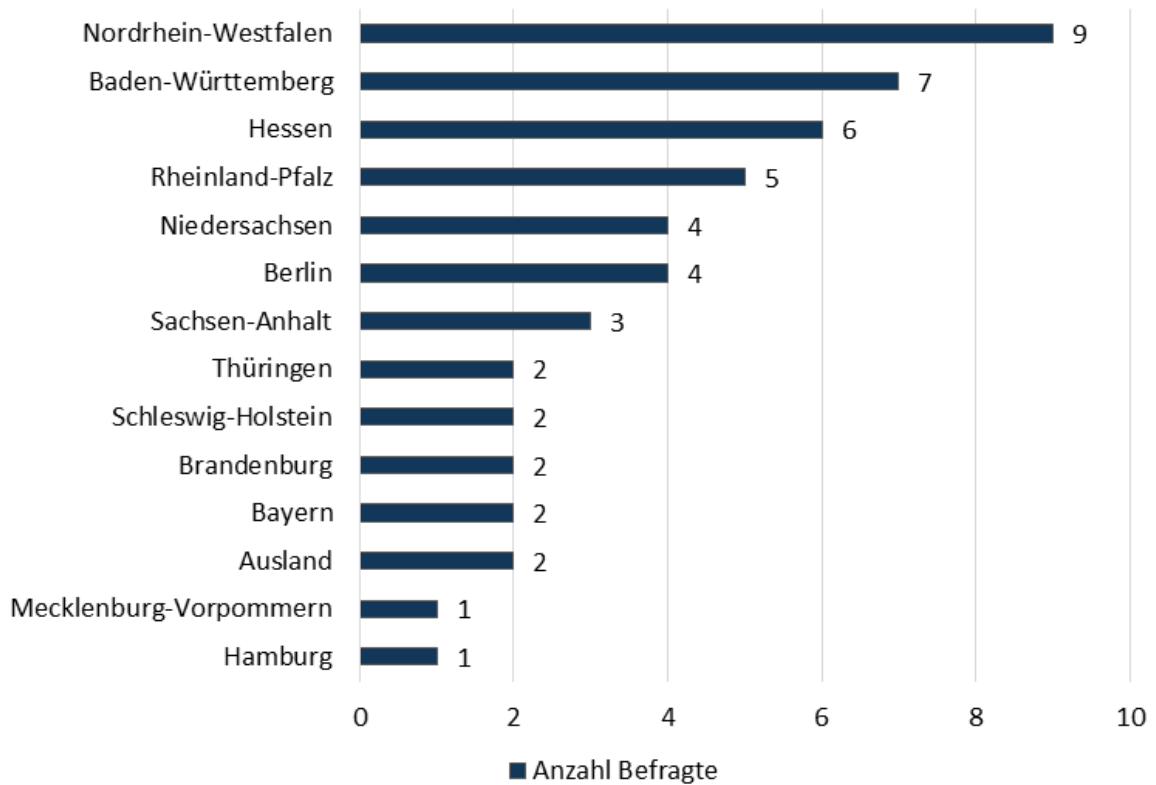
» **Abbildung 28: Geschlecht der Befragten**



» Anmerkung: Prozentuierung auf Basis gültiger Werte (N = 91).

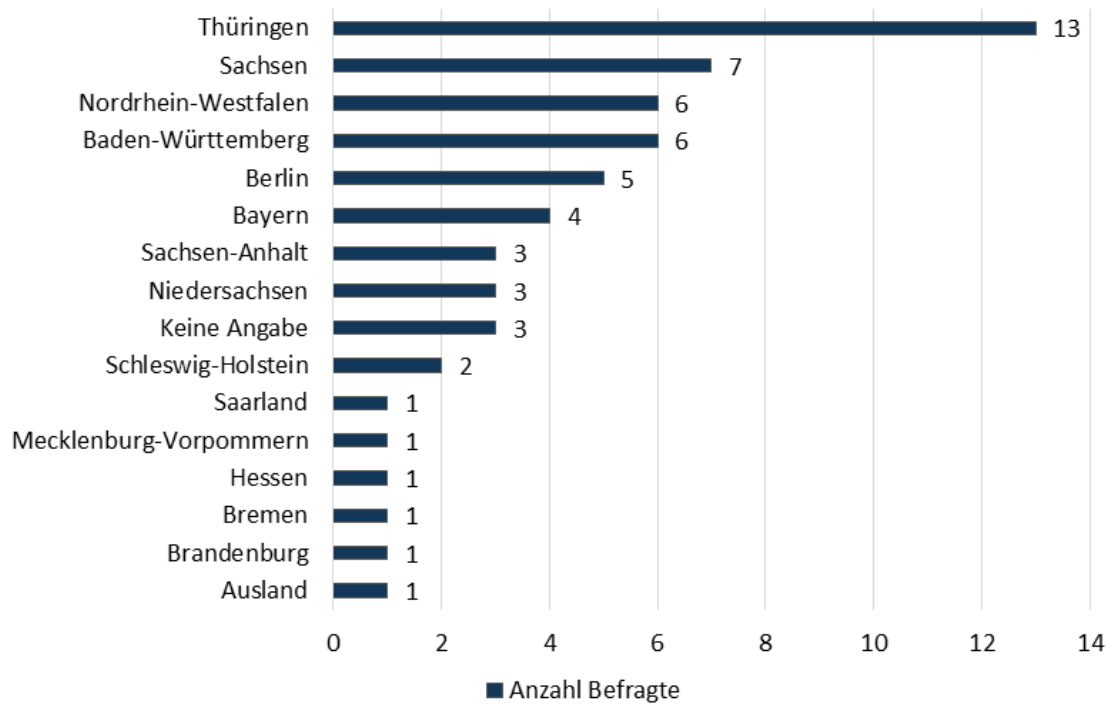
Die folgenden Abbildungen zeigen die Angaben der Teilnehmenden zum Wohnort sowie zum Arbeitsort. In Abbildung 29 wird ersichtlich, dass besonders viele der Befragten in Nordrhein-Westfalen, Baden-Württemberg oder Hessen lebten. Der häufigste Arbeitsort unter den Bundesländern war Thüringen (siehe Abbildung 30).

» **Abbildung 29: Wohnort der Befragten**



» Anmerkung: Berechnung auf Basis gültiger Werte (N = 50). Die Frage wurde ausschließlich Personen gestellt, die angaben, GovData.de privat zu nutzen oder zur Frage nach der privaten oder beruflichen Nutzung keine Angabe machen wollten.

» **Abbildung 30: Arbeitsort der Befragten**

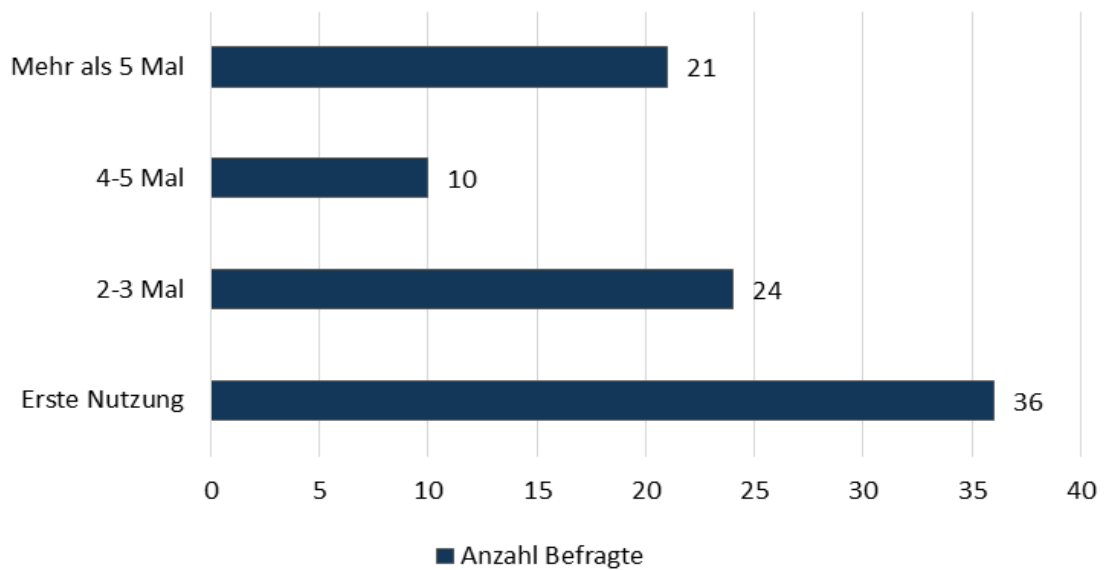


» Anmerkung: Berechnung auf Basis gültiger Werte (N = 58). Die Frage wurde ausschließlich Personen gestellt, die angaben, GovData.de beruflich zu nutzen.

3.4.2 Nutzung von GovData.de

38 Prozent der Teilnehmenden gaben an, GovData.de das erste Mal genutzt zu haben. Daraus lässt sich schließen, dass ein großer Teil der Befragten bisher recht wenige Erfahrungen mit der Plattform gesammelt hat. Wie in Abbildung 31 ersichtlich wird, hat mit 58 Prozent die Mehrheit der Teilnehmenden GovData.de jedoch mindestens zweimal genutzt.

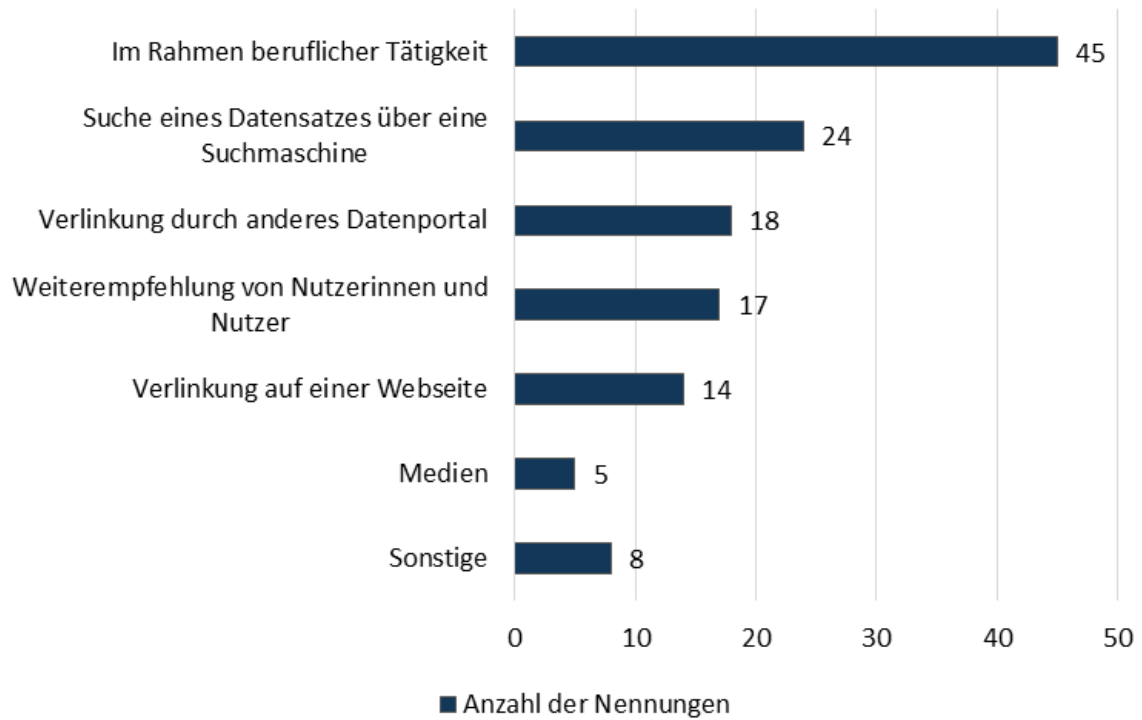
» **Abbildung 31: „Wie häufig haben Sie GovData in den letzten zwölf Monaten genutzt?“**



» Anmerkung: Berechnung auf Basis gültiger Werte (N = 91). Kleinere Abweichungen in den Fallzahlen gehen auf fehlende Antworten zurück, hierdurch kommt es auch zu Abweichungen zu den im Text genannten Prozentwerten.

Abbildung 32 zeigt, dass nahezu die Hälfte der Befragten durch ihre beruflichen Tätigkeiten von GovData.de erfahren hat. Etwa ein Viertel der Teilnehmenden gab zudem an, dass sie auf GovData.de gestoßen sind, als sie in einer Suchmaschine nach einem bestimmten Datensatz gesucht haben. Darüber hinaus wurden die Befragten auch durch die Verlinkung auf anderen Datenportalen oder Webseiten auf GovData.de aufmerksam gemacht. In einem Freitextfeld konnten die Teilnehmenden spezifizieren, um welches Datenportal oder um welche Webseite es sich dabei handelte. Genannt wurden in diesem Zusammenhang Data.gv.at, service-bw.de, data.europa.eu, opendata.bonn.de, geoportal.rlp.de, geoportal-hamburg.de, die Metadatenkataloge der GDI-DE, datenadler.de, geobserver.wordpress.com und das Bundeskriminalamt. Interessant ist zudem, dass einigen Befragten (18 Prozent) GovData.de von anderen Nutzerinnen und Nutzern empfohlen wurde.

» **Abbildung 32: „Wie sind Sie auf GovData aufmerksam geworden?“**

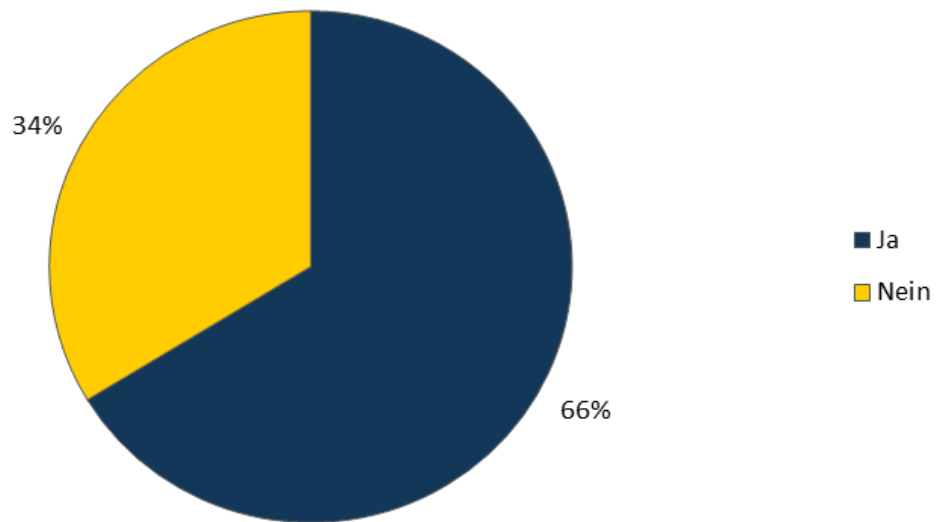


» Anmerkung: Berechnung auf Basis gültiger Werte (N = 94). Mehrfachantworten waren möglich.

Befragte, die über sonstige Wege von GovData.de erfahren haben, gaben in einem Freitextfeld als Informationsquelle das Studium, die Datenstrategie der Bundesregierung, einen Newsletter oder Veranstaltungen an.

Auf die Frage nach dem Kontext der Nutzung von GovData.de, gaben etwa zwei Drittel der Befragten an, GovData.de beruflich zu nutzen, während etwa ein Drittel GovData.de nicht im beruflichen Kontext nutzt (siehe Abbildung 33).

» **Abbildung 33: „Nutzen Sie GovData.de beruflich?“**

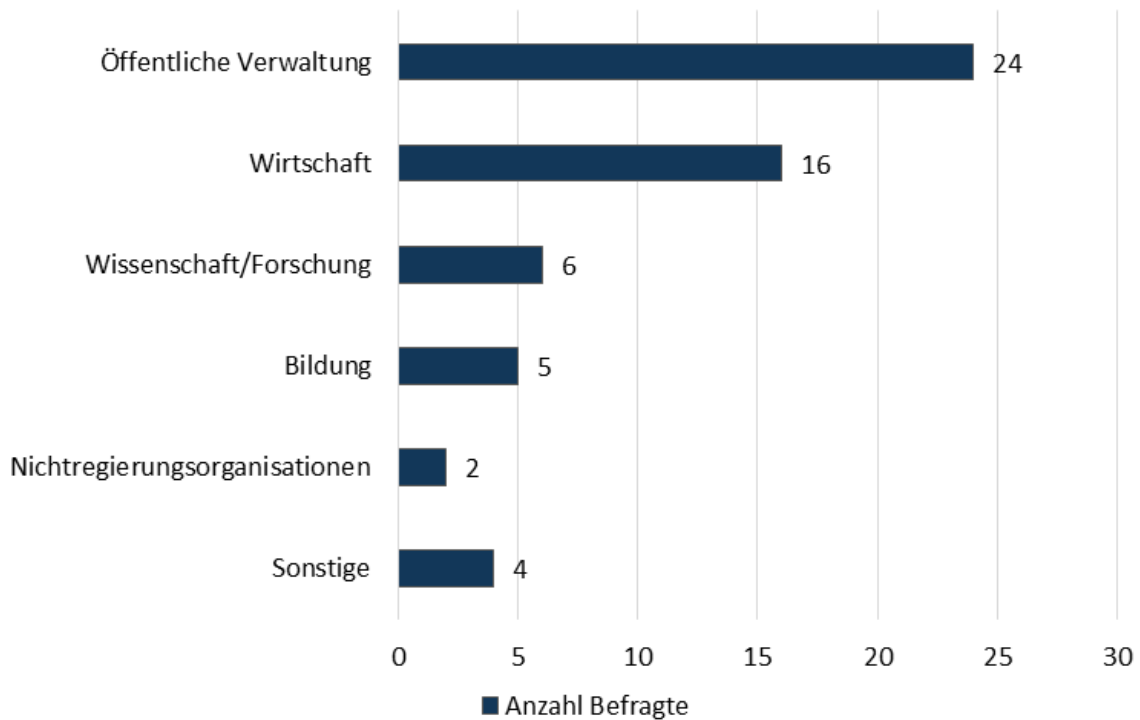


» Anmerkung: Prozentuierung auf Basis gültiger Werte (N = 89). Kleinere Abweichungen in den Fallzahlen gehen auf fehlende Antworten zurück, hierdurch kommt es auch zu Abweichungen zu den im Text genannten Prozentwerten.

3.4.2.1 Berufliche Nutzung

Mit 42 Prozent waren die meisten der insgesamt 57 Personen, die GovData.de im beruflichen Kontext nutzten, in der öffentlichen Verwaltung beschäftigt (siehe Abbildung 34). Etwas weniger arbeiteten in der Wirtschaft (28 Prozent). 11 Prozent waren zudem in der Wissenschaft und 9 Prozent im Bildungssektor tätig. Nur zwei Beschäftigte von Nichtregierungsorganisationen waren unter den Befragten vertreten. Während eine Person in diesem Zusammenhang keine genaueren Angaben zur Organisation machte, gab die andere Person in einem Freitextfeld an, auch für den Naturschutzbund Deutschland (NABU) zu arbeiten.

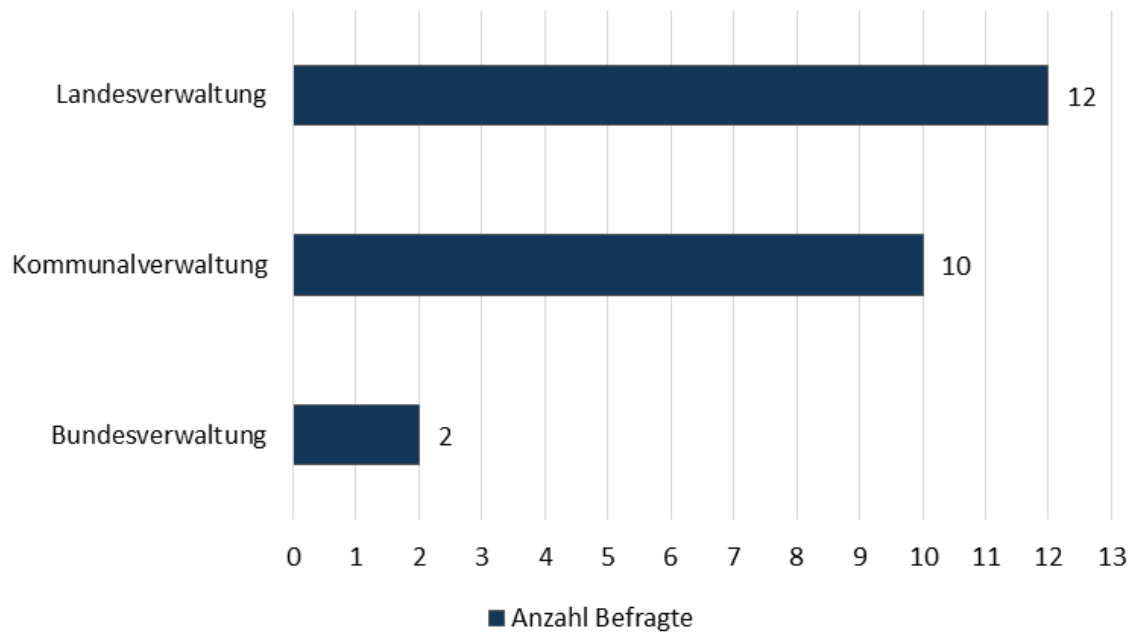
» **Abbildung 34: „In welchem Bereich sind Sie berufstätig?“**



» Anmerkung: Berechnung auf Basis gültiger Werte (N = 57). Die Frage wurde ausschließlich Personen gestellt, die angaben, GovData.de beruflich zu nutzen. Kleinere Abweichungen in den Fallzahlen gehen auf fehlende Antworten zurück.

Unter den beruflichen Nutzerinnen und Nutzern, die in der öffentlichen Verwaltung beschäftigt sind, waren zwölf in einer Landes- und zehn in einer Kommunalverwaltung tätig. Nur zwei Personen arbeiteten in der Bundesverwaltung (siehe Abbildung 35).

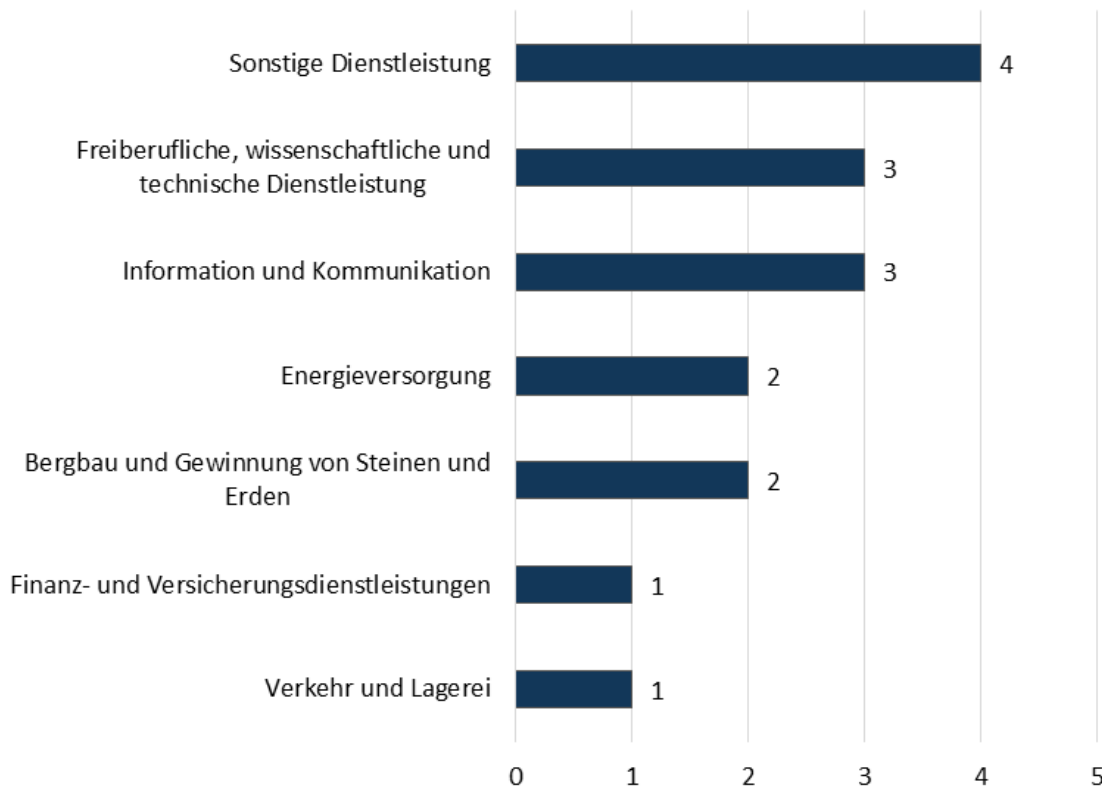
» **Abbildung 35: „Auf welcher Verwaltungsebene sind Sie tätig?“**



» Anmerkung: Berechnung auf Basis gültiger Werte (N = 24). Die Frage wurde ausschließlich Personen gestellt, die angaben in der öffentlichen Verwaltung zu arbeiten.

Unter den 16 Teilnehmenden, die in der Wirtschaft beschäftigt waren, gaben vier Personen an, sonstige Dienstleistungen zu erbringen, während jeweils drei Personen im Bereich „Information und Kommunikation“ oder „Freiberufliche, wissenschaftliche und technische Dienstleistungen“ arbeiteten (siehe Abbildung 36). Weitere vertretene Wirtschaftszweige waren „Bergbau und Gewinnung von Steinen und Erden“, „Energieversorgung“, „Verkehr und Lagerei“ sowie „Finanz- und Versicherungsdienstleistungen“.

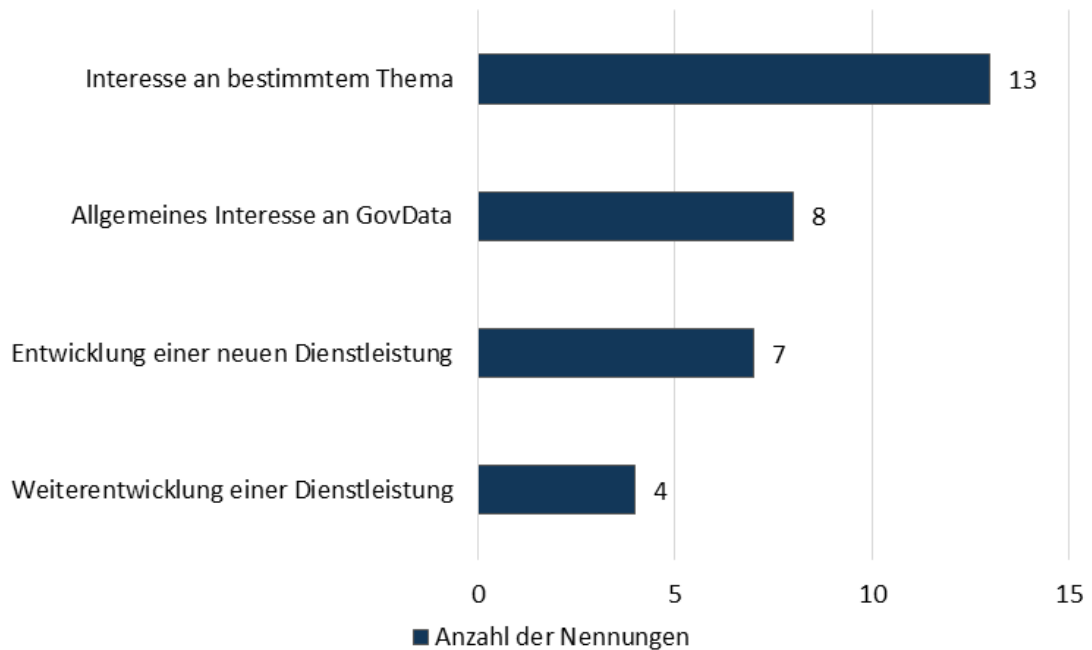
» **Abbildung 36: „In welchem Wirtschaftszweig sind Sie tätig?“**



» Anmerkung: Berechnung auf Basis gültiger Werte (N = 16). Die Frage wurde ausschließlich Personen gestellt, die angaben, in der Wirtschaft zu arbeiten. Kleinere Abweichungen in den Fallzahlen gehen auf fehlende Antworten zurück.

Besonders viele Befragte aus dem Bereich Wirtschaft nannten ein Interesse an einem bestimmten Thema als Grund für ihre berufliche Nutzung von GovData.de (siehe Abbildung 37). Während etwa die Hälfte der Teilnehmenden aus der Wirtschaft GovData.de zudem aufgrund eines generellen Interesses an GovData.de in Anspruch nahmen, gaben 44 Prozent jener Befragten an, GovData.de für die Entwicklung einer neuen Dienstleistung und ein Viertel für die Weiterentwicklung einer Dienstleistung zu nutzen. Die Nutzung für die (Weiter-) Entwicklung von Dienstleistungen deutet darauf hin, dass das Angebot von GovData.de für die wirtschaftliche Wertschöpfung durchaus von Bedeutung sein kann.

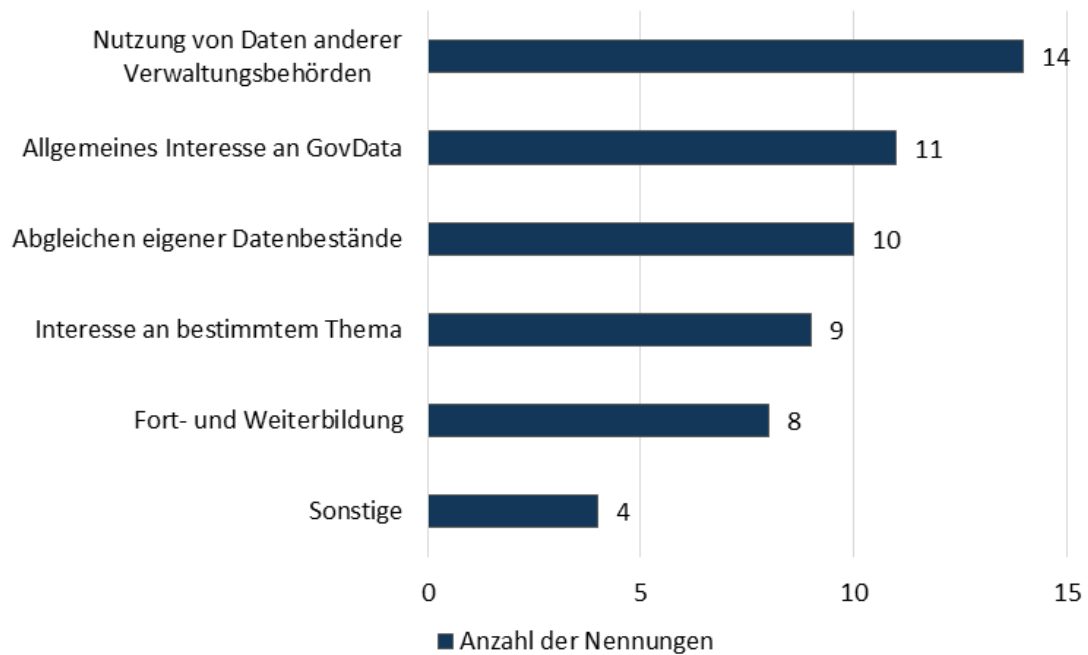
» **Abbildung 37: „Aus welchen beruflichen Gründen nutzen Sie GovData?“ (Beschäftigte in der Wirtschaft)**



» Anmerkung: Berechnung auf Basis gültiger Werte (N = 16). Die Frage wurde ausschließlich Personen gestellt, die angaben, in der Wirtschaft zu arbeiten. Mehrfachantworten waren möglich.

Während bei den Beschäftigten aus der Wirtschaft das Interesse an einem bestimmten Thema der mit Abstand häufigste Nutzungsgrund war, spielte dieser Faktor bei den Beschäftigten in der öffentlichen Verwaltung eine eher untergeordnete Rolle. Wie in Abbildung 38 ersichtlich wird, wurden in dieser Gruppe am häufigsten der Zugang zu Daten anderer Verwaltungsbehörden sowie ein allgemeines Interesse an GovData.de als Nutzungsgründe genannt.

» **Abbildung 38: „Aus welchen beruflichen Gründen nutzen Sie GovData?“ (Beschäftigte in der öffentlichen Verwaltung)**

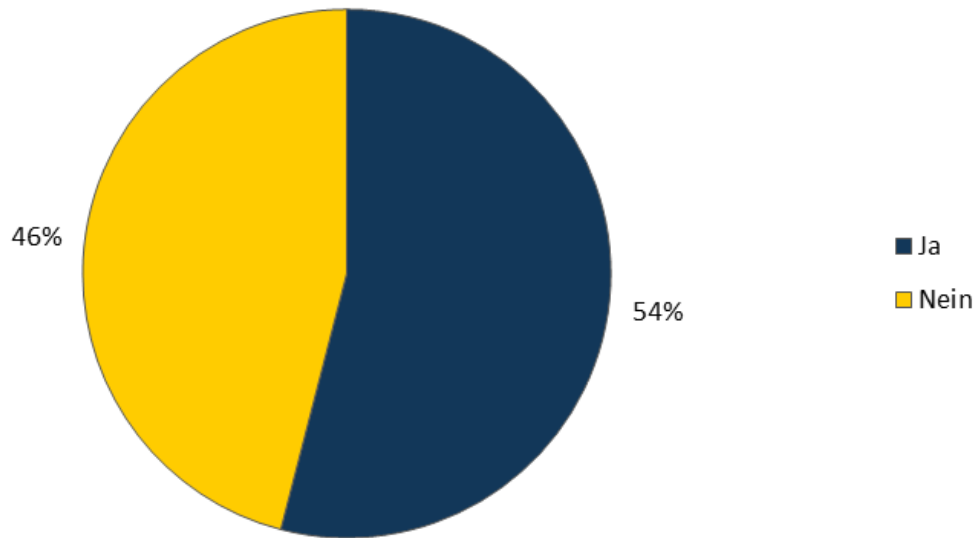


» Anmerkung: Berechnung auf Basis gültiger Werte (N = 24). Die Frage wurde ausschließlich Personen gestellt, die angaben, in der öffentlichen Verwaltung zu arbeiten. Mehrfachantworten waren möglich.

3.4.2.2 Private (oder ehrenamtliche) Nutzung

Insgesamt sagten 47 Befragte aus, GovData.de privat zu nutzen (siehe Abbildung 39).

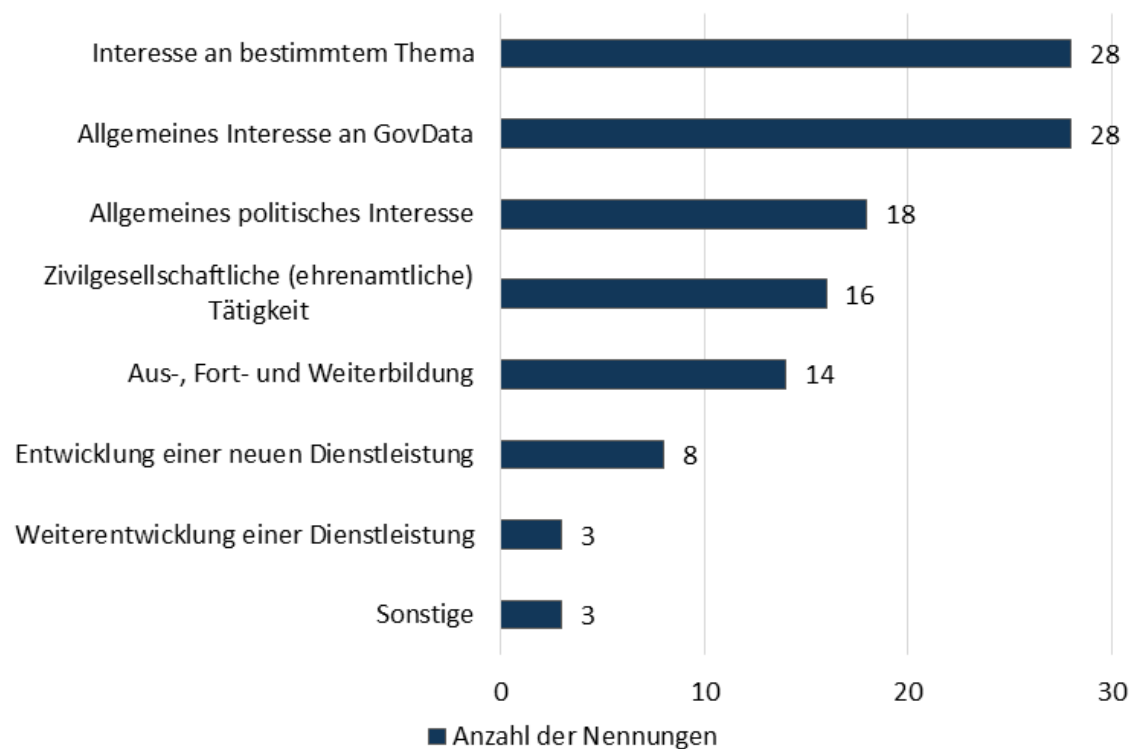
» **Abbildung 39: „Nutzen Sie GovData.de privat?“**



» **Anmerkung:** Prozentuierung auf Basis gültiger Werte (N = 87). Kleinere Abweichungen in den Fallzahlen gehen auf fehlende Antworten zurück.

Davon nannten jeweils mehr als die Hälfte das Interesse für ein bestimmtes Thema und ein allgemeines Interesse an GovData.de als Grund (siehe Abbildung 40). Dass darüber hinaus Faktoren wie zivilgesellschaftliches Engagement, Aus-, Fort- und Weiterbildung oder die Entwicklung neuer Dienstleistungen ebenso wichtige Gründe für die private Nutzung waren, verdeutlicht die enge Verknüpfung von Angeboten offener Verwaltungsdaten und sozialer sowie wirtschaftlicher Wertschöpfung.

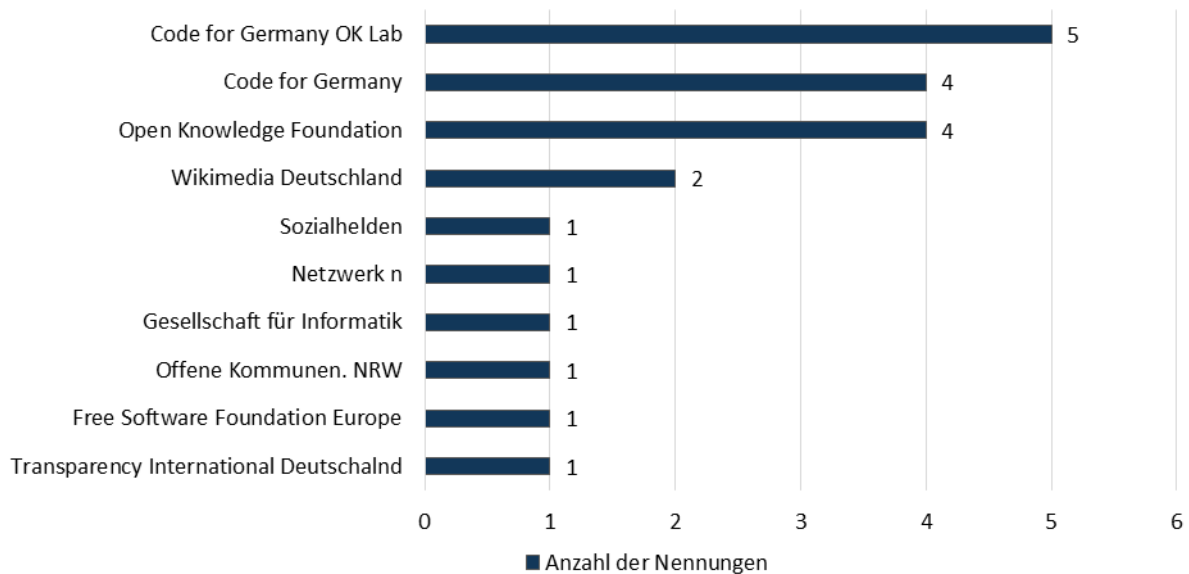
» **Abbildung 40: „Aus welchen privaten Gründen nutzen Sie GovData.de?“**



» **Anmerkung:** Berechnung auf Basis gültiger Werte (N = 47). Die Frage wurde ausschließlich Personen gestellt, die angaben, GovData.de privat zu nutzen. Mehrfachantworten waren möglich.

Abbildung 41 zeigt, dass zu den Organisationen, in denen sich die Befragten (ehrenamtlich) engagierten, vor allem Organisationen im Bereich Open Government zählten. Dieser Befund ist wenig überraschend, da offene Verwaltungsdaten ein Element des Open Government ist. Weitere Organisationen, die jeweils einmal im Freitextfeld angegeben wurden, waren die Freiwillige Feuerwehr, Kulturdb.de, NABU und die W3C OpenGov Community Group.

» **Abbildung 41: Zivilgesellschaftliche Organisationen, für die GovData.de genutzt wurde**

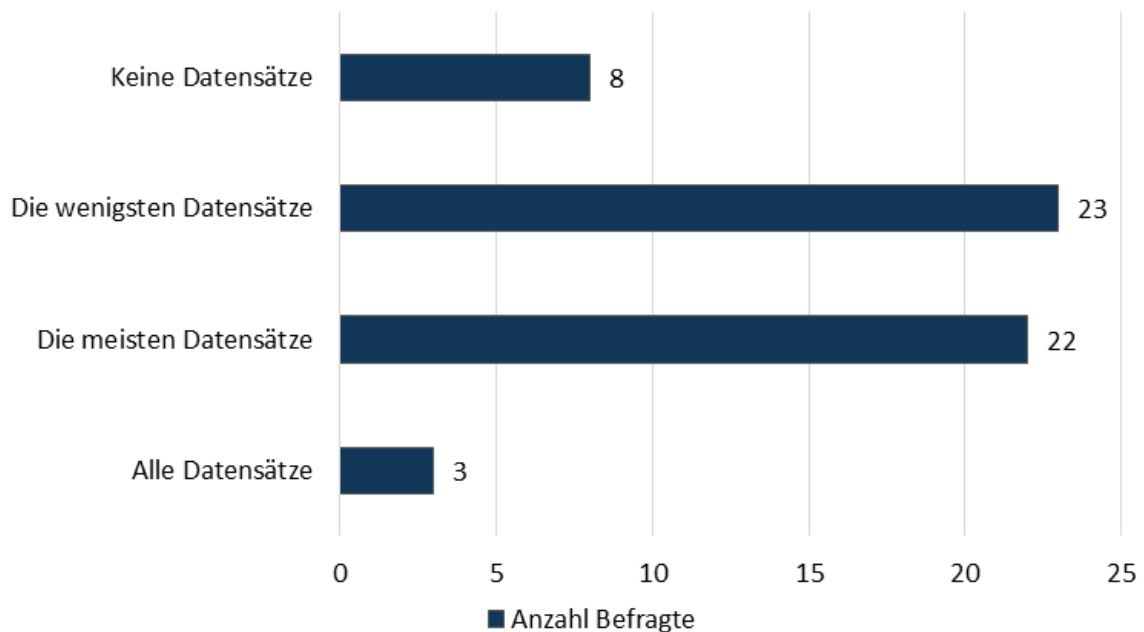


» **Anmerkung:** Berechnung auf Basis gültiger Werte (N = 16). Die Frage wurde ausschließlich Personen gestellt, die angaben, GovData.de im Rahmen einer zivilgesellschaftlichen (ehrenamtlichen) Tätigkeit zu nutzen. Mehrfachantworten waren möglich.

3.4.3 Bewertung von GovData.de

Um mehr darüber zu erfahren, wie zufrieden die Teilnehmenden mit GovData.de waren, wurden sie zunächst gefragt, ob sie die Datensätze, nach denen sie gesucht haben, auch gefunden haben. Von den 57 Teilnehmenden, die in den letzten zwölf Monaten nach mindestens einem bestimmten Datensatz gesucht haben, gaben 39 Prozent an, die meisten Datensätze gefunden zu haben (siehe Abbildung 42). Ebenso viele Befragte hatten bei Ihrer Suche nach Datensätzen jedoch nur in den wenigsten Fällen Erfolg. Während 14 Prozent der Befragten keine Datensätze gefunden haben, waren für fünf Prozent alle Datensätze auffindbar.

» **Abbildung 42: „Haben Sie den gesuchten Datensatz/die gesuchten Datensätze auf GovData.de gefunden?“**

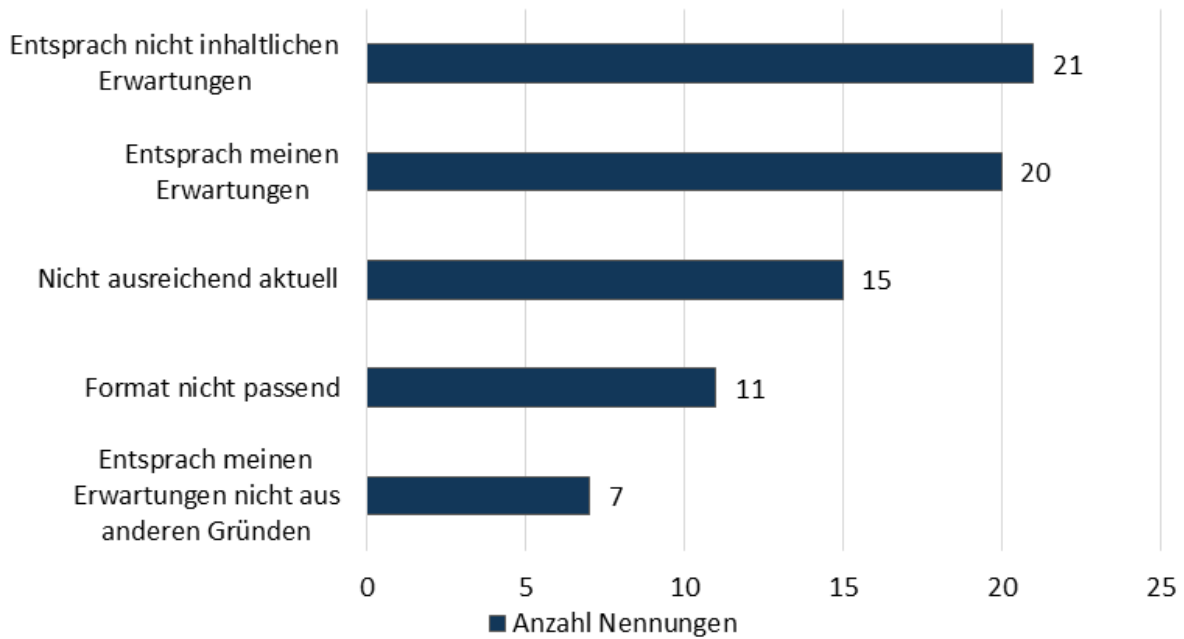


» **Anmerkung:** Berechnung auf Basis gültiger Werte (N = 56). Die Frage wurde ausschließlich Personen gestellt, die angaben, in den letzten zwölf Monaten auf GovData.de nach einen oder mehreren Datensätzen gesucht zu haben. Kleinere Abweichungen in den Fallzahlen gehen auf fehlende Antworten zurück.

Erfreulich ist, dass bei 41 Prozent der Befragten, die mindestens einen Datensatz gefunden haben, der Datensatz den eigenen Erwartungen entsprach (siehe Abbildung 43). Bei 85 Prozent der Befragten wurden die Erwartungen allerdings auch zumindest in Bezug auf einzelne Aspekte enttäuscht. In diesem Zusammenhang gaben die meisten Teilnehmenden an, dass die Daten nicht den inhaltlichen Erwartungen gerecht wurden. Unzufriedenheit zeigte sich zudem in Bezug auf die mangelnde Aktualität der Daten oder unpassenden Dateiformaten.

Befragte, die aus anderen Gründen nicht zufrieden waren, begründeten dies in einem Freitextfeld damit, dass der räumliche Filter nicht gepasst habe oder dass es für die Daten keine Lizenzen gebe. Darüber hinaus kritisierte eine Person, dass wesentliche Informationen fehlten und zu den Datensätzen lediglich Kategorien genannt würden. Ebenso beklagt wurde, dass ein INSPIRE-konformer Service noch eine Seltenheit sei, oder aber, dass der Link zu einem Webservice nicht funktioniere.

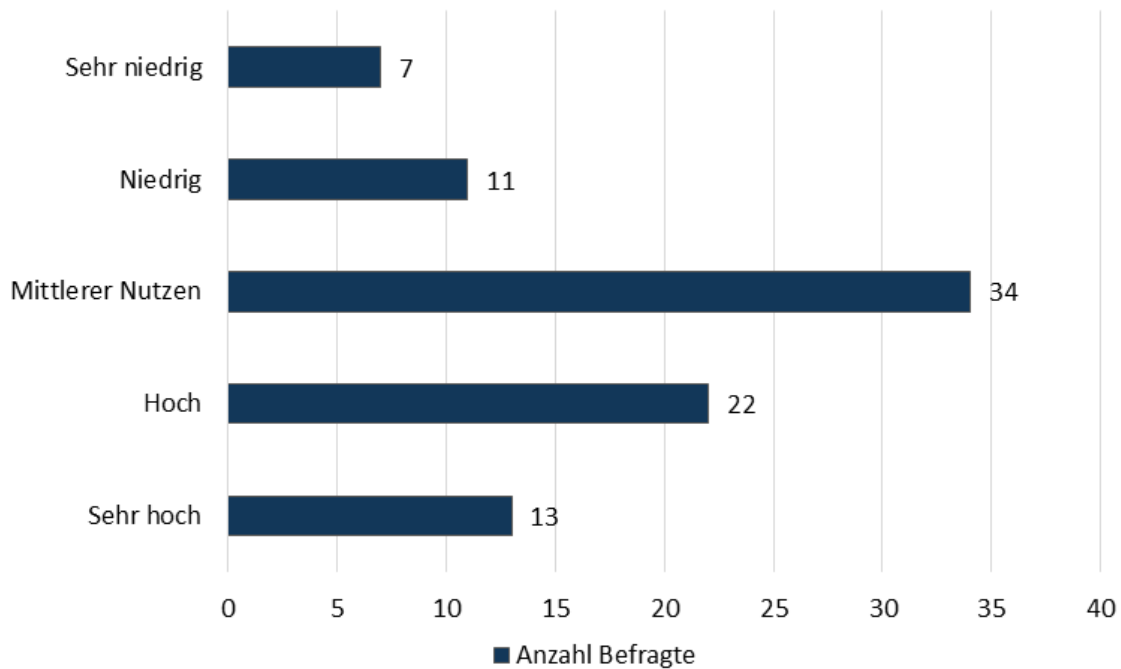
» **Abbildung 43: „Entsprach der Datensatz Ihren Erwartungen?“**



» Anmerkung: Berechnung auf Basis gültiger Werte (N = 49). Die Frage wurde ausschließlich Personen gestellt, die angaben, mindestens einen Datensatz gefunden zu haben. Mehrfachantworten waren möglich.

Des Weiteren wurden die Teilnehmenden gebeten, den Nutzen von GovData.de für ihre Zwecke einzuschätzen. Wie in Abbildung 44 ersichtlich, schätzten 37 Prozent der Befragten, dass GovData.de einen mittleren Nutzen hat. Als hoch oder sehr hoch bewerteten den Nutzen 38 Prozent der teilnehmenden Personen. Deutlich weniger Teilnehmende sahen die Plattform dagegen negativer: So gaben 20 Prozent an, dass GovData.de einen niedrigen oder gar einen sehr niedrigen Nutzen habe.

» **Abbildung 44: „Wie hoch schätzen Sie den Nutzen von GovData.de für Ihre Zwecke ein?“**



» Anmerkung: Berechnung auf Basis gültiger Werte (N =87). Kleinere Abweichungen in den Fallzahlen gehen auf fehlende Antworten zurück, hierdurch kommt es auch zu Abweichungen zu den im Text genannten Prozentwerten.

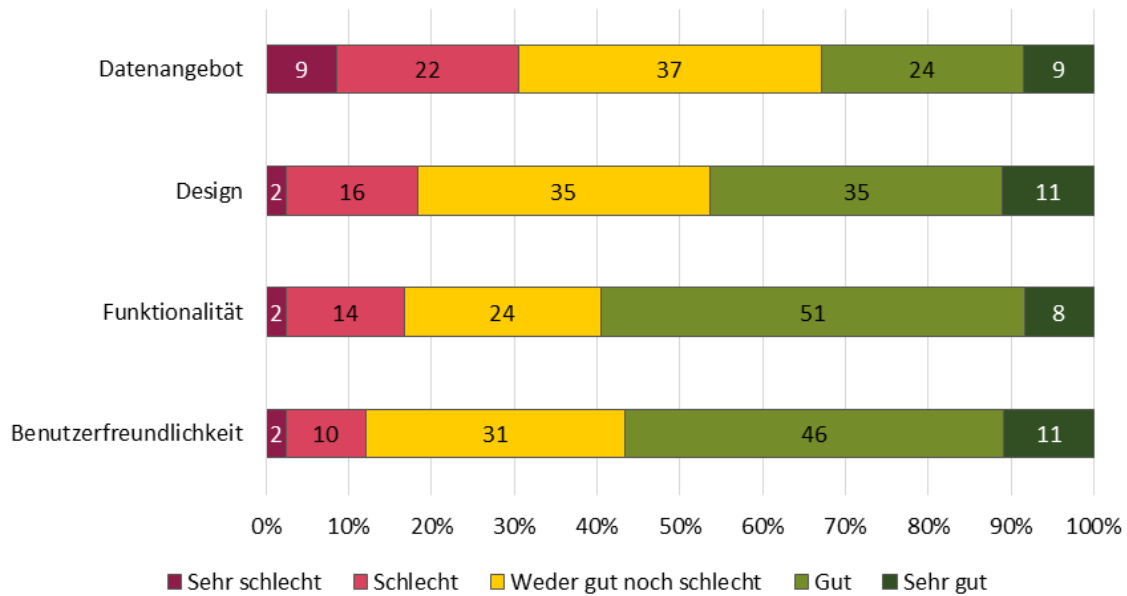
Die Befragten wurden zusätzlich gebeten, Ihre Einschätzung in einem Freitextfeld zu begründen. Unter denjenigen, die den Nutzen als sehr hoch oder hoch einschätzten, wiesen mehrere auf den Vorteil einer zentralen Plattform hin. Den hohen Nutzen sahen einige Befragte zudem für ihre Forschung – sowohl im akademischen Kontext als auch für private Zwecke. So war GovData.de für eine Person für die persönliche Ahnenforschung von hohem Nutzen.

Die Befragten, die den Nutzen von GovData.de als mittel, niedrig oder gar sehr niedrig einstufen, beklagten veraltete und wenig relevante Daten, ungeeignete Dateiformate sowie eine unübersichtliche Struktur der Plattform. Eine Person bemängelte zudem, dass die Daten nicht konsolidiert vorlägen. Zwei Befragte forderten, dass die Daten über eine API abrufbar sein müssten. Das sei laut einer Person notwendig, um „aufbauend auf den Daten Dienste und Visualisierungen entwickeln zu können“. Eine befragte Person wies darüber hinaus darauf hin, dass Datensätze (insbesondere Geobasisdaten) aufgrund der länderspezifischen Open Data-Initiativen „nur separat und keinesfalls länderübergreifend/flächendeckend angeboten“ würden. Die Recherche sei vor diesem Hintergrund mühsam.

Zudem wurden die Teilnehmenden gebeten, die verschiedenen Aspekte der GovData.de-Plattform zu bewerten. Bei den Aspekten handelte es sich um das Datenangebot, das Design, die Funktionalität und die Benutzerfreundlichkeit. Auffällig ist zunächst, dass das Datenangebot und somit die Kernaufgabe von GovData.de vergleichsweise schlecht abschnitt (siehe Abbildung 45). So wurden das Design, die Funktionalität und die Benutzerfreundlichkeit der Plattform häufiger als sehr gut oder gut bewertet als das Datenangebot. Ebenso wurde das Datenangebot häufiger als schlecht oder sehr schlecht beurteilt als die anderen Aspekte. Jeweils unter 20 Prozent der Teilnehmenden, die die Fragen zur Benutzerfreundlichkeit von GovData.de beantworteten, bewerteten das Design, die Funktionalität

oder die Benutzerfreundlichkeit von GovData.de als schlecht oder sehr schlecht. Beim Datenangebot fiel die Bewertung bei 31 Prozent schlecht oder sehr schlecht aus.

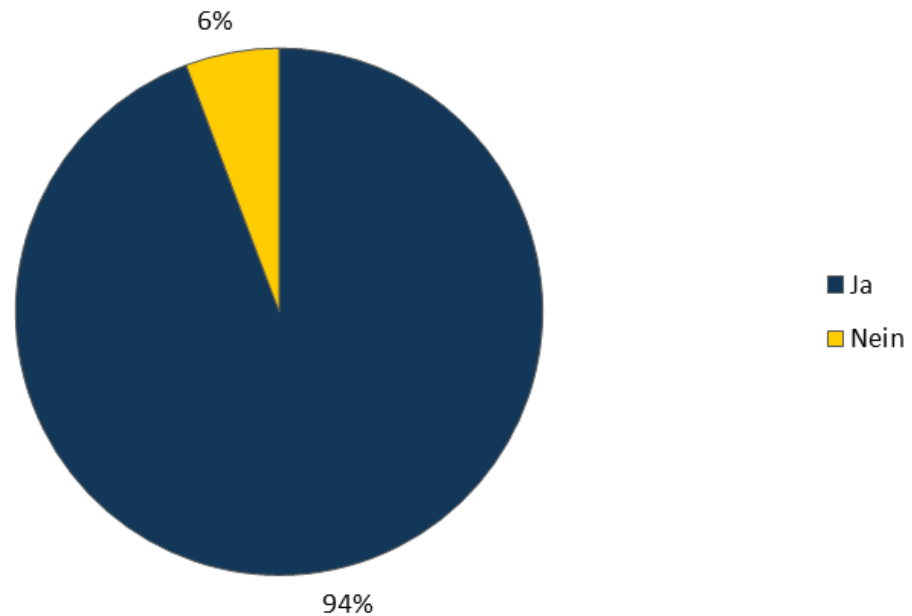
» **Abbildung 45: „Bitte bewerten Sie die folgenden Aspekte der Open Data-Plattform GovData.de“**



» Anmerkung: Prozentuierung auf Basis gültiger Werte (N = 88-91). Kleinere Abweichungen in den Fallzahlen gehen auf fehlende Antworten zurück.

Die Teilnehmenden wurden zudem gefragt, ob sie sich die Bereitstellung von mehr Daten wünschten. Das Ergebnis in Abbildung 46 zeigt, dass dies bei 88 Prozent der Befragten der Fall war.

» **Abbildung 46: „Wünschen Sie sich die Bereitstellung weiterer Daten?“**

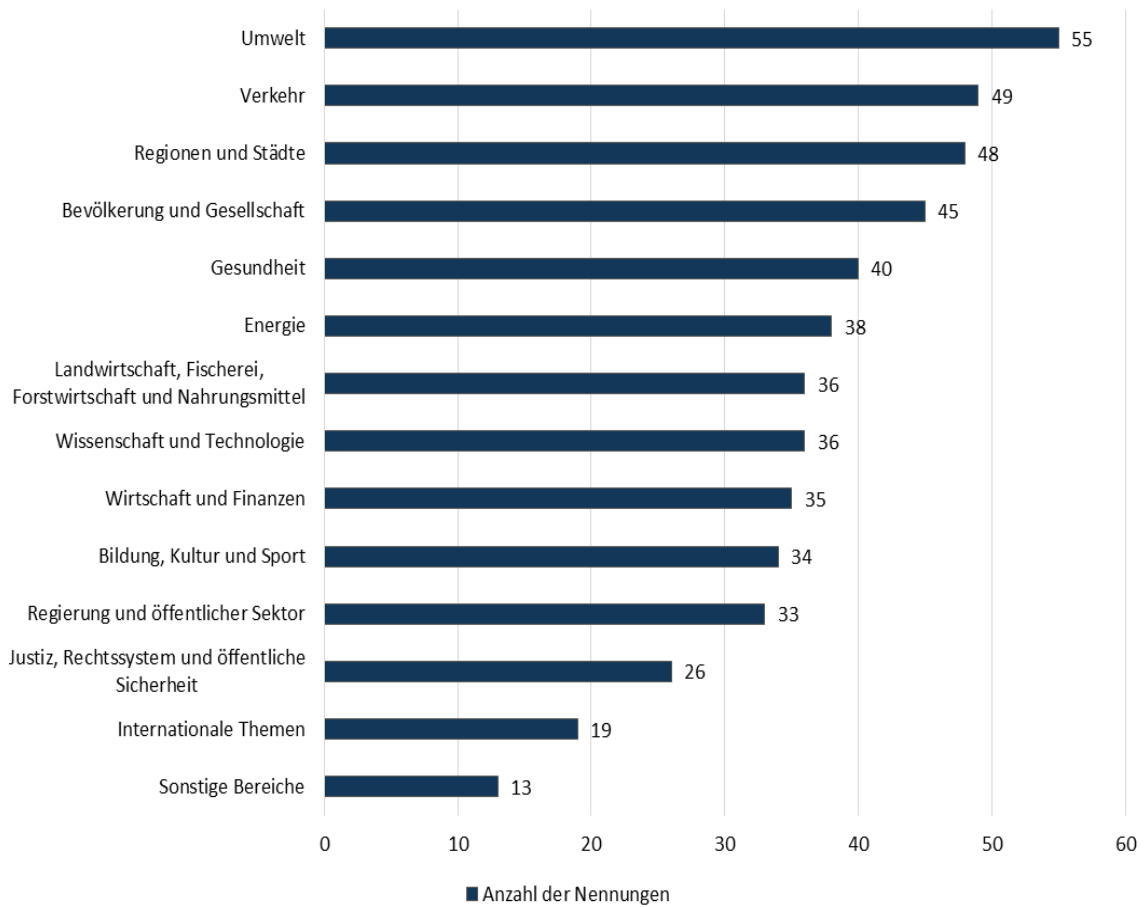


» Anmerkung: Prozentuierung auf Basis gültiger Werte (N = 87). Kleinere Abweichungen in den Fallzahlen gehen auf fehlende Antworten zurück.

Den Teilnehmenden wurde die Möglichkeit gegeben, die Bereiche, in denen sie sich mehr Daten wünschten, weiter zu spezifizieren. Mehrfach wurde in diesem Zusammenhang der Bedarf nach Geodaten angegeben. Weitere Bereiche waren Infrastruktur, Kinder, Bioterrorismus, Integration, Kriegsgräber und Denkmäler, Mobilität, Politik und Verwaltung. Eine Person wünschte sich zudem Daten von mit Steuermitteln unterstützten Studien und Forschungsergebnissen.

Auf die Frage, in welchem Bereich mehr Daten wünschenswert wären, gab ein besonders großer Anteil der Befragten den Bereich Umwelt an (siehe Abbildung 47). Bereiche mit einem großen Datenbedarf waren zudem Verkehr, Regionen und Städte, Bevölkerung und Gesellschaft sowie Gesundheit.

» **Abbildung 47: „In welchen Bereichen wünschen Sie sich weitere Daten?“**



» **Anmerkung:** Prozentuierung auf Basis gültiger Werte (N = 82). Die Frage wurde ausschließlich Personen gestellt, die angaben, sich mehr Daten zu wünschen. Mehrfachantworten waren möglich.

Zuletzt wurden die Teilnehmenden gefragt, ob sie Vorschläge für die Weiterentwicklung von GovData.de hätten. Einige dieser Vorschläge betrafen Funktionen, welche in die Plattform integriert werden könnten. So schlug eine teilnehmende Person die Möglichkeit vor, interessante Treffer zu sammeln (womöglich in einem eigenen Profil). Darüber hinaus wurde vorgeschlagen, eine Option zu integrieren, die es ermöglicht, Daten direkt auszuwerten, darzustellen und zu teilen. Eine Person empfahl zudem eine in der Plattform integrierte Datenvorschau, sowie eine Performance-/Verfügbarkeitsmessung von Webservices. Eine Empfehlung war darüber hinaus, auf dem Portal einfache Abfragen in den Datensätzen zu erlauben.

Ebenso wurde es von mehreren Befragten für sinnvoll erachtet, wenn Datensätze miteinander kombinierbar oder bereits miteinander verknüpft wären. In diesem Zusammenhang wurde von einer Person eine Verknüpfung der Datensätze mit „entsprechenden Initiativen auf EU- und UN-Ebene“ angeraten. Laut der Befragten oder dem Befragten könne GovData.de auch geschlossen werden, wenn alle Daten diesen Initiativen zur Verfügung gestellt werden würden. Mehrfach wurde angemerkt, dass GovData.de die Datensätze verschiedener Gebietskörperschaften aggregieren solle. Die Verknüpfung von Datensätzen könne über Linked Data im RDF-Format vereinfacht werden. Außerdem wurde eine Verknüpfung von GovData.de mit MetaVER vorgeschlagen.

Laut mehreren Befragten sollten die Daten zudem in einem einheitlichen und maschinenlesbaren Format vorliegen (wie XML oder JSON-ld). Eine Person bemerkte, dass nicht nur das Dateiformat, sondern ebenso die Erfassung der Daten einheitlich sein sollte. So werde bei Bebauungsplänen beispielsweise „teilweise der eigentliche Plan und teilweise nur der Umring bereitgestellt“. In Bezug auf Geodaten wünschte sich eine Person eine Verknüpfung mit geoportal.de oder einen Hinweis darauf, dass Daten auch als Geodaten (und nicht nur als Tabellen und Grafiken) vorliegen. Eine weitere Person forderte zudem mehr Formate für Geoinformationssysteme, wie GML und GEOJSON („nicht nur Schnittstellen wie WMS und WFS“). Generell wurde mehrfach vorgeschlagen, dass Daten über eine offene API abrufbar sein sollten.

In Bezug auf die Datensuche bemerkten mehrere Befragte, dass die räumliche Suche durch ein anwenderfreundliches Filtern der Daten vereinfacht werden sollte. Sehr konkret schlug eine Person eine Filterung nach dem 5-Sterne-Modell⁵⁰ für offene Daten vor. Die gefundenen Datensätze sollten auf einer Karte dargestellt werden. Auch eine weitere Person wünschte sich „eine Suchfunktion über eine Karte oder nach Bundesland“.

Eine weitere kleine Anmerkung, die sehr einfach umzusetzen wäre, war ein Hinweis auf die Dateigröße des Datensatzes vor dem Download. Vorgeschlagen wurde zudem mehr Netzwerk- und Öffentlichkeitsarbeit und die Entwicklung von Fortbildungs- und Professionalisierungsangeboten im Bereich Open Data. Bezüglich der Daten schlug eine weitere Person zudem die Bereitstellung kompletter Zeitreihen vor.

3.4.4 Zusammenfassung

Zwei Drittel der Befragten gaben an, im beruflichen Kontext auf GovData.de zurückzugreifen. Unter den Befragten, die GovData.de beruflich nutzten, waren besonders viele Personen, die in der öffentlichen Verwaltung (vor allem Landesverwaltung und Kommunalverwaltung) tätig waren. Diese nutzten GovData.de vor allem für den Zugang zu Daten von anderen Verwaltungsbehörden. Recht viele Befragte waren zudem in der Wirtschaft beschäftigt und gaben als Nutzungsgrund am häufigsten das Interesse an einem bestimmten Thema an. Nahezu die Hälfte der Teilnehmenden aus der Wirtschaft nutzten die Daten von GovData.de zudem für die Entwicklung einer neuen Dienstleistung – dieser Befund verdeutlicht den Nutzen von offenen Daten für die wirtschaftliche Wertschöpfung.

Privat wurde GovData.de von etwa der Hälfte der Befragten genutzt. Als häufigster Grund für die private Nutzung wurde das Interesse an einem bestimmten Thema oder allgemeines Interesse an GovData.de genannt. Dass in Bezug auf die private Nutzung jedoch auch zivilgesellschaftliches Engagement, Aus-, Fort- und Weiterbildung sowie die Entwicklung einer neuen Dienstleistung als Faktoren genannt wurden, könnte auf die Bedeutung von offenen Daten für die wirtschaftliche und soziale Wertschöpfung hindeuten.

In Bezug auf die Zufriedenheit mit GovData.de zeigt sich ein gemischtes Bild. Während 44 Prozent der Befragten, die in den letzten zwölf Monaten mindestens nach einem Datensatz gesucht hatten, die

⁵⁰ Siehe: <https://5stardata.info/de/> (Abrufdatum: 17. August 2021).

meisten oder alle Datensätze auch gefunden haben, gaben 40 Prozent an, die wenigsten und gar 14 Prozent keine Datensätze gefunden zu haben.

Positiv hervorzuheben ist, dass etwa 38 Prozent der Befragten den Nutzen von GovData.de als hoch oder sehr hoch bewerteten. Auch die Bewertung der einzelnen Aspekte von GovData.de fiel in der Tendenz positiv aus. Etwa die Hälfte der Teilnehmenden beurteilten die Funktionalität und die Benutzerfreundlichkeit von GovData.de als sehr gut oder gut. Das Datenangebot wurde insgesamt negativer bewertet. Insgesamt scheint der Bedarf an Daten bei den Befragten hoch zu sein: So äußerten 88 Prozent den Wunsch nach mehr Daten auf GovData.de. Der Bereich, in dem sich die meisten Teilnehmenden mehr Daten wünschten, war das Themenfeld Umwelt.

3.5 Analyse der Perspektiven verschiedener Nutzergruppen: Ergebnisse der Fokusgruppen und der Auswertung von Nutzerstatements

Um die Perspektiven verschiedener Nutzergruppen im Rahmen der vorliegenden Studie zu erfassen, wurden drei Fokusgruppen mit Nutzerinnen und Nutzern aus der öffentlichen Verwaltung, der Zivilgesellschaft und der Wirtschaft durchgeführt. Ergänzend wurden veröffentlichte Statements von Nutzerinnen und Nutzern und deren Interessensverbänden identifiziert und ausgewertet. Im Folgenden werden die Ergebnisse der Auswertung dieser Fokusgruppen und der betrachteten Statements erläutert. Die Ergebnisse sind nach den in den Fokusgruppen diskutierten Aspekten gegliedert.

3.5.1 Inhaltliche Aspekte

3.5.1.1 In den Fokusgruppen genannte inhaltliche Aspekte

Im ersten Themenblock der Fokusgruppen wurden inhaltliche Aspekte von offenen Daten adressiert. So wurden die Teilnehmerinnen und Teilnehmer aufgefordert zu erzählen, welche Daten sie vorwiegend nutzten und ob diese Daten der gewünschten Qualität entsprächen. In diesem Zusammenhang wurden insbesondere Geodaten, Sozialstrukturdaten und Mobilitätsdaten thematisiert.⁵¹ Bezüglich dieser unterschiedlichen Datenarten wurden jeweils Herausforderungen bei der Nutzung und Qualitätsaspekte diskutiert.

3.5.1.1.1 Geodaten

Geodaten wurden von Teilnehmerinnen und Teilnehmern aller Fokusgruppen genutzt. Die geschilderten Verwertungen dieser Daten waren vielseitig. Beispielsweise wurden von einem Teilnehmer in der Fokusgruppe Wirtschaft digitale Geländemodelle herangezogen, um das optimale Routing für Fahrradfahrende zu identifizieren. Auch für den Zweck der Bebauung oder anstehenden Bebauung in verschiedenen Verwendungszusammenhängen waren Geodaten von Interesse.

⁵¹ In der folgenden Darstellung werden unter die genannten Datenarten jeweils jene Daten gefasst, die die Teilnehmenden selbst den entsprechenden Kategorien zuordneten. Eine theoretische Eingrenzung dieser Kategorien erfolgte nicht.

Für den Abruf von Geodaten wurde unter anderem das Geodatenportal NRW als relevante Plattform erwähnt. Insgesamt variiere die Verfügbarkeit von Geodaten in den Bundesländern: Laut einem Teilnehmer sei beispielsweise Nordrhein-Westfalen sehr gut mit öffentlich zugänglichen Luftbildern abgedeckt, andere Bundesländer wie etwa Rheinland-Pfalz dagegen nicht.

Ein Teilnehmer wies darauf hin, dass die Bereitstellung von Geodaten durch die INSPIRE-Richtlinien auf EU-Ebene geregelt sei. Das Problem sei laut mehrerer Teilnehmerinnen und Teilnehmer jedoch, dass die Daten nicht zwangsläufig kostenfrei zur Verfügung stünden. Die Metadaten müssten aufgrund der INSPIRE-Richtlinien zwar an die EU gemeldet werden, der Datenzugang sei für die Nutzerinnen und Nutzer jedoch häufig mit hohen Kosten verbunden, da das Gesetz den Ländern hier einen Spielraum ließe. In Deutschland würden einige Bundesländer ihre Geodaten freigeben und andere nicht. Einige Teilnehmerinnen und Teilnehmer bemängelten in diesem Zusammenhang, dass ihre Bundesländer keine kostenfreien Geodaten zur Verfügung stellten. Teilnehmerinnen und Teilnehmer der Fokusgruppen wiesen zudem darauf hin, dass das Angebot an Geodaten der Nachfrage nicht gerecht würde. Darüber hinaus wurde angemerkt, dass Nutzerinnen und Nutzern oftmals der Zugang zu den Daten erschwert würde. Man könne sich auf Portalen zwar häufig Geodaten ansehen, aber sobald man diese importieren möchte, um einen eigenen Dienst aufzusetzen, werde es kompliziert. Ein Teilnehmer der Fokusgruppe öffentliche Verwaltung verwies auf ein weiteres Beispiel, das die Nachnutzung von Geodaten vor allem in Bezug auf eine wirtschaftliche Wertschöpfung einschränke: Planungsbüros seien an Geodaten interessiert, könnten für eine kommerzielle Nutzung jedoch nicht kostenfrei auf sie zugreifen. Laut dem Geodateninfrastrukturgesetz seines Bundeslandes (Schleswig-Holstein) müssten Daten nur für eine nicht-kommerzielle Weiterverwendung kostenfrei zur Verfügung gestellt werden.

Ein Teilnehmer bemerkte, dass man in anderen europäischen Ländern schon viel weiter sei bei der Bereitstellung offener Geodaten. Gerade in den nordeuropäischen Ländern seien fast alle vorhandenen Verwaltungsdaten in diesem Bereich Open Data. Zudem würden Geodaten in anderen europäischen Ländern viel mehr genutzt werden und seien auch oftmals viel einheitlicher und vollständiger. Vor allem Estland sei hier sehr fortschrittlich und stelle Daten strukturiert und gebündelt bereit. Der Teilnehmer wies zudem darauf hin, dass es im Bereich Geodaten auf internationaler Ebene zwar schon OpenStreetMap gebe – das sei allerdings etwas anderes als Open Data. Die Daten seien hier auch nur bedingt vollständig: Während Metropolen und Urlaubsregionen sehr gut abgedeckt seien, sei das bei ländlichen Regionen eher nicht der Fall.

3.5.1.1.2 Sozialstrukturdaten

Neben Geodaten wurden in den Fokusgruppen auch Sozialstrukturdaten thematisiert. Vor allem für einen Mitarbeiter eines Forschungsinstituts waren solche Daten für wissenschaftliche Untersuchungen relevant. Von Interesse seien für ihn beispielweise Bildungs- bzw. Schul- und Hochschulstatistiken des statistischen Bundesamtes. Eine Teilnehmerin aus einer Unfallversicherung gab an, dass Sozialdaten für die Präventionsarbeit von Unfallversicherungen bedeutsam sein könnten. Darüber hinaus wurden von den Teilnehmerinnen und Teilnehmern Zensusdaten angesprochen. Hierbei stellte ein Teilnehmer die Frage, wann diese wieder erhoben werden sollen, da bislang nur Zensusdaten von 2011 vorlägen.⁵²

⁵² Die alle zehn Jahre stattfindende Zenserhebung war zuletzt für das Jahr 2021 angesetzt, wurde aber aufgrund der Covid-19-Pandemie auf das Jahr 2022 verschoben, vgl. https://www.zensus2022.de/DE/Was-ist-der-Zensus/_inhalt.html (Abrufdatum: 25. Februar 2022).

Als Quelle für Sozialstrukturdaten wurde von einem Teilnehmer die Datenbank Indikatoren und Karten zur Raum- und Stadtentwicklung (INKAR-Datenbank) hervorgehoben, in der unter anderem das Bundesamt für Bauwesen und Raumordnung Daten bereitstelle. Der Mitarbeiter des Forschungsinstituts erläuterte, dass in der INKAR-Datenbank Daten bis 2017 auf Kreisebene mit Indikatoren für alle möglichen Bereiche vorlägen. Der Datensatz, den er für seine wissenschaftliche Untersuchung heranziehe, sei sehr gut, da er alle Kreise in Deutschland abdecke und keine fehlenden Werte vorlägen; allerdings vermisse er teilweise aktuellere Daten. Ein weiterer Teilnehmer nannte als Quelle Destatis.

In Bezug auf Bildungsstatistiken bemängelte ein Teilnehmer, dass zentrale Angaben fehlten, wie beispielsweise der Migrationshintergrund von Schülerinnen und Schülern. Es werde hier nur zwischen deutschen und ausländischen Schülerinnen und Schülern unterschieden. Einen Mangel gebe es laut einem anderen Teilnehmer darüber hinaus an „hoch aufgelösten“ Arbeitsplatzdaten. So gebe es wenige offene Daten zur Verortung von Arbeitsplätzen und der Frage, wohin gependelt wird. Solche Daten gebe es lediglich auf Gemeindeebene, aber nicht höher aufgelöst. Baden-Württemberg habe solche Daten wiederum auf Kommunalebene, aber das wäre für größere Städte nicht detailliert genug.

Kritisiert wurde, wie bereits bei den Geodaten, dass andere europäische Länder, insbesondere die skandinavischen Länder, weiter seien als Deutschland. In skandinavischen Ländern könne man beispielweise Sozialversicherungsdaten und Erwerbsbiographien zusammentragen, um daraus Erwerbsmuster abzuleiten. Während dort dementsprechend sämtliche Sozialversicherungsdaten in Form einer Vollerhebung öffentlich vorlägen, gebe es in Deutschland in diesem Zusammenhang nur Daten aus Befragungen.

3.5.1.1.3 Mobilitätsdaten

Weitere Daten, die in den Fokusgruppen Erwähnung fanden, waren Mobilitätsdaten. Ein Teilnehmer aus der Fokusgruppe Wirtschaft äußerte in diesem Zusammenhang beispielweise einen Bedarf an Daten (auch Echtzeitdaten) vom öffentlichen Personennahverkehr (ÖPNV). Eine Teilnehmerin aus der Branche der Unfallversicherung hielt Verkehrsdaten für interessant, um Präventionsmaßnahmen zu entwickeln. Ein Teilnehmer wies in einem ähnlichen Zusammenhang darauf hin, dass er für sein Start-Up-Unternehmen im Mobilitätsbereich viel mit Daten zu Fahrradunfällen von Destatis arbeite. Hier werden deutschlandweit Unfallschwerpunkte gesucht und nach verschiedenen Unfallarten geclustert, um zu schauen, wo unfallträchtige Stellen seien. Das helfe auch Fahrradinitiativen und Städten in ihrer Arbeit, da Problemstellen identifiziert und entschärft werden könnten.

Bei offenen Daten im Bereich der Mobilität sei laut einem Teilnehmer für manche Städte eine langsame Verbesserung feststellbar: So seien in den letzten drei Monaten für seine Stadt neue Daten von Verkehrszählungen für die letzten drei Jahre veröffentlicht worden. Diese Daten reichten bis Mai 2021 und würden Auskunft über die Verkehrsauslastung pro Monat und sogar pro Stunde geben. Davor gab es laut dem Teilnehmer nur Daten zur Verkehrsauslastung von Hauptstraßen. Auch die Messungen hätten nur an ausgewählten Tagen stattgefunden. Darüber hinaus wurde von einem Teilnehmer auf das Open Transport Meetup verwiesen, in dem sich virtuell regelmäßig bundesweit Vertreterinnen und Vertreter der Zivilgesellschaft, Forschung, Verwaltung und Verkehrsunternehmen trafen, um sich zu unterschiedlichen Themen auszutauschen. Verkehrsbetriebe würden hier beispielweise einen Einblick geben, wie es mit der Datenverarbeitung läuft. Er sei glücklich, dass in diesem Kontext ein Forum entstanden ist, in dem ein „toller Austausch“ stattfinde.

Von einem Teilnehmer wurde bemängelt, dass für ÖPNV-Betriebe keine Bereitstellungspflicht von Echtzeitdaten bestünde, weil auf Lobbyinteressen von Verkehrsverbänden Rücksicht genommen

worden sei. Dabei seien Echtzeitdaten für die Entwicklung von Mobilitätsangeboten entscheidend – hier habe man eine echte Chance verpasst.

3.5.1.1.4 Sonstige Daten

Laut einer Teilnehmerin der öffentlichen Verwaltung gebe es einen hohen Bedarf nach Kulturdaten. Für die Veröffentlichung von Kulturdaten sei man jedoch auf die Kulturinstitutionen angewiesen und hier müsse man viel Überzeugungsarbeit leisten, um die Einrichtungen von der Veröffentlichung von Daten zu überzeugen. Einige Institutionen habe man bislang schon überzeugen können, Daten bereitzustellen (z. B. Fotografien von Exponaten, Audiodaten oder Videodateien) und man hoffe, dass weitere Einrichtungen folgen. Zudem sei der Hackathon für offene Kulturdaten „Coding da Vinci“ hervorzuheben. Ein weiteres vielversprechendes Projekt für die Bereitstellung von Kulturdaten sei d:kult in Düsseldorf, das in einem E-Museum Daten veröffentliche. Nach und nach werden diese Daten nun auch als offene Daten zur Verfügung gestellt; darauf könne man gut aufbauen. Auch aus anderen Städten habe es bereits Interesse an dem Projekt als Vorlage für weitere Projekte gegeben.

Einige Teilnehmerinnen und Teilnehmer der Fokusgruppen hoben zudem den besonderen Wert von Wetterdaten hervor. Laut einer Teilnehmerin aus der Energiebranche sei es für Energieunternehmen beispielweise nicht nur wichtig sehen zu können, wo Dachflächen für Solaranlagen verfügbar sind, sondern sie müssten ebenso durch Wetterdaten erkennen können, wo eine ausreichende Sonneneinstrahlung gegeben ist. Darüber hinaus seien Wetterdaten auch für die Bereiche Wartung und Instandhaltung wichtig. Eine andere Teilnehmerin erzählte von einer App, die auf Basis von Regendaten entwickelt wurde. Laut einem Teilnehmer würden insbesondere mit Blick auf den zunehmenden Starkregen offene Daten zum Thema Umwelt und Klima fehlen.

3.5.1.2 In den Statements genannte inhaltliche Aspekte

In den betrachteten Statements fanden sich kaum Äußerungen zu unterschiedlichen inhaltlichen Aspekten offener Daten. Dabei ist zu beachten, dass die Auswahl der Statements anhand des Bezuges zu offenen Verwaltungsdaten erfolgte. Eine Suche nach Statements zu unterschiedlichen Datenarten z. B. Geodaten, Sozialstrukturdaten etc. erfolgte hingegen nicht.

In den Statements zum Gesetzesentwurf des Zweiten Open-Data-Gesetzes wurde aus Nutzersicht die Wichtigkeit verschiedener Datenarten hervorgehoben. Andererseits fanden sich auch Äußerungen, die sich gegen eine Veröffentlichung bestimmter Daten aussprachen. Diese Einwände wurden, wie zu erwarten, vor allem aus Sicht der Bereitstellerinnen und Bereitsteller geäußert,⁵³ die bei der Auswertung der Statements nicht näher betrachtet wurden. Allerdings fanden sich auch einige Argumente, die eine Bereitstellung bestimmter Daten mit Verweis auf mögliche negative Auswirkungen auf die Gesamtgesellschaft oder bestimmte gesellschaftliche Gruppen ablehnten. Diese Gegenargumente wurden in der Betrachtung berücksichtigt.

⁵³ Als Einschätzung aus Sicht der bereitstellenden Behörden wurden Argumente gewertet, die sich auf den mit der Bereitstellung verbundenen Aufwand bezogen oder die Wichtigkeit der Wahrung der Rechte der Datenbereiterinnen und -bereitstellern anführten.

3.5.1.2.1 Mobilitätsdaten und Daten der Daseinsvorsorge

Ein betrachtetes Gutachten, welches im Rahmen der Begleitforschung zur Innovationsinitiative mFund des Bundesministeriums für Digitales und Verkehr entstanden ist, nennt die Beschaffung von Daten als zentrale Herausforderung für Projekte im Mobilitätsbereich. Trotz regulatorischer Verbesserungen seien immer noch zu wenige relevante Daten frei verfügbar.⁵⁴ Als besonders relevante Datenarten für Mobilitätsprojekte nennt das Gutachten ein vielfältiges Spektrum an Daten aus verschiedenen Bereichen: „Verkehrs-, Klima- und Wetterdaten, Pegelstände, Fernerkundungsdaten, Daten aus dem Bereich der bzw. zur Luft- und Raumfahrt, amtliche Geobasisdaten, Bewegungsdaten, Bevölkerungsdaten, Verkehrs- und Fahrplandaten einschließlich Echtzeitinformationen und Daten zur Infrastruktur.“⁵⁵ .

In der Entwurfsfassung des Zweiten Open-Data-Gesetzes wird der Anspruch zur Veröffentlichung von Daten auch auf öffentliche Unternehmen im Bereich der Daseinsvorsorge (und damit auch des öffentlichen Nahverkehrs) ausgeweitet.⁵⁶ Ein mehrfach gegen diese Ausweitung vorgebrachter Einwand ist, dass es zu einer Wettbewerbsverzerrung zwischen öffentlichen Unternehmen und privatwirtschaftlichen Unternehmen, die im selben Sektor tätig sind, kommen könne, zum Nachteil der öffentlichen Unternehmen.⁵⁷ Wenn beispielsweise öffentliche Unternehmen im öffentlichen Nahverkehr dazu verpflichtet würden, Daten zur Auslastung ihrer Verkehrsmittel zu veröffentlichen, könnten konkurrierende private Anbieter bei ihrem Streckenangebot gezielt weniger profitable Bereiche aussparen. Dies könne die Qualität des öffentlichen Nahverkehrs bzw. der Daseinsvorsorge beeinträchtigen, so beispielsweise die Einschätzung der Gewerkschaft Verdi.⁵⁸ Mehrfach wurde dabei auch die Befürchtung geäußert, es würden hauptsächlich internationale Unternehmen von der Datenbereitstellung profitieren.⁵⁹ Die Bitkom begrüßte zwar grundsätzlich die Bereitstellung offener Daten im Hinblick auf die Daseinsvorsorge, wies aber ebenfalls auf die Möglichkeit von Wettbewerbsverzerrungen hin.⁶⁰

Als mögliche Lösung dieses Problems wurde daher eine Ausweitung des Datenbereitstellungsanspruches auf private Unternehmen im Bereich der öffentlichen Daseinsvorsorge vorgeschlagen.⁶¹

54 Vgl. mFUND (2019), S.2.

55 mFUND (2019), S.3.

56 Vgl. Bundesregierung (2021b), S.8

57 Vgl. Verband kommunaler Unternehmen (2021), S.4; Verdi (2021), S.1.

58 Vgl. Verdi (2021), S.2f.

59 Vgl. Verband kommunaler Unternehmen (2021), S.13; Smart City Dialog (2020).

60 Vgl. Bitkom (2021), S.4.

61 Vgl. Max-Planck-Institut für Innovation und Wettbewerb (2021), S.4; OKF (2021), S.4.

3.5.1.2.2 Gesundheitsdaten

Im Bitkom Statement zur Entwurfsfassung des Zweiten Open-Data-Gesetz wird zudem auf den wirtschaftlichen und gesamtgesellschaftlichen Nutzen von offenen Daten im Gesundheitsbereich in Hinblick auf die Covid-19-Pandemie hingewiesen: „Die [Covid-19] Krise zeigt auch, dass der datenschutzkonformen Bereitstellung und Nutzung von Daten aus dem Bereich der Gesundheitsversorgung eine hohe gesellschaftliche Bedeutung zukommt. Wenn entsprechende (Echtzeit-) Daten (z. B. Bettenkapazitäten, Infektionszahlen, Terminanfragen) in anonymisierter Form Dritten zugänglich gemacht werden, können dadurch Innovationen befördert werden, welche die Reaktionsfähigkeit staatlicher und privater Entscheidungsträger verbessern. Politik, Forschung und Wirtschaft werden dadurch auch auf zukünftige Krisensituationen besser vorbereitet sein.“⁶². Auch der Medienkonzern Relx und die Konrad-Adenauer-Stiftung begrüßten die Aufnahme von Gesundheitsdaten in den Regelungsbereich des Zweiten Open-Data-Gesetzes (in der Entwurfsfassung).⁶³ Diese sah im Gesundheitsbereich auch die Bereitstellung von offenen Daten durch private Einrichtungen vor.

Die Open Knowledge Foundation Deutschland e. V. (OKF) begrüßte zwar grundsätzlich die Berücksichtigung von Gesundheitsdaten, betonte aber die hohe Wichtigkeit des Datenschutzes in diesem Bereich, dieser müsse im gesamten Prozess der Datenbereitstellung gewahrt bleiben.⁶⁴ Auch die Bitkom beschrieb die besondere Wichtigkeit der Anonymisierung von Gesundheitsdaten, zudem forderte die Organisation die Ausweitung des Datenbereitstellungsanspruchs auf die Selbstverwaltungskörperschaften im Gesundheitsbereich, insbesondere auf die Krankenkassen.⁶⁵

Der betrachtete Artikel von Netzpolitik.org äußerte sich eher skeptisch zur Aufnahme von Gesundheitsdaten. Die Aufnahme dieser Daten mit der Begründung der besonderen gesellschaftlichen Relevanz sei zwar vor dem Hintergrund der Covid-19-Pandemie nachvollziehbar, wirke aber planlos.⁶⁶ Vor allem sei unklar, welche Daten welcher Einrichtungen genau gemeint seien. Zudem wurde in dem Beitrag argumentiert, dass neben den aktuell durch die Pandemie besonders im Fokus stehenden Gesundheitsdaten weitere Daten aus anderen Bereichen eine vergleichbar hohe gesellschaftliche Relevanz aufweisen und daher aus Sicht der Autorin ebenfalls veröffentlicht werden sollten, genannt wurden beispielhaft Daten aus dem Bildungsbereich und finanzmarktbezogene Daten.⁶⁷

3.5.1.2.3 Wissenschaftliche Basisdaten

Relx betonte in seinem Statement die Bedeutung der Veröffentlichung von Forschungsdaten: „Die gemeinsame Nutzung von Forschungsdaten kann die Wirkung, Validität, Reproduzierbarkeit, Effizienz und Transparenz der wissenschaftlichen Forschung erhöhen. Die gemeinsame Nutzung von Daten ermöglicht es anderen, die Ergebnisse von Experimenten wiederzuverwenden und unterstützt die

62 Bitkom (2021), S.4.

63 Vgl. RELX (2021), S.2; Konrad-Adenauer-Stiftung (2021), S.1.

64 Vgl. OKF (2021), S.4.

65 Vgl. Bitkom (2021), S.3.

66 Vgl. Litta (2021).

67 Vgl. Litta (2021).

Schaffung neuer Wissenschaft, die auf früheren Erkenntnissen aufbaut, wodurch der Forschungsprozess effizienter wird.“⁶⁸

Mehrere Organisationen aus der Wissenschaft wiesen im Hinblick auf die PSI-Richtlinie und das Zweite Open-Data-Gesetz darauf hin, dass die Nutzerinnen und Nutzer von Daten zu wissenschaftlichen Zwecken andere Anforderungen an die geteilten Daten stellen als andere Nutzergruppen. Auch seien in der Wissenschaft bereits vielfältige Bemühungen erfolgt, Daten diesen Ansprüchen entsprechend zu veröffentlichen und auszutauschen.⁶⁹ Der Rat für Sozial- und Wirtschaftsdaten äußerte in seinem Statement zum Zweiten Open-Data-Gesetz die Befürchtung, die bisherige Praxis der Bereitstellung von Daten in der Wissenschaft könne durch die als „größer“ wahrgenommene Formen der Bereitstellung im Sinne der Open-Data-Prinzipien ersetzt werden.⁷⁰

3.5.1.2.4 Sonstige Daten

Neben den bisher dargestellten Dateninhalten, auf die das Zweite Open-Data-Gesetz Bezug nimmt, wurde eine Ausweitung auf Daten aus dem Justizbereich sowie Daten im Bildungsbereich gefordert.

Aus Sicht von Open Legal Data sollten auch die Urteile der Bundesgerichte als offene Daten veröffentlicht werden. Dies könne die Transparenz staatlichen Handelns weiter verbessern: „Für eine Bereitstellung von Urteilen als Open Data spricht insbesondere, dass Urteile im Namen des Volkes ergehen und das Demokratiegebot es erfordert, dass die Bürger*innen jederzeit in Erfahrung bringen müssen, welche Rechte und Pflichten sich aus einem Gesetz ergeben.“⁷¹

Das deutsche Jugendinstitut forderte hingegen die Aufnahme von Daten zur Schulbildung in den Anwendungsbereich des zweiten Open-Data-Gesetzes. Es sei nicht nachvollziehbar, warum gerade dieser Bereich nicht berücksichtigt worden sei: „Im Rahmen der Schulpflicht erfüllt das Schulsystem wichtige öffentliche Aufgaben, von deren transparenterer Darstellung und Dokumentation Forschung, Bildungsadministration und -praxis enorm profitieren würden.“⁷² Auch in dem betrachteten Artikel von Netzpolitik.org wurden Daten aus dem Bildungsbereich als gesellschaftlich relevant eingestuft.⁷³

3.5.2 Nutzungsaspekte

3.5.2.1 In den Fokusgruppen angesprochene Nutzungsaspekte

3.5.2.1.1 Probleme bei der Nutzung offener Daten

Die Teilnehmerinnen und Teilnehmer der Fokusgruppen wiesen auf unterschiedliche Probleme hin, die sie bei der Nutzung offener Daten erlebt haben. Ein grundlegendes Problem, das von mehreren

68 RELX (2021), S.3.

69 Vgl. Deutsches Jugendinstitut (2021), S.1; Rat für Sozial- und Wirtschaftsdaten (2021), S.1; Allianz der Wissenschaftsorganisationen (2018), S.1.

70 Vgl. Rat für Sozial- und Wirtschaftsdaten (2021), S.1.

71 Open Legal Data (2021), S.2.

72 Deutsches Jugendinstitut (2021), S.4.

73 Vgl. Litta (2021).

Teilnehmerinnen und Teilnehmern genannt wurde, war die schwierige Auffindbarkeit von Daten und die Unbekanntheit von Open-Data-Portalen. Eine Teilnehmerin erzählte, dass sie die unterschiedlichen Portale, auf denen offene Daten hinterlegt werden, lange Zeit nicht kannte. Auch ein Mitarbeiter eines Forschungsinstituts merkte an, dass das Angebot an offenen Daten seinen Kolleginnen und Kollegen kaum bekannt sei und wenig genutzt würde. Vor diesem Hintergrund regte er an, gezielte Angebote in Forschungsinstituten und Universitäten zu forcieren, um die Angebote offener Daten in Deutschland bekannter zu machen. So könne die Verwendung offener Daten beispielweise Teil von Statistikvorlesungen werden. Er selbst habe mehrere Jahre studiert und promoviert, bis er auf das Angebot an offenen Daten aufmerksam wurde. Schwer auffindbar seien offene Daten auch laut eines Teilnehmers aus der Zivilgesellschaft, der sich die Daten für seine App laut seiner Aussage aus unterschiedlichen Quellen mühsam zusammensuchen musste. Teilweise seien Daten auf den Seiten von Institutionen sehr schwer auffindbar oder nur auf Nachfrage verfügbar. Kritisiert wurde zudem, dass für die Nutzung mancher Portale eine Registrierung erforderlich sei. Es sei laut einem Teilnehmer zwar verständlich, dass diejenigen, die die Portale betreiben, sehen möchten, wer die offenen Daten nutzt – der Zugang zu den Daten sei dann jedoch nicht mehr als offen zu bewerten.

Die mangelnde Aktualität vieler Daten wurde als weiteres Problem thematisiert. Im Begutachtungsprozess wissenschaftlicher Studien werde dies von Fachzeitschriften gerne bemängelt und gefragt, warum keine neueren Verwaltungsdaten vorlägen. Ein Teilnehmer betonte, dass vor allem Daten zum Verkehr sehr aktuell sein müssten: Wenn man wissen wolle, wie sich der Verkehr in einer Stadt entwickle, könne man nicht auf Daten aus dem Jahr 2014 zurückgreifen. Dafür wachse eine Stadt wie zum Beispiel Berlin viel zu schnell. Ein anderer Teilnehmer merkte an, dass bei Verkehrsdaten sogar oftmals Echtzeitdaten notwendig seien. In manchen Anwendungen sei es zudem so, dass sich verschiedene Datensätze aufeinander bezögen – wenn in solchen Fällen einer der Datensätze erst verspätet bereitgestellt wird, beträfe das auch die Verwertbarkeit der anderen Daten.

Weiterhin wurde kritisiert, dass es selten Feedbackmöglichkeiten zu Datensätzen gebe. Generell wünschte sich einer der Teilnehmenden mehr Austausch zwischen den Datenbereiterinnen und -bereitstellern und den Nutzerinnen und Nutzern von offenen Verwaltungsdaten z. B. zur Automatisierung und zu Datenformaten. Er regte dabei eine Orientierung an der Feedbackkultur im Open Source Bereich an, wo Nutzerinnen und Nutzer häufig direkt mit den Bereitstellerinnen und Bereitstellern von Dienstleistungen über Portale wie GitHub kommunizieren könnten. Im Open Data Bereich gebe es hingegen selten eine praktikable Feedbackfunktion zu den bereitgestellten Daten. Ein anderer Teilnehmer erwiderte, es sei nicht ungewöhnlich, dass Nutzerinnen und Nutzer offene Daten kommentieren könnten. Allerdings habe die Stadt Berlin auf ihrem Portal bspw. ein solches Kommentarsystem, ohne dass jemand davon wisse, da dafür nicht geworben würde. Seiner Meinung nach haben Portale so etwas häufig, aber niemand nutze diese Möglichkeit. Ein anderer Teilnehmer erzählte von einem Vorfall, bei dem er wegen häufiger Abrufe eines Datensatzes gesperrt wurde, aber keine Kontaktdaten zum Datenbereiterinnen und -bereitstellern aus der Verwaltung angegeben waren. Erst über Twitter und mit der Hilfe von Journalisten sei er an den Kontakt gekommen und konnte entsperrt werden.

Vor dem Hintergrund der derzeit mangelnden Kommunikation zwischen der Verwaltung und den Nutzerinnen und Nutzern offener Verwaltungsdaten forderte ein Teilnehmer aus der Zivilgesellschaft stärker aufeinander zuzugehen und mehr Empathie für einander zu zeigen. Auf der Seite der Zivilgesellschaft fehle häufig das Verständnis dafür, warum von der Verwaltung bestimmte Daten nicht bereitgestellt würden. Hier würde es helfen, wenn die Zivilgesellschaft die Gründe kennen würde, wie bspw. ein Gesetz, das die Veröffentlichung verbietet. Die Angebots- und Nutzerseite müssten stärker zusammengebracht werden; die Zivilgesellschaft müsse die Einschränkungen der Verwaltung besser

verstehen. Die Verwaltung solle eine Veröffentlichung nicht ausschließen, weil davon ausgegangen wird, dass niemand etwas mit den Daten anfangen könne. Gleichzeitig solle die Zivilgesellschaft nicht pauschal denken, dass „die blöde Verwaltung“ die Daten einfach nicht bereitstellen wolle. Darüber hinaus vermutete er, dass den Datenbereiterinnen und -bereitstellern oft nicht bewusst sei, dass Nutzerinnen und Nutzer mit den Daten funktionierende Anwendungen bereitstellen möchten und dies wieder aufgegeben werden müsste, wenn Daten nicht zur Verfügung gestellt würden. Es müsse zudem darauf vertraut werden können, dass die Verwaltung die benötigten Daten auch in der Zukunft zuverlässig bereitstellt, damit sich die Entwicklung derartiger Anwendungen lohnt. Ein Teilnehmer bemerkte hierzu, dass der Austausch zwischen Zivilgesellschaft und Verwaltung durch Hackathons nicht reichen würde, da es sich hierbei nicht um repräsentative Nutzerinnen und Nutzer handeln würde.

Ein weiteres Problem sahen die Teilnehmerinnen und Teilnehmer in Bezug auf fehlende Daten. Laut eines Teilnehmers sei die Flächenabdeckung von Datensätzen sehr ungleichmäßig; teilweise würden große Lücken bestehen. Auch ein anderer Teilnehmer bemerkte, dass feinräumige Daten oftmals nur auf regionaler Ebene vorhanden seien. Es sei ein Problem, wenn z. B. von Nordrhein-Westfalen Luftbilder vorlägen, aber die Luftbilder eines anderen Bundeslandes 200.000 Euro kosten würden. Auch fehlende Informationen innerhalb von Datensätzen kämen vor: Eine Teilnehmerin berichtete von einem Projekt, bei dem auf offene Daten zu Bäumen in der Stadt zurückgegriffen wurde und in einigen Fällen Bäume gefehlt hätten.

Ein Teilnehmer erzählte des Weiteren, dass er bei der Recherche nach Gesundheitsdaten teilweise auf sehr alte Datensätze gestoßen sei, die nicht mehr weitergeführt wurden. Kommunen oder Ländern hätten hier angefangen, Daten maschinenlesbar online zur Verfügung zu stellen und dann wieder aufgegeben. In diesem Zusammenhang betonte ein weiterer Teilnehmer, dass die Kontinuität und Verlässlichkeit von offenen Daten zentrale Punkte seien. Ohne eine Verlässlichkeit der Fortschreibung von offenen Daten könne man Start-Ups an den Rand des Zusammenbruchs bringen. Eine weitere Person vermutete zudem, dass das statistische Landesamt seines Bundeslandes viel mehr Daten habe, als es online zur Verfügung stelle. Dabei sei es wichtig, mehr Details zu haben, um Bevölkerungsentwicklungen in einzelnen Stadtteilen nachvollziehen zu können. Ebenso kritisiert wurde, dass nicht immer Rohdaten bereitgestellt würden, sondern aufbereitete und plausibilisierte Daten, wodurch fünf bis zehn Prozent der Daten verloren gingen. Es sei wichtig, dass der Zivilgesellschaft die Rohdaten nicht vorenthalten würden, weil die Verwaltung denke, die Zivilgesellschaft könne damit nicht umgehen.

Von einer Teilnehmerin wurde zudem eine mangelnde Barrierefreiheit von Open-Data-Angeboten angesprochen. Für Menschen mit Sehbehinderung sei es beispielweise wichtig, dass alles verschriftlicht und nicht nur graphisch aufbereitet wird, damit eine Nutzung mit dem Screen Reader möglich ist. Um die Nutzung offener Daten für viele Menschen zu ermöglichen, sollte auch über Angebote in leichter Sprache nachgedacht werden. Ein Teilnehmer erzählte, dass es sich für ihn sehr schwierig gestaltet habe, sich das Wissen anzueignen, wie man offene Daten barrierefrei aufbereitet. Hierzu gebe es wenige frei zugängliche Informationen; er selbst habe dafür dann zwei sehr technisch affine Programmierer mit Sehbehinderung kontaktiert.⁷⁴

74 Ein Teilnehmer wies zudem auf die Seite <https://github.com/transportkollektiv/opendata-antipatterns> hin, auf der viele Probleme hinsichtlich der Nutzung offener Daten bereits gesammelt worden sind.

3.5.2.1.2 Herausforderung für Datenbereiterinnen und Datenbereiter

In den Fokusgruppen wurde auch über Herausforderungen diskutiert, denen die öffentliche Verwaltung bei der Bereitstellung offener Daten begegnet. In der Fokusgruppe öffentliche Verwaltung hatten mehrere Teilnehmende bereits Erfahrung in der Verwendung und der Bereitstellung von offenen Daten und konnten somit sowohl aus Perspektive von Nutzerinnen und Nutzern als auch Bereitstellerinnen und Bereitstellern berichten. Auch in der Fokusgruppe Zivilgesellschaft fand sich eine Person, die in der öffentlichen Verwaltung tätig war, wodurch eine Diskussion über das Verhältnis zwischen Datenbereiterinnen und -bereitstellern und Nutzerinnen und Nutzer zustande kam.

Eine zentrale Herausforderung, die in diesem Zusammenhang thematisiert wurde, war der Mangel an Ressourcen. So wurde von den Teilnehmerinnen und Teilnehmern darauf hingewiesen, dass der Verwaltung sowohl die finanziellen als auch die personellen Ressourcen fehlten, um offene Daten in hinreichender Form bereitzustellen. Laut einem Teilnehmer seien beispielsweise die Kommunen damit überfordert, die INSPIRE-Richtlinien zu erfüllen und Geodaten bereitzustellen. Es handele sich um eine Vorschrift, an die sich die wenigsten halten würden. Um die Richtlinien umsetzen zu können, müssten die Kommunen eine bessere finanzielle Unterstützung erhalten. Ein Teilnehmer aus der Zivilgesellschaft berichtete zudem, dass er im Kontakt mit Datenbereiterinnen und Datenbereitstellern erfahren habe, dass die Aufbereitung der Daten sehr aufwendig sei und in diesem Zusammenhang ein Personalproblem vorläge. Auch laut einer anderen Teilnehmerin fehlten in der Verwaltung IT-Verantwortliche, die über das notwendige Wissen verfügten, um Daten zu bereinigen und öffentlich zugänglich machen zu können. Darüber hinaus wurde angemerkt, dass die technische Infrastruktur in der Verwaltung oftmals nicht vorhanden sei um bspw. Daten in Echtzeit zu veröffentlichen.

Ein mangelndes Wissen in der Verwaltung bezüglich Open Data wurde von Teilnehmerinnen und Teilnehmern in weiteren Kontexten beklagt. So fehlte laut einer Teilnehmerin kommunalen Einrichtungen oftmals das Gespür für Daten: Vor allem wenn es um Kulturdaten ginge, müsste Einrichtungen zunächst vermittelt werden, was Daten sind und wie man sie digitalisiert. Ein anderer Teilnehmer bemerkte, dass die Mitarbeiterinnen und Mitarbeiter in der Verwaltung offenbar nicht wüssten, was für die technische Seite wichtig ist. Die Namen von Variablen müssten beispielweise immer gleichbleiben und dürften sich nicht bei der Fortschreibung von Datensätzen ändern. Eine Teilnehmerin plädierte dafür, Open Data in den Behörden zunächst einmal bekannter zu machen. Dazu seien Informationsveranstaltungen und Seminare sinnvoll, damit die Mitarbeiterinnen und Mitarbeiter Wissen erwerben können und eine stärkere Sensibilisierung stattfindet. Bei ihrer Behörde würden beispielweise Schulungen angeboten, um Open Data stadintern bekannter zu machen. Zudem sei ein „open-by-default“-Ansatz in den Kommunen sinnvoll.

Ein weiteres Problem sei laut den Teilnehmerinnen und Teilnehmern, dass den Mitarbeiterinnen und Mitarbeitern der öffentlichen Verwaltung die Nutzergruppen von offenen Verwaltungsdaten oftmals nicht bekannt seien. In diesem Zusammenhang erzählte eine Teilnehmerin aus eigener Erfahrung, dass es in der öffentlichen Verwaltung jedoch durchaus den Wunsch gebe zu wissen, wer die Nutzerinnen und Nutzer von bereitgestellten Daten sind. Oft gebe es die Vorstellung, dass es sich um Hacker, Trolle oder Programmierer handle, für die man „einfach irgendwelche Daten“ online stellen müsse, aus denen die dann „coole Sachen“ machen. Der Verwaltung sei oftmals nicht bewusst, dass es bei den Nutzerinnen und Nutzern offener Verwaltungsdaten eine große Vielfalt gebe. Eine andere Teilnehmerin bemerkte, dass Daten teilweise auch nicht öffentlich bereitgestellt werden, weil man sie für zu kompliziert halte und davon ausgehe, dass niemand etwas mit den Daten anfangen könne. Auch

eine weitere Teilnehmerin bestätigte, es sei eine riesige kommunikative Aufgabe, in der Verwaltung intern zu vermitteln, wie divers die Nutzerinnen und Nutzer sind.

Eine Teilnehmerin aus der öffentlichen Verwaltung stellte sich vor diesem Hintergrund die Frage, wie ein Austausch zwischen den Bereitstellerinnen und Bereitstellern von Daten und den Nutzerinnen und Nutzern geschaffen werden könne. Es gebe keine etablierten Kanäle für Feedback, daher wisse die Verwaltung oft nicht, welche Bedarfe bei der Nutzung offener Daten vorliegen. Somit entstehe der Eindruck, dass man Daten „in ein schwarzes Loch werfe“: Man wisse nicht was passiert, weil niemand seine Datennutzung melden müsse und man daher keine Rückmeldung bekäme. Daher fehle auch das Verständnis dafür, wozu und wieso man die Daten bereitstellt und ob es sich lohnte. Ein Teilnehmer vermutete, dass die Verwaltung offene Daten eher als Grundlage für die Forschung sehe und weniger als Grundlage für tagesaktuelle Anwendungen, für die Echtzeitdaten benötigt werden. Die Verwaltung denke außerdem oft nicht daran, wie wertvoll diese Daten sein können. Auch eine weitere Teilnehmerin hielt es für notwendig, das Verständnis in der Verwaltung dafür zu verbessern, was offene Verwaltungsdaten in ihrer Anwendung für Schätze seien.

Thematisiert wurden darüber hinaus die mangelnden Anreize für die öffentliche Verwaltung, offene Daten bereitzustellen. Möglich sei es laut einer Teilnehmerin, Mitarbeiterinnen und Mitarbeitern der öffentlichen Verwaltung zur Bereitstellung offener Verwaltungsdaten zu überzeugen, in dem man den Musterdatenkatalog für Kommunen der Bertelsmann Stiftung verwendet. Da könne man sehen, welche Daten interessant und bereits veröffentlicht worden sind. Man könne somit argumentieren: Andere machen das auch! Denn laut einiger Teilnehmerinnen und Teilnehmer gehöre die Bereitstellung offener Daten nicht zu den Prioritätsaufgaben der öffentlichen Verwaltung. In diesem Zusammenhang erzählte eine Teilnehmerin, dass die Daten oftmals händisch eingetragen werden müssten und man diese Aufgabe gerne nach hinten verschiebe. Manche Behörden würden die Daten aus Bequemlichkeit nicht bereitstellen; es läge zudem sowohl ein Generationen- als auch teilweise ein Verantwortlichkeitsproblem vor.

Darüber hinaus vermutete ein Teilnehmer, dass Daten teilweise nicht öffentlich zugänglich gemacht werden, da sich die Behörden angreifbar machen würden, falls bei den Daten Fehler vorliegen oder sie nicht vollständig sind. Um den Behörden diese Sorge zu nehmen, könnte man beispielweise qualifizierte Stellen mit geschultem Personal einrichten. Ein weiterer möglicher Grund für die zögerliche Bereitstellung sei laut diesem Teilnehmer auch, dass einige Behörden die erstellten Daten erst selbst auswerten möchten, bevor sie der Öffentlichkeit preisgegeben werden. Bis zur Veröffentlichung der Daten könnten dann einige Jahre vergehen. Dadurch sicherten sich die datenhaltenden Stellen die Interpretationshoheit über die Daten. Ein weiteres Hemmnis, Daten öffentlich zugänglich zu machen sei, dass man mit der Veröffentlichung von Daten einen Standard setze, an dem man sich auch weiterhin halten müsse.

Des Weiteren berichteten Teilnehmerinnen und Teilnehmer, die selbst schon Erfahrungen in der Bereitstellung von Daten gesammelt hatten, dass sie Daten teilweise auch aus juristischen Gründen nicht öffentlich bereitstellen dürften. Daten zur Infrastruktur und Versorgung von Gas, Wasser oder Strom seien laut einem Teilnehmer empfindliche Daten, die dem Geheimschutz unterlägen. Wenn jemand einen Anschlag planen würde und dafür auf Wasserdaten zugreifen könne, wäre es leicht, neuralgische Punkte zu identifizieren. Es sei daher oftmals schwierig zu unterscheiden, was man veröffentlichen dürfe und was nicht. Eine weitere Teilnehmerin bemerkte in diesem Zusammenhang, dass man bei der Veröffentlichung von Daten eine große Verantwortung trage. Vor allem aus Sicherheitsgründen sei es oft ein sehr schmaler Grat, welche Daten z. B. zur Infrastruktur veröffentlicht werden können und welche nicht.

3.5.2.1.3 Anwendungsbeispiele offener Daten

Um einen Eindruck davon zu gewinnen, wie offene Verwaltungsdaten in den unterschiedlichen identifizierten Gruppen von Nutzerinnen und Nutzern zum Einsatz kommen und wie dadurch eine wirtschaftliche und soziale Wertschöpfung generiert werden kann, wurden die Teilnehmerinnen und Teilnehmer nach dem Mehrwert offener Daten sowie nach konkreten Anwendungen gefragt, von denen einige im Folgenden beispielhaft vorgestellt werden sollen.

3.5.2.1.3.1 Nutzerinnen und Nutzer aus der öffentlichen Verwaltung

Laut einer Teilnehmerin hätten offene Verwaltungsdaten einen großen Mehrwert für die Kommunikation zwischen der öffentlichen Verwaltung und den Bürgerinnen und Bürgern. Die Arbeit der Verwaltungsmitarbeiterinnen und -mitarbeiter würde durch die Verfügbarkeit offener Daten vereinfacht: Wenn Anfragen von Bürgerinnen und Bürgern, aus der Wissenschaft oder der Politik kämen, wisse man, worauf man verweisen könne. In ihrer Stadt würde der Nutzen von offenen Verwaltungsdaten allerdings noch nicht systematisch genutzt. Alle Dezernate schrieben zwar ihre Berichte, aber die Daten würden noch nicht in die Portale wandern. Wenn für Berichte Daten koordiniert gesammelt werden sollen, müssten im Einzelnen die jeweiligen Stellen angeschrieben werden, um die Daten verfügbar zu machen. Der Nutzen für verschiedene Fachabteilungen wäre sehr groß, wenn man nicht immer nachfragen müsste, sondern in einem Datenkatalog definiert wäre, welche Daten der Verwaltung vorliegen. Möglich sei es auch, Dashboard-Lösungen zu integrieren, um einen Mehrwert für die Fachabteilungen zu schaffen, weil sie selbst ihre Berichte für die Einwohnerinnen und Einwohner, die diese Informationen benötigen, damit bereichern könnten.⁷⁵

Bezüglich des Nutzens von Dashboards merkte ein anderer Teilnehmer an, dass es in Bundesländern kompliziert sein könne, ein Dashboard zu bauen und die Daten dafür zu bekommen. In Kommunen sei dies ggf. einfacher als in der Landesverwaltung. Der Prozess der Entscheidung, welche Daten ins Dashboard kommen, sowie hinsichtlich der Freigabe der Daten sei schwieriger, je mehr unterschiedliche Akteure an der Bereitstellung beteiligt seien. Wenn die Daten aber schon als offene Verwaltungsdaten zur Verfügung stehen, sei diese Entscheidung schon gefallen. Dann müsse der Transport ins Dashboard auch nicht mehr so gesichert werden und viele Prozesse seien deutlich verkürzt, dies sei häufig auch mit Kosteneinsparungen verbunden.

Neben den beschriebenen Vorteilen von offenen Verwaltungsdaten für die Bereitstellung von Informationen und Dienstleistungen für die Bürgerinnen und Bürger sahen Teilnehmerinnen und Teilnehmer der Fokusgruppe öffentliche Verwaltung in offenen Daten einen weiteren Mehrwert für die Verwaltung selbst, da durch sie doppelte Arbeit vermieden werden könne. Ein Teilnehmer verwies hierbei auf das Beispiel des Wirtschaftsministeriums seines Bundeslandes, in dem es eine Abteilung für Tourismus gebe, die eine Liste mit Denkmälern führe. Gleichzeitig führe aber auch das Landesamt für Denkmalpflege eine Denkmalliste. Dadurch, dass das Landesamt die Liste nun digital zugänglich mache, könne die Ministeriumsabteilung darauf zugreifen und es entstehe keine doppelte Arbeit. Laut dem Teilnehmer sei es ein grundsätzliches Problem, dass diverse Behörden die gleichen Daten erheben. Die Hoffnung sei es, einen zentralen Datenbestand zu entwickeln, auf den alle zugreifen

⁷⁵ Als Dashboard wird gemeinhin eine interaktive grafische Oberfläche verstanden, in der Daten und darin enthaltene Zusammenhänge dargestellt werden.

können. Eine Teilnehmerin fügte hinzu, dass es auch wichtig sei, Verwaltungen zu befähigen, Daten zu sammeln (auch von Privatunternehmen) und anschließend zu poolen. Auf Basis einer solchen Datengrundlage könnten sich dann neue Ideen entwickeln. Gepoolte Daten könnten beispielsweise Start-ups helfen, die ständig auf der Suche nach neuen Ideen wären, aber nicht einfach Daten einkaufen könnten.

Ein weiterer Teilnehmer hob den Nutzen von offenen Verwaltungsdaten am Beispiel der durch EU-Gelder geförderten Landwirtschaftsbetriebe hervor. Bei der EU-Direktförderung bekämen Landwirtinnen und Landwirte Geld dafür, dass sie bestimmte Dinge auf ihren Feldern anbauten. Um zu überwachen, dass sich die Landwirtinnen und Landwirte an die Auflagen hielten, würden die sogenannten Kopernikus-Daten genutzt, für die die EU Satelliten ins All geschossen habe. Durch die Satellitenaufnahmen, die die EU kostenfrei als Open Data zur Verfügung stelle, würde man viel Zeit und Personal sparen, da man die geförderten Felder zur Kontrolle nicht mehr befliegen oder begehen müsse. Der Vorteil, diese Daten öffentlich zugänglich zu machen sei aber nicht nur, dass die Verwaltung davon profitiert habe, sondern auch, dass sich Forschende damit befassen konnten. Mit den Kopernikus-Daten haben Forscherinnen und Forscher beispielsweise ermittelt, welche Pflanzen in unterschiedlichen Regionen wachsen. Viele Erkenntnisse konnten nur entstehen, weil Studentinnen und Studenten durch offene Daten die Möglichkeit hätten, mit diesen zu experimentieren. Es sei daher sehr gut, dass die EU sich dafür entschieden habe, diese für viel Geld beschafften Daten freizugeben.

Einige Teilnehmerinnen und Teilnehmer der öffentlichen Verwaltung sahen in offenen Daten spezifische Potenziale, die im Rahmen ihrer Arbeit jedoch noch nicht ausgeschöpft würden. Eine im gesetzlichen Versicherungszweig tätige Teilnehmerin äußerte beispielweise Interesse an Unfalldaten, da diese für die Präventionsarbeit von Unfallversicherungen bedeutsam seien. Wenn durch die Daten bspw. deutlich wird, dass Unfälle regional ansteigen, könne man auf dieser Basis Druck machen für sicherere Verkehrswege. Noch kämen solche Daten jedoch nicht zum Einsatz. Eine weitere Teilnehmerin erhoffe sich offene Daten für die bedarfsgerechtere Entwicklung von Apps für Menschen mit Behinderung nutzen zu können.

3.5.2.1.3.2 Nutzerinnen und Nutzer aus der Wirtschaft

Auch von einigen Teilnehmerinnen und Teilnehmern in der Fokusgruppe Wirtschaft wurde berichtet, dass sie derzeit das Potenzial von offenen Daten für ihre Arbeit noch ausloten müssten. Einen Teilnehmer aus der IT-Branche beschäftigte beispielweise, inwieweit es möglich sei, Smart Metering mit Open Data zu verknüpfen bzw. anzureichern. Bei Smart Metering werden Haushalte und Unternehmen mit intelligenten Messsystemen ausgestattet. Es handele sich hierbei um einen wichtigen Bestandteil der Digitalisierung und der Energiewende. Da es sich jedoch um personenbezogene Daten, wie etwa den individuellen Stromverbrauch handele, bestehe aufgrund des Datenschutzes ein Gegensatz zu Open Data. Ein weiteres Potenzial sah der Teilnehmer zudem darin, individuelle Daten von Stromverbraucherinnen und Stromverbrauchern mit Daten anzureichern, die schon öffentlich sind (z. B. Wetterdaten), um mehr Informationen über den eigenen Stromverbrauch zu erhalten.

Andere Teilnehmerinnen und Teilnehmer konnten offene Verwaltungsdaten im Rahmen ihrer Tätigkeit schon einsetzen und konkrete Beispiele nennen. Ein Teilnehmer aus dem Marketingbereich verwies auf den Nutzen von Luftbildern für seine Arbeit, da diese beispielsweise angewandt werden könnten, wenn eine Kundin oder ein Kunde fragt, ob in der Nähe einer Koordinate z. B. ein Garten sei. Des Weiteren könne man beispielweise von erkennbaren Schaukeln auf einer Koordinate eines Luftbildes ableiten, dass dort vermutlich Familien lebten. Grundsätzlich seien Luftbilder notwendig, um prüfen zu können, ob sich bestimmte Entwicklungen lohnen. Ein Teilnehmer aus der Energiewirtschaft hob

ebenso den Nutzen von Luftbildern hervor. In der Energiewirtschaft könnten Luftbilder beispielweise darauf hinweisen, wo man Solaranlagen anbringen kann. Es mache wenig Sinn Solaranlagen zu planen, wenn man dafür keine Dachflächen habe. Ähnliches gelte für Stromtrassen: Durch Luftbilder könne man erkennen, wo Flächen bzw. Korridore vorhanden seien, um eine Stromtrasse zu legen. Laut dem Teilnehmer seien Daten hier sehr wichtig, lägen allerdings häufig nicht vor.

Für den Teilnehmer aus dem Marketing seien des Weiteren Daten relevant, mit denen die jeweilige Zielgruppe möglichst präzise, aber datenschutzkonform beschrieben werden können. Dies können etwa regionale, lokale Informationen, wie die Wohnumfelder, die Pendlerstrukturen oder die Arbeitslosenquoten sein. Darüber hinaus seien im Bereich Smart City Daten zur Auslastung öffentlicher Parkhäuser oder des öffentlichen Personennahverkehrs beispielweise für private Taxiunternehmen von Bedeutung. Diese seien oft Anschlüsse für die öffentlichen Verkehrsmittel und hätten daher unter anderem Interesse daran zu wissen, wie viele Leute ankommen oder wo sie umsteigen.

Ein Teilnehmer wies darauf hin, dass offene Verwaltungsdaten auch dazu dienen, einen „neutralen Blick“ zu bekommen, indem man die offenen Daten mit den eigenen Daten abgleiche. In diesem Zusammenhang betonte ein weiterer Teilnehmer, dass man offene Daten beispielweise auch zur Validierung von Ergebnissen aus Umfragen heranziehen könne, da in der Umfrageforschung häufig eine Verzerrung vorläge. Der Vorteil von prozessgenerierten Daten sei, dass sie – auch wenn sie nicht zu 100 Prozent stimmten – näher an der Wahrheit seien als das, was man mit Umfragen herausbekomme.

In der Fokusgruppe Wirtschaft wurde auch der Einsatz von künstlicher Intelligenz diskutiert. Ein Teilnehmer hob in diesem Zusammenhang hervor, dass KI nur zuverlässig funktioniere, wenn sie mit passenden Datensätzen trainiert werde. Das Problem sei jedoch, dass sich nur wenige die notwendigen Daten und somit den Einsatz von KI leisten könnten. Dies seien vor allem diejenigen, die selbst schon über die Daten verfügen, oder große Unternehmen, die über deutlich mehr finanzielle Mittel verfügen als kleinere Unternehmen und Kommunen und die benötigten Daten daher einkaufen können. Dies führe zu einem starken Ungleichgewicht, da große Unternehmen somit deutlich effizienter KI nutzen könnten. Durch offene Daten könne dieses Ungleichgewicht in Teilen behoben werden.

3.5.2.1.3.3 Nutzerinnen und Nutzer aus der Zivilgesellschaft

Die Teilnehmerinnen und Teilnehmer in der Fokusgruppe Zivilgesellschaft berichteten über unterschiedliche Projekte, die auf Basis offener Daten entstanden seien. Ein ausführlich diskutiertes Thema war dabei die Rolle von offenen Daten in der Pandemiebekämpfung. So erzählte ein Teilnehmer von einer App, die zum Ziel habe, Menschen zu helfen, durch die Pandemie zu kommen. Es handele sich hierbei um ein privates Projekt, das gemeinsam mit Freunden initiiert wurde. Mit einem Bot wurde auf unterschiedliche offene Daten zur Pandemie zurückgegriffen. Die Ergebnisse konnten über verschiedene Messenger-Plattformen abonniert werden und wurden für die Nutzerin oder den Nutzer graphisch aufbereitet. Ebenso mit Bezug auf die Covid-19-Pandemie erzählte eine Teilnehmerin, dass mithilfe offener Daten eine Karte mit der Breite von Gehwegen entwickelt wurde, um zu zeigen, wo es besonders leicht oder schwierig ist, einen Mindestabstand von anderen Personen einzuhalten. Solche Karten könnten laut der Teilnehmerin auch aufzeigen, in welchen Stadtvierteln es sich lohnen würde, gegebenenfalls Straßen für Fußgänger zu öffnen oder Spielstraßen zu schaffen.

Auch zu anderen gesellschaftlichen Problemfeldern berichteten die Teilnehmerinnen und Teilnehmer von Apps, durch die ein positiver Nutzen für die Gesellschaft geschaffen werden sollte. In Berlin wurde beispielweise auf Basis verschiedener Datenquellen eine Erfrischungskarte entwickelt. Um den Nutzerinnen und Nutzern dieser Karte in den heißen Sommertagen besonders schattige und kühle Orte

für bestimmte Uhrzeiten vorzuschlagen, wurde unter anderem berücksichtigt, wie hoch Gebäude sind; auf dieser Grundlage habe man ausrechnen können, wo und wann es an einem Sommertag Schatten geben würde. Auch die Windstärke wurde in die Berechnungen einbezogen, um besonders erfrischende Orte zu identifizieren.

Ein weiterer Teilnehmer berichtete von einem Navigationsprojekt, in dem Nutzerinnen und Nutzer u. a. Kombinationen unterschiedlicher Transportmittel vorgeschlagen werden. Solche Projekte würden laut dem Teilnehmer allerdings von der Verfügbarkeit von Echtzeitdaten leben und hier gebe es noch viel nicht ausgeschöpftes Potenzial. So könnten beispielweise Daten zur Parkplatzverfügbarkeit oder zum Straßenwetter hilfreich sein. Ebenso bräuchte man aktuelle Prognosezeiten, weil es wenig bringe, zu wissen, dass in fünf Minuten der Bus kommen sollte, dieser aber eigentlich ausgefallen ist. Sobald diese Lücken behoben seien, könnten derartige Anwendungen einen Beitrag zur Verkehrswende leisten, um individuelle Mobilitätslösungen bei Nutzung öffentlicher Verkehrsmittel aufzuzeigen.

Darüber hinaus wurde von einem Projekt erzählt, in dem in einem Wettbewerb verschiedene Städte aufgefordert wurden, Photovoltaik-Anlagen auszubauen. Dadurch, dass die Daten zu diesem Wettbewerb öffentlich zugänglich gemacht wurden, wurde laut dem Teilnehmer ein weiterer Anreiz geschaffen, sich für die Energiewende zu engagieren. Die Daten seien wöchentlich abgerufen worden, um auf dieser Datengrundlage ein aktuelles Ranking zu erstellen.

Ähnlich wie auch in der Fokusgruppe Wirtschaft betonte ein Teilnehmer in der Fokusgruppe Zivilgesellschaft den Wert von offenen Verwaltungsdaten zur Kontrolle anderer Daten oder Behauptungen. Gerade bei Fake News im Bereich von Energiethemen könnten offene Daten entlarvend sein. Transparenz und Nachprüfbarkeit seien notwendig, wenn sich jemand beispielweise hinstelle und sage, sein Bundesland sei führend beim Ausbau von erneuerbaren Energien. Hier sei es wichtig, dass man Daten danebenhalten könne, um den Wahrheitsgehalt solcher Aussagen zu überprüfen. Dies träfe laut einem anderen Teilnehmer u. a. auch auf das Thema Kinderarmut zu. Hierbei handle es sich um ein hochpolitisches Thema und man bekomme je nach Definition ganz unterschiedliche Quoten. Auf der Grundlage solcher Zahlen würde jedoch Sozialpolitik gemacht. Es sollte selbstverständlich sein, dass in diesem Zusammenhang Daten öffentlich zugänglich sind, um Studien zu verifizieren. Offene Verwaltungsdaten hätten in diesem Sinne einen großen Mehrwert für die Demokratie und für das Demokratieverständnis, weil der Staat nicht einfach nur Daten sammelt, sondern sie der Gemeinschaft auch zur Verfügung stellt. Der Staat mache sich somit angreifbar und das solle auch so sein.

3.5.2.2 In den Statements angesprochene Aspekte zur Nutzung offener Daten

3.5.2.2.1 Problem bei der Nutzung offener Daten

Neben den in den folgenden Kapiteln beschriebenen Problemlagen zu rechtlichen und technischen Aspekten wurden in den betrachteten Statements insbesondere eine allgemein unzureichende Bereitstellung offener Daten durch die öffentliche Verwaltung kritisiert. Als mögliche Ursachen für dieses Problem wurden häufig eine fehlende Ressourcenausstattung bei den bereitstellenden Organisationen sowie das Fehlen einer „Open-Data-Kultur“ genannt.⁷⁶ Eine Ressourcenaufstockung

⁷⁶ Vgl. z. B. Open Legal Data (2021), S.2; Wiebe (2020), S.5.

wurde insbesondere für die Nationale Metadatenplattform GovData.de⁷⁷ und das CCOD gefordert⁷⁸. Mehrere der betrachteten Statements forderten auch einen generellen Ausbau der Ressourcenausstattung zur Datenbereitstellung in Behörden auf allen föderalen Ebenen. Die erfolgreiche Bereitstellung offener Verwaltungsdaten erfordere fachlich geschultes Personal und eine geeignete digitale Infrastruktur, beides sei bisher in den meisten Behörden nicht ausreichend vorhanden.⁷⁹

Als Beleg für die These, dass die mangelnde Bereitstellung insbesondere auf fehlende Ressourcen zurückzuführen sei, führten mehrere der Statements die Ergebnisse der Behördenbefragung an, die im ersten Fortschrittsbericht Open-Data der Bundesregierung veröffentlicht wurden.⁸⁰ Die Befragung zeigte, dass 57 Prozent der befragten Bundesbehörden nach eigener Einschätzung nicht über ausreichend personelle und finanzielle Ressourcen für die angestrebte Bereitstellung offener Daten verfügen.⁸¹ In mehreren der betrachteten Statements wurde die Einführung einer Open-Data-Koordinatorin bzw. eines Open-Data-Koordinators als Schritt in die richtige Richtung gesehen, wobei angeregt wurde, die genaue Funktion und Qualifikation dieser neu geschaffenen Stelle näher zu definieren.⁸² In dem betrachteten Beitrag von Netzpolitik.org wurden konkrete Punkte zu den möglichen Befugnissen dieser neugeschaffenen Stellen vorgeschlagen: „Im Gesetzentwurf wird es vermieden, Analogien zu den behördlichen Datenschutzbeauftragten herzustellen. Dabei ist es durchaus denkbar, sich die Open-Data-Koordinator:innen perspektivisch auch mit ähnlichem Mandat vorzustellen.“⁸³

Neben den im Gesetz vorgeschlagenen Open-Data-Koordinatorinnen und Koordinatoren wurden in den verschiedenen Statements auch die Einrichtung weiterer Stellen gefordert bzw. vorgeschlagen, die zur vermehrten Bereitstellung von offenen Verwaltungsdaten beitragen könnten. Angeregt wurde vor allem die Schaffung bzw. der Ausbau von zentralen Stellen zum Thema Open Data. Als behördenübergreifende Lösung wurde die Einrichtung einer zentralen Ombudsperson angeregt.⁸⁴ In einer Stellungnahme der OKF zum Zweiten Open-Data-Gesetz wurde als mögliche Funktion dieser Person ein Vergleich zur slowenischen Dateninformationsbeauftragten gezogen. „Die Dateninformationsbeauftragte der slowenischen Regierung kann beispielsweise Einspruch erheben, wenn aus nicht triftigen Gründen von Datenbereitstellungen abgesehen wird und betreffende Behörden um Stellungnahme bitten.“⁸⁵ In dem betrachteten Zehn-Punkte-Plan für Open-Government-Data der Bitkom wird ein Konzept mit mehreren neu geschaffenen Stellen zur Verbesserung der Bereitstellung und Koordination offener Datenangebote skizziert. Dieses Konzept umfasst die Einrichtung von Experimentierräumen, die Schaffung eines „Open Data Instituts“ und die Schaffung

⁷⁷ Siehe [Kapitel 3.5.3](#).

⁷⁸ Vgl. Bundesverbands Digitale Wirtschaft e. V. (2020), S.4.

⁷⁹ Vgl. Open Legal Data (2021), S.2.

⁸⁰ Vgl. D64 (2021), S.1; Wiebe (2020), S.4.

⁸¹ Vgl. Bundesregierung (2019b), S.27.

⁸² Vgl. Rat für Sozial- und Wirtschaftsdaten (2021), S.3; Bundesverbands Digitale Wirtschaft e. V. (2020), S.3.

⁸³ Litta (2021).

⁸⁴ Vgl. OKF (2021), S.2.

⁸⁵ OKF (2021), S.3.

von „Daten-Manager“-Stellen in Behörden auf allen föderalen Ebenen.⁸⁶ Das angedachte Open-Data-Institut soll laut 10-Punkte-Plan sowohl zur Vernetzung der beteiligten Akteure beitragen, als auch Beratungs- und Schulungsangebote bereitstellen, hierzu sei die Unabhängigkeit einer entsprechenden Einrichtung essenziell. Diese erfordere es, dass die Finanzierung des Instituts Personen und Organisationen aus allen Gesellschaftsbereichen offenstehe, sowohl bezogen auf die generelle Finanzierung des Instituts als auch auf die Möglichkeit einer projektbezogenen Beauftragung des Instituts.⁸⁷ Der bundDEV forderte ebenfalls die Einrichtung einer unabhängigen Einrichtung in Form einer „Verbrauchsstiftung Open Data“⁸⁸.

Neben der generellen unzureichenden Bereitstellung von offenen Daten wurde weiterhin eine mangelnde Beteiligung der Nutzerinnen und Nutzer bei der Gestaltung der offenen Daten als problematisch beschrieben. Die Vorteile eines vermehrten Austausches wurden im 10-Punkte für Open Government Data-Plan sowohl auf Seite der Nutzerinnen und Nutzer als auch der Datenbereiterinnen und -bereitstellern gesehen: „Einerseits können Data Provider aus der öffentlichen Hand besser nachvollziehen, wie ihre bereitgestellten offenen Daten genutzt und weiter verwertet werden. Gleichzeitig erfahren sie aber auch, welche Herausforderungen die Nutzung durch verschiedene Nutzergruppen (interessierte Bürgerinnen und Bürger, Hobby-Coder, professionelle Entwicklerinnen und Entwickler) erschweren und können so gezielt die Bereitstellung für Nicht-Fachleute wie auch für Profis nutzerfreundlicher gestalten. Andererseits erhalten Data Consumer Einblicke, wie Verwaltungshandeln funktioniert, welche Datenbestände ggf. noch erschlossen werden und welche Hürden für eine Bereitstellung genommen werden müssen.“⁸⁹ In anderen Statements wurde zudem die stärkere Berücksichtigung von Nutzerinnen und Nutzern bei weiteren regulatorischen Prozessen im Bereich offene Verwaltungsdaten. So kritisierte die OKF in einer Stellungnahme, dass Erfahrungen von Akteuren außerhalb der öffentlichen Verwaltung in den Open-Data-Fortschrittsbericht nicht einbezogen worden seien.⁹⁰ Das Max-Planck-Institut für Innovation und Wettbewerb forderte in einem Statement zur zweiten Änderung des Open-Data-Gesetzes, dass zukünftig auch gesetzlich eine Einbeziehung von Nutzerinnen und Nutzern in die Erstellung weiterer Verordnungen zum Thema Open Data festgeschrieben werden sollte.⁹¹

3.5.2.2 Angesprochene Nutzungszusammenhänge

Der Nutzen offener Daten wurde in den betrachteten Statements generell als hoch eingeschätzt. So heißt es im 10-Punkte Plan für Open Government Data-Plan: „Der freie Zugang zu und die breite Nutzung von Daten bilden eine wichtige Säule für die Digitalisierung der Verwaltung, Wirtschaft, Wissenschaft und Zivilgesellschaft. Denn nicht nur die Bedeutung des Digitalen wird in allen

86 Vgl. Bitkom (2020).

87 Vgl. Bitkom (2020) S.7f.

88 bundDEV (2021).

89 Vgl. Bitkom (2020), S.2.

90 Vgl. OKF (2021), S.4.

91 Max-Planck-Institut für Innovation und Wettbewerb (2021), S.3.

gesellschaftlichen Belangen immer grundlegender, sondern auch die bei der Benutzung erzeugten und verwendeten Daten werden immer wertvoller.“⁹²

Mehrere der betrachteten Statements betonten den hohen wirtschaftlichen Wert offener Verwaltungsdaten. Als Beleg für den zu erwartenden wirtschaftlichen Nutzen wurde auf die Studie „The Economic Impact of Open Data – Opportunities for value creation in Europe“ der Europäischen Union hingewiesen⁹³, ebenso wie auf eine Studie der Konrad-Adenauer-Stiftung aus dem Jahr 2016: „Open Data. The Benefits“⁹⁴. Als besonders relevant wurden offene Daten von mehreren Akteuren für die Innovationsentwicklung beschrieben.⁹⁵ Beispielsweise im Bereich KI könnten offene Daten wesentlich zur Entwicklung innovativer Anwendungen beitragen.⁹⁶ Weitere genannte Anwendungsbereiche waren die Block-Chain-Technologie, die wissenschaftliche Forschung sowie der Mobilitäts- und der Gesundheitsbereich.⁹⁷

Bezogen auf den Nutzen von offenen Daten innerhalb der öffentlichen Verwaltung betonte die Bundesvereinigung der kommunalen Spitzenverbände in einem Statement, dass auch die Kommunalverwaltung Nutzen aus der Bereitstellung offener Daten ziehen könne, beispielsweise im Bereich der Geodaten.⁹⁸ Zudem wurde in einem Thesenpapier zur PSI-Richtlinie angemerkt, dass offene Daten auch die verwaltungsinterne Kommunikation vereinfachen können, da oft nicht bekannt sei, welche Daten in anderen Behörden bereits vorhanden sind.⁹⁹ In der Studie „Open Data in Deutschland und Europa“ der Konrad-Adenauer-Stiftung wurde zudem argumentiert, die Verwaltung könne auch „von der Weiterbearbeitung und „Veredelung“ der bereitgestellten Daten profitieren.“¹⁰⁰

Im Hinblick auf den zivilgesellschaftlichen Nutzen von offenen Daten wurde eine gesteigerte Transparenz von Regierungs- und Verwaltungshandeln von mehreren Akteuren als zentraler Vorteil genannt.¹⁰¹ Als weitere zivilgesellschaftliche Vorteile wurden die Korruptionsbekämpfung sowie eine verbesserte Informationsgrundlage zur Entlarvung von Fake-News angeführt.¹⁰²

3.5.3 Technische Aspekte

3.5.3.1 In den Fokusgruppen genannte technische Aspekte zur Nutzung offener Daten

Im dritten inhaltlichen Block wurde in den Fokusgruppen über technische Aspekte von Open-Data-Angeboten gesprochen, die einen Einfluss auf die Nutzungserfahrungen haben. Der Übergang vom

92 Vgl. Bitkom (2020), S.1.

93 Vgl. OKF (2021), S.2.

94 Vgl. Wiebe (2020), S.4.

95 Vgl. Bitkom (2020), S.1; Wiebe (2020), S.10.

96 Vgl. Bitkom (2020), S.1.

97 Vgl. Bitkom (2020), S.1; RELX (2021), S.1; siehe auch [Kapitel 3.5.1.1](#).

98 Vgl. Bundesvereinigung der kommunalen Spitzenverbände (2021), S.2.

99 Vgl. Smart-City-Dialog (2020), S.20.

100 Vgl. Wiebe (2020), S.10.

101 Vgl. dbb (2021), S.1; Wiebe (2020), S.11; Smart-City-Dialog (2020), S.20.

102 Vgl. Wiebe (2020), S.11.

zweiten zum dritten inhaltlichen Block war fließend, da die meisten genannten technischen Problemstellungen im engen Zusammenhang mit den sonstigen geschilderten Problemen bei der Nutzung stehen. Insgesamt wurde in den Fokusgruppen über technische Aspekte zu Dateiformaten, zur Einheitlichkeit von Datenstrukturen sowie zu Bedarfen bei der Dokumentation von Daten und der Einrichtung von Schnittstellen zum automatisierten Abruf gesprochen. Die getätigten Äußerungen zur Seite GovData.de werden ebenfalls in diesem Berichtsabschnitt vorgestellt.

In den Vorstellungsrunden wurde nur selten Interesse an einer Diskussion über technische Aspekte von offenen Daten angegeben, das Forschungsteam fragte daher gezielt nach Erfahrungen und wahrgenommenen Problemen in diesem Bereich, da dieser Punkt im Sinne des Forschungsvorhabens als besonders wichtig eingestuft wurde. In den unterschiedlichen Fokusgruppen wurden daraufhin in unterschiedlicher Intensität verschiedene technische Aspekte diskutiert. In der Fokusgruppe Zivilgesellschaft wurde zu technischen Aspekten vor allem über die Dokumentation von bereitgestellten Daten und die Möglichkeiten zum automatisierten Abruf gesprochen. Im Gegensatz dazu wurde insbesondere in den Fokusgruppen öffentliche Verwaltung und Wirtschaft über Herausforderungen und Erfahrungen mit verschiedenen Dateiformaten und der Standardisierung von Daten gesprochen. In der Fokusgruppe Zivilgesellschaft wurden diese Aspekte hingegen nur kurz erwähnt.

3.5.3.1.1 Dateiformate

Grundsätzlich wurden in den Fokusgruppen wenige Probleme zur Beschaffenheit der angebotenen Dateiformate von Open-Data-Angeboten angemerkt. Als bevorzugte Formate wurden plattformunabhängige Datei-Formate wie CSV- und JSON-Dateien angegeben. Mehrere der Teilnehmenden äußerten hierzu, dass dies bereits erfahrungsgemäß durch die Plattformbetreibenden umgesetzt würde. Gerade das CSV-Format habe sich für einfach strukturierte Daten inzwischen durchgesetzt. Lediglich ein Teilnehmender berichtete von einem Fall, indem vom Robert Koch-Institut bereitgestellte Daten nur als Excel-Tabelle vorhanden waren, was aufgrund deren Aufbau zu Problemen beim Einlesen der Daten geführt habe.

Auch diejenigen der Befragten, die selbst Daten bereitstellen, äußerten, bisher hätten sie keine Anfragen nach Dateiformaten erhalten, die sie nicht hätten umsetzen können. Einer dieser Befragten berichtete, er stimme sich gelegentlich mit verschiedenen Nutzerinnen und Nutzern seiner Daten zu gewünschten Dateiformaten ab. Dabei sei seiner Erfahrung nach bei Akteuren aus dem zivilgesellschaftlichen Bereich das JSON-Format für komplexer strukturierte Daten tendenziell beliebt, wohingegen Akteure aus der öffentlichen Verwaltung häufiger den Wunsch nach einer Bereitstellung im XML-Format nach XÖV-Standard äußerten. Allgemein seien XML-Dateien in der Verwaltung üblicher. XÖV steht für XML in der Öffentlichen Verwaltung und bezeichnet einen deutschen Standard zum Austausch von XML-Dateien innerhalb der öffentlichen Verwaltung, welcher durch die Koordinierungsstelle für IT-Standards gepflegt wird.¹⁰³ Gelegentlich käme auch von Nutzerinnen und Nutzern aus der öffentlichen Verwaltung der Wunsch vorhandene CSV-Dateien auch als Excel-Dateien bereitzustellen.

¹⁰³ Vgl. <https://www.xoev.de/xoev-4987> (Abrufdatum: 25. Februar 2022).

Auch bezogen auf Geodaten wurden durch die Teilnehmenden keine grundlegenden Probleme hinsichtlich der Dateiformate geschildert. Hier seien SHAPE-Dateien das gängige Format zur Datenbereitstellung.¹⁰⁴ Dieses Dateiformat wurde von den Teilnehmenden der Fokusgruppen als grundsätzlich geeignet eingeschätzt.

Als problematisch wurde hingegen durch die Teilnehmenden beurteilt, wenn Daten in nicht maschinenlesbarer Form bereitgestellt werden. Es wurde hierzu ein Fall geschildert, in denen benötigte Daten ausschließlich im PDF-Format bereitgestellt wurden, aus dem die benötigten Informationen erst mühsam ausgelesen werden mussten. Die Teilnehmenden aus der entsprechenden Fokusgruppe (Zivilgesellschaft), waren sich darüber einig, dass eine solche Bereitstellung vermieden werden sollte.

3.5.3.1.2 Standardisierung von Datenstrukturen

Deutlich mehr Probleme als im Zusammenhang mit den bereitgestellten Dateiformaten wurden hinsichtlich der inhaltlichen Strukturierung von offenen Verwaltungsdaten durch die Teilnehmenden beschrieben. Obwohl diese Punkte im Block „technische Aspekte“ genannt wurden, sind die geschilderten Problemlagen eher organisatorischer Natur. Nach Einschätzung mehrerer Teilnehmenden sei eine fehlende Standardisierung von Daten ein relevantes Problem bei der Nutzung von Open-Data-Angeboten. Die geschilderten Probleme zur Einheitlichkeit bezogen sich dabei teilweise auf die Einheitlichkeit von bereitgestellten Daten über die Zeit und teilweise auf die Einheitlichkeit von bereitgestellten Daten ähnlichen Inhalts durch verschiedene Bereitstellerinnen und Bereitsteller.

Nach Erfahrung der teilnehmenden Nutzerinnen und Nutzer kommt es vor, dass bereitgestellte Daten ohne Ankündigung oder Erklärung in ihrer Struktur verändert werden und verschiedene Versionen der gleichen Daten unterschiedlich strukturiert sind, z. B. durch unterschiedliche Variablennamen, abweichende Sortierungen etc. Dies stelle die Nutzerinnen und Nutzer vor erhebliche Herausforderungen, da die auf den ursprünglichen Daten basierenden Auswertungen entsprechend angepasst werden müssen. Werden bestehende Daten ohne Ankündigung geändert, kann es zudem zu unbemerkten Fehlern in den darauf basierenden Anwendungen der Nutzerinnen und Nutzern kommen. Dieses Problem steht also in engem Zusammenhang zu den in [Kapitel 3.5.2](#) geschilderten Problemen in der Kommunikation zwischen Nutzerinnen und Nutzern und Bereitstellenden offener Daten.

Noch häufiger wurden Probleme in Hinblick auf fehlende Standardisierung bei der Bereitstellung von inhaltlich ähnlichen Daten durch verschiedene Organisationen beschrieben. Ein Teilnehmender nannte beispielsweise Probleme bei der Bereitstellung von Sozialstrukturdaten durch verschiedene Bundesländer, bei denen verschiedene Metriken leicht anders definiert seien. Teilweise verhindere dies eine Zusammenführung der entsprechenden Daten. Zudem sei die Prüfung auf Konsistenz extrem Zeit- und Ressourcenaufwändig. Einer der Teilnehmenden berichtete hierzu aus der Perspektive der Bereitstellerinnen und Bereitstellern, dass auch für die Datenzulieferer eine fehlende Standardisierung

¹⁰⁴ SHAPE ist ein Dateiformat für die Speicherung von geographischen Informationen, welches durch einen Anbieter für Geoinformationssysteme namens Esri entwickelt wurde. Vgl. <https://desktop.arcgis.com/de/arcmap/10.3/manage-data/shapefiles/what-is-a-shapefile.htm>.

Mehraufwand verursache. So müsse seine Verwaltungseinheit Daten zur Wasserqualität an unterschiedliche Behörden liefern, die diese wiederum in unterschiedlichen Formaten verlangen. Generell sah dieser Teilnehmer insbesondere im Bereich der Umweltdaten dringenden Standardisierungsbedarf.

Bei der Bereitstellung von Geodaten besteht laut einem Teilnehmenden eine unterschiedliche Herangehensweise in der Darstellung zwischen öffentlicher Verwaltung und Zivilgesellschaft. Geographische Koordinaten können auf unterschiedliche Arten und Weisen angegeben werden, je nach zugrundeliegendem Koordinatensystem. In der deutschen öffentlichen Verwaltung sei nach Auskunft des Teilnehmenden das Gauß-Krüger-Modell üblich, während in den internationalen ausgerichteten Bereichen hingegen auch in der öffentlichen Verwaltung das allgemein gebräuchlichere UTM-Format verwendet werde. Grundsätzlich würden die unterschiedlichen Referenzsysteme gerade bei Einsteigern in die Thematik häufig zu Verständnisproblemen führen.

In der Fokusgruppe Wirtschaft, in der die Standardisierungsaspekte ausführlicher als in den anderen Gruppen diskutiert wurden, waren sich die Teilnehmenden einig, dass mehr Bemühungen zur Vereinheitlichung der Datenstrukturen von Daten mit ähnlichem Inhalt durch unterschiedliche Bereitstellerinnen und Bereitsteller wünschenswert seien. Die entsprechenden Anpassungen sollten jedoch aus Sicht der Teilnehmenden nach Möglichkeit in Zusammenarbeit mit den Nutzerinnen und Nutzern vorgenommen werden, da auch eine Anpassung mit dem Ziel der Vereinheitlichung zu Problemen mit den auf den alten Datenstrukturen basierenden Anwendungen führen können.

3.5.3.1.3 Dokumentation der Daten und Zugang über Programmierschnittstellen

Ein weiterer Punkt zur technischen Organisation offener Verwaltungsdaten, der insbesondere in der Fokusgruppe Zivilgesellschaft diskutiert wurde, waren die Bedarfe in Hinblick auf die Dokumentation von Daten und die Organisation von versionierten Daten. Dies sei in einem geschilderten Fall problematisch gewesen, da die Variablennamen von Daten in Tabellen nicht selbsterklärend gewesen seien. Es sei in diesem und ähnlichen Fällen schwierig, die entsprechenden Dokumentationen zu finden, da Daten und Dokumentation nicht zusammenhängend organisiert seien.

Die Teilnehmenden waren sich darüber bewusst, dass eine gute Dokumentation mit erheblichem Aufwand für die datenbereitstellende Organisationen verbunden sein kann. Daher forderten die Teilnehmenden auch in diesen Bereich einen stärkeren Einbezug der Nutzerinnen und Nutzer bei der Entscheidung, welche Daten in welcher Form dokumentiert werden sollten, um eine bedarfsgerechte Priorisierung sicherzustellen.

Die beschriebenen Probleme zur unübersichtlichen Strukturierung von Daten verschärfen sich aus Sicht der Teilnehmenden noch, wenn es sich um sogenannte dynamische Daten handelt, die sich durch eine häufige Aktualisierung auszeichnen. In diesen Fällen sei laut den Beteiligten ein möglicher Zugriff über eine API die beste Wahl. Auf diese Weise können die Nutzerinnen und Nutzern die Daten selbstständig in für sie passenden Zeitabständen abrufen und die Gefahr von Fehlern bei der Zusammenführung von verschiedenen Versionen der Daten sei minimiert. Zudem müsse so nicht ständig manuell oder mit einer eigenen Scraping-Anwendung überprüft werden, ob die verwendete Version der Daten die aktuellste sei und es sei allgemein ein einfacherer und zielgenauerer Abruf der benötigten Daten möglich.

3.5.3.1.4 Erfahrungen mit der Webseite GovData.de

Zum Einstieg in die drei Fokusgruppen wurde jeweils kurz der Aufbau des Forschungsprojekts vorgestellt. In diesem Zusammenhang wurde erläutert, dass in der ersten Projektphase primär die Webseite GovData.de betrachtet wurde, während in der zweiten Projektphase der Blickwinkel erweitert werden sollte auf die generelle Nutzung von Open-Data-Angeboten in Deutschland. Den Teilnehmenden war also das Forschungsinteresse an dieser Seite bekannt und einige der Teilnehmenden erwähnten daher auch schon in der Vorstellungsrunde, dass sie bereits mit GovData.de gearbeitet hätten. In allen drei Fokusgruppen wurde aber auch nochmal gezielt durch das Forschungsteam nach bisherigen Erfahrungen mit dieser Webseite gefragt.

Die Plattform war nicht allen Teilnehmenden gleichermaßen bekannt. Einige Teilnehmende gaben an, die Plattform gelegentlich zu nutzen, mehrere Teilnehmende hatten jedoch noch nie von der Plattform gehört. Anderen Teilnehmenden war die Plattform wiederum grundsätzlich bekannt, die dort vorhandenen Inhalte wurden jedoch für die eigenen Nutzungsbereiche als irrelevant eingeschätzt.

Ein Nutzungszusammenhang dieser Webseite, der von mehreren Befragten aus dem Bereich der öffentlichen Verwaltung mehrfach geschildert wurde, war der Abgleich von Datenstrukturen vor der Bereitstellung von Daten. Die entsprechenden Teilnehmenden gaben an, dass sie gerne auf GovData.de nachschauen, wie inhaltlich ähnliche Daten von anderen Bereitstellerinnen und Bereitstellern strukturiert sind, um sich daran bei der Aufbereitung von eigenen Daten zu orientieren. In diesem Sinne ist die Plattform also ein Hilfsmittel, das zu einheitlicheren Bereitstellung von Daten verschiedener Verwaltungseinheiten beitragen kann.

Der zentrale Kritikpunkt zu GovData.de, der von mehreren Teilnehmenden in ähnlicher Weise geäußert wurde, war, dass die auf GovData.de bereitgestellten Daten zu uneinheitlich seien. Dies schließt an die geschilderten Probleme zur fehlenden Standardisierung an. Neben Problemen mit der Strukturierung von auf GovData.de verlinkten Daten wurde auch ein Standardisierungsproblem in Bezug auf die auf GovData.de hinterlegten Metadaten angemerkt. Alleine durch die unterschiedliche Benennung von Datensätzen sei es für die Nutzerinnen und Nutzern nur schwer zu erkennen, welche Datensätze ähnliche Inhalte enthalten.

Ein Teilnehmer nannte einen konkreten Beispielfall. Bei der Suche nach dem Begriff „Leistungsempfänger“ fänden sich bei der Suche auf GovData.de zwar Daten zur Stadt Dortmund, zur nahegelegenen Stadt Duisburg jedoch nicht unter diesem Suchbegriff. Somit müssten die Nutzerinnen und Nutzer bei Interesse an einem größeren geographischen Gebiet für jede kommunale Einheit erneut nach Daten suchen, da die Daten jeweils unterschiedlich benannt und auch mit unterschiedlichen weiteren Metadaten versehen seien. Der Grund hierfür sei, dass die jeweiligen Daten von den unterschiedlichen zuständigen Behörden selbst eingepflegt werden und diese Angaben nach Einschätzung mehrerer Fokusgruppenteilnehmende zu uneinheitlich seien. Dies mache die Suche nach mehreren Datensätzen mit gleichem inhaltlichem Bezug sehr aufwändig. Das Problem werde auch durch die vergebenen Schlagwörter nicht behoben, da auch diese durch die Bereitstellerinnen und Bereitsteller selbstständig und ohne erkennbare Richtlinien vergeben werden.

Weiterhin kritisch bewertet wurden insbesondere die Metadaten zu den Ansprechpartnern für die Daten und die Angaben zum Dateiformat, letzteres wurde von einem Teilnehmenden als „wie Kraut und Rüben“ beschrieben. Durch die uneinheitlichen Angaben zum Dateiformat sei es kaum möglich, hiernach sinnvoll zu filtern. Weiterhin sei auch problematisch, dass Angaben in den Metadaten teilweise sehr minimal gehalten seien oder ganz vergessen wurden. Insgesamt wurde auch bezogen auf GovData.de der Wunsch nach mehr Standardisierung bei der Datenstruktur der verlinkten Daten und den hinterlegten Metadaten geäußert.

Von den Teilnehmenden gewünscht wurde außerdem eine verbesserte Sortiermöglichkeit der Suchergebnisse auf GovData.de in Bezug auf die geografische Einheit und die inhaltliche Zuordnung der Daten. Es wurde eine Filtereinstellung zur Suche nach geografischen Einheiten an Stelle der Kartensuche gewünscht.¹⁰⁵ Weiterhin wurde durch einen Teilnehmenden das Anlegen von inhaltlichen Kategorien nach Nutzungszusammenhang vorgeschlagen z. B. Daten für den Telekommunikationsbereich, für Energieanbieter etc. In diesem Bereich des lizenzkonformen Umgangs mit den angebotenen Daten wünschte sich einer der Teilnehmenden zudem weitere Informations- und Beratungsangebote von GovData.de.

3.5.3.2 In den Statements geäußerte technische Aspekte zur Nutzung offener Daten

Ebenso wie in den Fokusgruppen, wurde auch in den betrachteten Statements der Wunsch nach mehr Standardisierung, sowohl in Bezug auf die Strukturierung von Daten wie auch zugehörigen Metadaten geäußert.¹⁰⁶ Dabei wurde auch die Wichtigkeit der Berücksichtigung von bestehenden Standards bei weiteren Standardisierungsbemühungen hervorgehoben, weiterhin seien Doppelstrukturen zu vermeiden.¹⁰⁷

Mehrere Stakeholder betonten die Wichtigkeit der Standardisierung von Open-Data-Angeboten auf europäischer Ebene. So stellt z. B. Bitkom aus fehlender Standardisierung erwachsende Probleme fest: „Die praktische Erfahrung zeigt, dass offene Daten, die von verschiedenen Regierungsstellen innerhalb der EU zur Verfügung gestellt werden hinsichtlich Format, Detailtiefe und Datenmodell stark variieren können. Dies stellt erhebliche Hindernisse für die Interoperabilität und damit für die grenzüberschreitende Nutzung dar.“¹⁰⁸

In einer Studie zu Herausforderungen bei der Nutzung offener Daten, die im Rahmen des Projektes mFUND erstellt wurde, wurde ebenfalls fehlende Standardisierung als eines der Hindernisse zur erfolgreichen Nutzung offener Daten beschrieben. Gefordert wurde im Rahmen dessen die Standardisierung von Datenstrukturen, Nutzungsbedingungen und bereitgestellten Metadaten. In Bezug auf die Metadaten hebt die Studie die Bedeutung des Standards DCAT-AP.de hervor: „Durch Etablierung einheitlicher Standards, insbesondere für Metadaten, kann das Auffinden von Datensätzen zusätzlich erleichtert werden. Hier wurde mit der Festlegung des gemeinsamen deutschen Metadatenmodells zum Austausch von offenen Verwaltungsdaten DCAT-AP.de durch den IT-Planungsrat im Juni 2018 (<https://www.govdata.de/standardisierung>) ein wichtiger Schritt getan auf dem Weg zu einer einheitlichen Beschreibung und damit zur Verbesserung der Auffindbarkeit von Open (Government) Data.“¹⁰⁹ Zudem ist durch die Kompatibilität mit dem europäischen Metadatenstandard DCAT-AP.de auch eine Zusammenführung der entsprechenden Metadaten mit

¹⁰⁵ Der Wunsch nach vereinfachter geografischer Filterung wurde in der Zwischenzeit von GovData.de bereits umgesetzt, sodass inzwischen auch nach dem Bundesland auf das sich die Daten beziehen gefiltert werden kann.

¹⁰⁶ Vgl. z. B. Bitkom (2020), S.6; mFUND (2019), S.6.

¹⁰⁷ Vgl. mFUND (2019), S.6; Deutscher EDV-Gerichtstag e. V. (2021), S.3.

¹⁰⁸ Bitkom (2021), S.7.

¹⁰⁹ mFUND (2019), S.6.

Metadaten aus anderen europäischen Ländern möglich, was ebenfalls positiv in der entsprechenden Studie hervorgehoben wurde.¹¹⁰

3.5.3.2.1 Anmerkungen zu Dateiformaten

Der in den Fokusgruppen geäußerte Wunsch nach einer Bereitstellung von Daten in Form von plattformunabhängigen und maschinenlesbaren Dateiformaten findet sich auch in ähnlicher Form in mehreren der betrachteten Statements von anderen Nutzerinnen und Nutzern offener Datenangebote in Deutschland. So heißt es beispielsweise im 10-Punkte für Open Data-Plan der Bitkom: „Open Data zeichnet sich insbesondere durch eine einfache Auffindbarkeit und eine maschinenlesbare Form aus. Daraus folgt, dass eine nutzerfreundliche Bereitstellung von Open Data nicht in proprietären Formaten (z. B. eher CSV als XLS) oder in gänzlich ungeeigneter Form (z. B. PDF) erfolgen kann.“¹¹¹

Im Hinblick auf die anstehende Änderung des Open-Data-Gesetzes wurde in den betrachteten Statements mehrfach gewarnt, der Anspruch auf Maschinenlesbarkeit und Bereitstellung in offenen Formaten sei im Gesetzentwurf zu weich formuliert. Die Kritik wendet sich gegen eine Formulierung in § 7 Abs. 2 der Entwurfsfassung des Zweiten Open-Data-Gesetzes, nach der Daten und Metadaten nur „soweit möglich und sinnvoll, (...) elektronisch und in nach den anerkannten Regeln der Technik offenen, maschinenlesbaren, zugänglichen, auffindbaren und interoperablen Formaten zusammen mit den zugehörigen Metadaten bereitzustellen“¹¹² sind. Somit handle es sich nach Einschätzung mehrerer Nutzergruppen nur um eine kann-Bedingung ohne tatsächliche Wirkung, bei der zudem unklar bliebe, wann eine entsprechende Bereitstellung als „möglich und sinnvoll“ zu bewerten sei.¹¹³

3.5.3.2.2 Anmerkungen zur Organisation und Dokumentation offener Daten

Während der Aspekt der offenen Schnittstellen für den Abruf offener Verwaltungsdaten in den Fokusgruppen eher am Rande angesprochen wurde, wurde in mehreren der betrachteten Statements die Forderung nach einer Bereitstellung von Daten über offene Schnittstellen „API“ bzw. „Open API“ nachdrücklich betont.¹¹⁴ In der Bitkom Stellungnahme zum Zweiten Open-Data-Gesetz findet sich die folgende Erklärung zu dem Begriff Open API: „Durch Open API werden bisher unsichtbare Backend-Systeme für Daten-Konsumenten (Entwickler) sicht- und nutzbar. Durch den Zusammenschluss einzelner Dienste mittels Open API entstehen digitale Wertschöpfungsnetzwerke, wodurch mehr Kundenreichweite erzielt und Innovationen stimuliert werden können. Der Zusatz »Open« stellt somit zugleich eine Beschreibung und einen Appell dar, Schnittstellen nicht nur technisch zu verwenden, sondern diese der Allgemeinheit bekannt zu machen (APIs müssen nach Themen und Relevanz such- und auffindbar sein) und ein aktives Community-Management zu betreiben.“¹¹⁵ Um Open API in Deutschland flächendeckend umzusetzen bedarf es laut des 10-Punkte für Open Data-Plan: „einer

110 mFUND (2019), S.6.

111 Bitkom (2020), S.3.

112 Vgl. Bundesregierung (2021b), S. 12.

113 Vgl. Bitkom (2020), S.5; Rat für Sozial- und Wirtschaftsdaten (2021), S.2.

114 Vgl. Open Legal Data (2021), S.2; Bitkom (2020) S.3.

115 Bitkom (2021) S.5.

gemeinsamen Anstrengung und eines Diskurses von Verwaltung, Wirtschaft, Wissenschaft und Zivilgesellschaft.“¹¹⁶

Um aktiv auf dieses Ziel hinzuwirken, haben sich zudem bereits Aktivistinnen und Aktivisten unter der Bezeichnung „bundDEV“ zusammengefunden. Diese Gruppe dokumentiert vorhandene Schnittstellen, um sie so für die allgemeine Öffentlichkeit nutzbar zu machen. In einem 5-Punkte-Plan fordert die Gruppe in Hinblick auf die Erreichbarkeit von Daten via API: „Dokumentation und Zugänglichmachung aller bestehenden Schnittstellen des Bundes bis Ende 2021.“ sowie „Ergänzung aller Leistungen des Onlinezugangsgesetzes um Programmierschnittstellen bis 2023.“¹¹⁷

Neben der Forderung nach Bereitstellung von Daten über eine API findet sich in den betrachteten Statements auch die Forderung nach Bulk-Downloads.¹¹⁸ In den von Nutzerinnen und Nutzern zusammengetragenen Open-Data-Antipatterns, also festgestellten problematischen Mustern im Bereich Open Data, wird zudem auf zwei Problemstellungen im Zusammenhang mit den automatisierten Abrufen hingewiesen. Zum einen sei es problematisch, wenn verschiedene Versionen der Daten unregelmäßig unter neuen URL bereitgestellt würden, bei denen für die Nutzerinnen und Nutzer nur schwer oder unmöglich erkennbar sei, unter welcher Adresse die nächste Version der Daten veröffentlicht wird. Als Lösung wurde in diesem Zusammenhang vorgeschlagen, die jeweils aktuellste Version der Daten unter einer konstanten URL zu veröffentlichen, die als solche klar gekennzeichnet werden sollte. Als weiteres Antipattern wurde bei dynamischen Daten auch festgestellt, dass Daten jeweils im gleichen Datensatz aktualisiert werden und gleichzeitig für die Nutzerinnen und Nutzer vor dem Download nicht erkennbar ist, ob sich seit dem letzten Download eine Änderung im Datensatz ergeben hat. Die Nutzerinnen und Nutzer müssten in diesem Fall auf Verdacht die Daten regelmäßig herunterladen, nur um dann teilweise festzustellen, dass sich nichts geändert hat. Dieses Problem sei durch angemessene Servereinstellungen zu beheben.¹¹⁹

3.5.3.2.3 Äußerungen in Bezug auf GovData.de

Zur Entwurfsfassung des Zweiten Open-Data-Gesetz, welche den Anspruch formuliert, Metadaten „soweit möglich und sinnvoll“ (§ 7 Abs. 2 S. 1 EGovG) auf der Metadatenplattform GovData.de bereitzustellen, bewerteten mehrere Statements den Zusatz „soweit möglich und sinnvoll“ kritisch.¹²⁰ Der Rat für Sozial- und Wirtschaftsdaten argumentierte, mindestens die Bereitstellung von Metadaten in übersichtlicher und maschinenlesbarer Form müsse gegeben sein, damit eine sinnvolle Nutzung offener Verwaltungsdaten möglich sei.¹²¹ Der D64 – Zentrum für Digitalen Fortschritt e. V. (D64) forderte zudem nicht nur die Metadaten in entsprechender Form auf einer zentralen Plattform bereitzustellen, es sei vielmehr auch eine Bereitstellung der offenen Verwaltungsdaten selbst über eine zentrale Plattform notwendig. Dies habe mehrere Vorteile gegenüber dem aktuellen dezentralen

¹¹⁶ Bitkom (2020) S.3.

¹¹⁷ bundDEV (2021).

¹¹⁸ Vgl. Open Legal Data (2021), S.2.

¹¹⁹ Vgl. Antipatterns (2021).

¹²⁰ Vgl. Bitkom (2021), S.5f.

¹²¹ Vgl. Rat für Sozial- und Wirtschaftsdaten (2021), S.5.

Modell: „Eine zentrale Plattform spart nicht nur Kosten bei der Ausgestaltung und Betrieb von Servern, Webseiten und APIs, sondern erleichtert auch die Suche verfügbarer Datensätze sowohl für Unternehmen (Wettbewerbsvorteil) als auch für die Zivilgesellschaft.“¹²²

Weiterhin wiesen mehrere der betrachteten Statements darauf hin, dass die Ressourcenausstattung von GovData.de erhöht werden sollte, insbesondere damit die Plattform das erwartete Mehraufkommen an Metadaten durch die Umsetzung des zweiten Open-Data-Gesetzes bewältigen könne. In einem Beitrag auf der Seite Netzpolitik.org heißt es hierzu: „Wie gut und einfach Nutzer Daten auf dem Portal [GovData.de] finden können, bestimmt auch, wie häufig das Portal genutzt wird. Im Gesetz und seiner Begründung fehlt aber sowohl eine Zukunftsvision für das Portal und seine Rolle im Open-Data-Institutionengeflecht als auch ein Bekenntnis, es nachhaltig und ausreichend auszustatten.“¹²³

Sowohl die betrachtete mFUND-Studie zur Nutzung offener Daten, als auch der 10-Punkte-Plan für Open Data stellen fest, dass bisher (Stand 11. Januar 2022) noch nicht alle Bundesländer an die Plattform angebunden sind, dies gelte es schnellstmöglich aufzuholen.¹²⁴

3.5.4 Rechtliche Aspekte

3.5.4.1 In den Fokusgruppen diskutierte rechtliche Aspekte zur Nutzung offener Verwaltungsdaten

Rechtliche Aspekte zur Nutzung offener Verwaltungsdaten wurden als einer der vier inhaltlichen Blöcke vom Forschungsteam gesetzt. Zum Zeitpunkt der Fokusgruppen lag das Inkrafttreten des zweiten Open-Data-Gesetzes nur wenige Monate zurück. Daher lag die Vermutung nahe, dass die rechtlichen Änderungen die Nutzerinnen und Nutzer zum Zeitpunkt der Fokusgruppen beschäftigten. Dennoch wurden zu diesem inhaltlichen Block in allen drei Fokusgruppen die wenigsten Punkte angesprochen. Dabei muss jedoch auch beachtet werden, dass dieser Block als letzter mit den Teilnehmenden diskutiert wurde, dementsprechend können auch Ermüdungserscheinungen die Anzahl der Beiträge reduziert haben.

Auch zeigten sich in diesem inhaltlichen Block große Unterschiede in den Interessen der jeweiligen Nutzergruppen. Die angesprochene Änderung des Open-Data-Gesetzes wurde z. B. nur in der Fokusgruppe öffentliche Verwaltung näher diskutiert. Aspekte des Datenschutzes wurden in der Fokusgruppe Zivilgesellschaft nicht angesprochen, dafür aber ausführlich in den beiden anderen Fokusgruppen diskutiert. Dafür zeigten die Teilnehmenden der Fokusgruppe Zivilgesellschaft ein besonderes Interesse an der Lizenzierung von offenen Verwaltungsdaten, einem Themenbereich, der in den anderen beiden Gruppen nur am Rande erwähnt wurde.

¹²² D64 (2021), S.2.

¹²³ Litta (2021).

¹²⁴ Vgl. Bitkom (2020), S.3; mFUND (2019) S.7.

3.5.4.1.1 Datenschutz

Der Aspekt Datenschutz wurde ausschließlich in den Fokusgruppen Wirtschaft und öffentliche Verwaltung diskutiert. In der Fokusgruppe Zivilgesellschaft wurde der Punkt nicht durch die Teilnehmenden angesprochen. Auch auf Anregung des Evaluationsteams wurde zu diesem Themenbereich nichts geäußert. In den anderen zwei Gruppen wurde das Thema eher generell bezogen auf die Arbeit mit Daten beschrieben und somit nicht als ein spezielles Problem bei der Nutzung offener Verwaltungsdaten.

In keiner der Fokusgruppen wurde die Notwendigkeit des Datenschutzes per se in Frage gestellt. Angesprochen wurden hingegen vor allem vorhandene Unsicherheiten bezogen auf die korrekte Anwendung der geltenden Richtlinien. Welche Daten eine im Sinne des Datenschutzes unzulässige Personalisierung enthielten, sei nicht immer klar zu sagen bzw. erfordere großes Fachwissen. Eine Person gab an, das Problem habe sich in letzter Zeit verschärft, ein anderer Teilnehmer meinte, das Thema Datenschutz sei vor allem seit der Einführung der Datenschutz-Grundverordnung ein „großes Thema“.

Das Problem der unklaren Rechtslage wurde sowohl bezogen auf die Nutzung in einem wirtschaftlichen Zusammenhang als auch bezogen auf die Nutzung und Bereitstellung von offenen Daten in der öffentlichen Verwaltung angesprochen. Als Beispiel für einen unklaren Fall bei der Bereitstellung wurden Daten zur Covid-19-Pandemie genannt. Diese konnten nur bis zu einer gewissen geographischen Ebene genau weitergegeben werden, da diese Angaben zu dem Alter der Erkrankten und deren Geschlecht enthielten. Würden diese Angaben pro Straße oder auch nur pro Ortschaft veröffentlicht werden, könnte dies bereits die Deanonymisierbarkeit der entsprechenden Daten zur Folge haben. Wo aber genau die Grenze zu ziehen ist, sei dabei nicht immer ganz klar.

Die Unsicherheiten beim Datenschutz seien unter anderem auch darauf zurückzuführen, dass sich aus als „anonymisiert“ geltenden Daten mit der Anreicherung durch Daten aus anderen Quellen teilweise wieder ein Personenbezug herstellen ließe. Dies führe aus Sicht eines der Befragten häufig dazu, dass die zuständigen Datenschutzbeauftragten einer Veröffentlichung von Daten sehr kritisch gegenüberständen. Dies betreffe auch die Nutzungsseite, da im Rahmen der Nutzung in der Regel ein Zusammenführen von Daten aus verschiedenen Quellen erforderlich ist, wodurch Daten angereichert werden. So berichtete eine Person, welche offene Verwaltungsdaten zur Optimierung von Marketing-Maßnahmen nutzt, dass im zugehörigen Unternehmen bei jeder datenbasierten Maßnahme zuerst eine Sensibilitätsprüfung erfolge. Im Rahmen dessen werde geprüft, welche Daten zur Erreichung der angestrebten Ziele erforderlich seien und ab welchem Punkt die Daten nicht mehr weiter angereichert werden dürfen, um die Anonymisierung der Daten nicht aufzuheben. Neben der eigentlichen Prüfung sei auch eine entsprechende Dokumentation erforderlich, um im Zweifelsfall eindeutig nachweisen zu können, dass die entsprechende Prüfung erfolgt ist.

Die beschriebenen Unsicherheiten wurden von den Nutzerinnen und Nutzern jedoch generell bezogen auf den Umgang mit Daten insgesamt wahrgenommen, nicht ausschließlich bezogen auf den Umgang mit offenen Verwaltungsdaten im Speziellen. Ein Teilnehmer in der Fokusgruppe Wirtschaft gab an, dass bei offenen Daten zumindest der Vorteil bestünde, dass davon ausgegangen werden könne, dass die Daten in der veröffentlichten Form in sich keinen Personenbezug enthielten, somit sei dies in der Prüfung für die Nutzerinnen und Nutzer etwas einfacher als bei Daten aus anderen Quellen. Ein Teilnehmer aus der Gruppe für die Nutzung innerhalb der Verwaltung hatte zudem den Eindruck, dass der Personenbezug häufig wenig interessant für die Nutzerinnen und Nutzer sei. Dass Daten häufig nicht anonymisiert veröffentlicht werden, weil der veröffentlichten Stelle die Ressourcen oder die

Motivation für den aufwändigen Prozess der Anonymisierung fehle, wurde als viel problematischer eingeschätzt als eine Veröffentlichung mit fehlendem Personenbezug.

3.5.4.1.2 *Lizensierung*

Das Thema Lizenzierung wurde ausschließlich in der Fokusgruppe Zivilgesellschaft ausführlich diskutiert, in der Fokusgruppe öffentliche Verwaltung wurde das Thema zwar von einer Person kurz angesprochen, aber nicht tiefergehend behandelt. Die entsprechende Person berichtete von einem Lizenzproblemen bei der Plattform PegelOnline. Dabei handelt es sich um eine Plattform der Wasserstraßen- und Schifffahrtsverwaltung des Bundes, produziert durch das ITZ-Bund, auf der Wasserpegelstände und andere gewässerkundliche Parameter bereitgestellt werden.¹²⁵ Laut den Erfahrungen des Teilnehmenden gab es Probleme bei der Nutzung der Daten, da die verwendete Lizenz nur eine unveränderte Weiterverwendung der Daten erlaubte, also eine „share-alike“ Regelung. Diese verhinderte eine Weitergabe von durchschnittlichen Pegelständen für einen längeren Zeitraum, welche aus den angegebenen Pegelständen von einzelnen Tagen ermittelt wurde.

In der Fokusgruppe Zivilgesellschaft wurde von einer Person das Thema Lizenzierung als „das Rechtsthema schlechthin“ im Zusammenhang mit offenen Verwaltungsdaten eingeschätzt. Als besonders problematisch wurde von den Teilnehmerinnen und Teilnehmern dieser Gruppe die Vielzahl an unterschiedlichen Lizenzen bewertet, dadurch seien die Nutzungsbedingungen für die Nutzerinnen und Nutzer häufig unklar. So war einer teilnehmenden Person beispielsweise nicht klar, warum die Datenlizenz Deutschland eingeführt wurde und welchen Mehrwert diese gegenüber der international angewendeten Lizenz „Creative Common“ habe.

Das Problem der unklaren Lizenzbedingungen wird nach Ansicht eines Beteiligten auch dadurch verschärft, dass Datenbereitstellende Behörden sich zwar an gängigen Lizenzbedingungen orientierten, aber eigene Anpassungen für die Lizenzierung ihrer Daten vornehmen. Dies könne zur Folge haben, dass umfangreiche juristische Prüfungen vor der Verwendung von veröffentlichten Daten notwendig sind, was den Open-Data-Prinzipien widerspricht. Insbesondere sei hiervon auch die Nutzung zu zivilgesellschaftlichen Zwecken betroffen, da in diesem Bereich selten die notwendigen finanziellen Mittel zur Beauftragung von juristischen Prüfungen vorhanden seien. Generell wurde auch in diesem Bereich, ähnlich wie bei der Anwendung des Datenschutzes, eine gewisse Unsicherheit bezüglich der rechtskonformen Anwendung der geltenden Regelungen ausgedrückt.

3.5.4.1.3 *Anmerkungen zur Änderung des Open-Data-Gesetzes*

Die zum Zeitpunkt der Fokusgruppen aktuelle Änderung aufgrund des Inkrafttretens des Zweiten Open-Data-Gesetzes und des Datennutzungsgesetzes wurde zwar in allen Fokusgruppen durch die Moderatorinnen des FÖV angesprochen, allerdings äußerten die Teilnehmenden zu diesem Thema nur wenige Punkte.

Die getätigten Äußerungen bezogen sich vor allem auf die Sorge der kommunalen Unternehmen und Unternehmen der Daseinsvorsorge: Das Zweite Open-Data-Gesetz sieht eine Ausweitung des

¹²⁵ Vgl. <https://www.pegelonline.wsv.de/gast/start> (Abrufdatum: 25. Februar 2022).

Anspruches auf Bereitstellung von Daten in diesen Bereichen vor und zum Zeitpunkt der Fokusgruppen herrschte nach Auskunft einer Vertreterin aus einem Unternehmensverband große Unsicherheit darüber, wie dies zu interpretieren sei und welche Daten konkret veröffentlicht werden sollten. Auch sei unklar, ob sich aus dem Gesetz rechtlich bindende Bereitstellungspflichten ergeben. Weitere Unsicherheiten aus Nutzungsperspektive zur Gesetzesänderung wurden durch die Teilnehmenden nicht angesprochen.

3.5.4.2 In den Statements genannte rechtliche Aspekte

Während in den Fokusgruppen im Vergleich zu den anderen inhaltlichen Blöcken eher wenige Punkte zu rechtlichen Aspekten genannt wurden, hatte eine Mehrzahl der betrachteten Statements eine juristische Ausrichtung. Die meisten der betrachteten Statements kommentiert den Gesetzesentwurf zum Zweiten Open-Data-Gesetz, wobei auch weitergehende Aspekte zur rechtlichen Gestaltung genannt wurden.

Als grundsätzliche Problemlage in Hinblick auf die rechtliche Regelung zu Open-Data wurde in mehreren Statements eine unübersichtliche Gesetzeslage im Bereich der Transparenzgesetzgebung genannt, so heißt es in einem Statement des Max-Planck-Instituts für Innovation und Wettbewerb: „Bezeichnend für sie [die Transparenzgesetzgebung in Deutschland] ist Fragmentierung, Föderalismus, unionsrechtliche Überformung, gerade im Sektorenrecht, und nicht zuletzt die Arkantradition und der dahinterstehende, systemische Interessenkonflikt der öffentlichen Hand, ihre eigenen Informationen bereitzustellen und dem Publikum nutzbar machen zu müssen.“¹²⁶ Auch der dbb beamtenbund und tarifunion (dbb) sieht die Überkomplexität der Gesetzgebung sowohl als Hinderungsgrund für die Bereitstellung wie auch für die Nutzung offener Daten.¹²⁷

3.5.4.2.1 Aussagen zum Datenschutz

In den betrachteten Statements wurde als Problem in Hinblick auf Datenschutz und offene Verwaltungsdaten hervorgehoben, dass sich eine mögliche Deanonymisierung von Daten nur schwer ausschließen lasse, da durch eine Kombination mit Daten aus anderen Quellen und die Möglichkeiten der KI-Technologie diverse Möglichkeiten bestünden einen Personenbezug wieder herzustellen.¹²⁸ Dieses Problem ist im Zusammenhang mit Open Data laut Einschätzung des D64 besonders schwerwiegend; „Bislang ist eine effektive Anonymisierung regelmäßig nur möglich, wenn der Einsatzzweck der anonymisierten Daten bekannt ist. Bei Open Data kann eine bereitstellende Behörde aber nicht abschließend absehen, zu welchen Zwecken die Daten genutzt werden. Einmal veröffentlichte Daten können technisch nicht zurückgenommen werden.“¹²⁹ Daher wurde durch den Verein gefordert, die Forschungs- und Entwicklungsbemühungen im Bereich der Anonymisierungsverfahren weiter voranzutreiben. Zudem forderte der Verein eine Verpflichtung zur

¹²⁶ Max-Planck-Institut für Innovation und Wettbewerb (2021), S.2.

¹²⁷ dbb (2021), S.2.

¹²⁸ Vgl. Deutscher EDV-Gerichtstag e. V.(2021), S.4; D64 (2021), S.3.

¹²⁹ D64 (2021), S.3.

Durchführung einer Datenschutzeinschätzung bei der Veröffentlichung von Daten, die Ergebnisse derselben sollten zudem mit den Metadaten veröffentlicht werden.¹³⁰

Der Bundesverbands Digitale Wirtschaft e. V. weist in seinem Statement zum Zweiten Open-Data-Gesetz ebenfalls auf die Schwierigkeiten bei der Anonymisierung offener Daten hin. Auf diesem Hintergrund fordert der Verband eine Beratungspflicht für die Datenschutzaufsichtsbehörden.¹³¹ Zudem nannte er als Überarbeitungsvorschlag für das Open-Data-Gesetz eine konkrete Nennung von Mindeststandards zur Anonymisierung von offenen Daten.¹³²

3.5.4.2.2 Anmerkung zur Lizenzierung

Wie in den Fokusgruppen auch, wurde in den Statements mehrfach festgestellt, dass eine uneinheitliche und unübersichtliche Lizenzlage ein erhebliches Problem bei der Nutzung offener Daten darstelle.¹³³ Insbesondere sei durch die unterschiedlich gestalteten Lizenzbedingungen die Zusammenführung von Daten aus unterschiedlichen Quellen nur eingeschränkt möglich.¹³⁴ In den betrachteten Antipatterns wurde zudem kritisiert, es käme vor, dass Daten zwar scheinbar unter einer offenen Lizenz veröffentlicht werden, dann aber in einem „Beschreibungstext“ doch bestimmte Nutzungsarten ausgeschlossen werden. Damit sei für die Nutzerinnen und Nutzer nicht erkennbar, welche der widersprüchlichen Regelungen gelte.¹³⁵

Als Lösung für das Problem der uneinheitlichen Lizenzregelungen wurde eine Einigung auf eine oder wenige Standardlizenzen vorgeschlagen.¹³⁶ Auch die rechtliche Festlegung der entsprechenden Lizenzen im Rahmen des Datennutzungsgesetzes wurde gefordert, um der als problematisch wahrgenommenen Vielfalt der Nutzungsbedingungen entgegen zu wirken.¹³⁷ Dabei wurden verschiedene Anforderungen an eine entsprechende Standardlizenz beschrieben. Die Nutzungsbedingungen sollten keine Nutzungsgruppe diskriminieren, keine zeitlichen Einschränkungen der Nutzung definieren und eine kostenlose Nachnutzung erlauben.¹³⁸ Generell wurde angeregt, die Kriterien für offene Lizenzen der opendefinition.org zu übernehmen.¹³⁹ Als grundsätzlich geeignet für eine Bereitstellung offener Daten wurde die „Datenlizenz Deutschland“ eingeschätzt ebenso wie die Lizenzen nach den creative common Prinzipien.¹⁴⁰

130 D64 (2021), S.3.

131 Bundesverbands Digitale Wirtschaft e. V. (2020), S.4.

132 Bundesverbands Digitale Wirtschaft e. V. (2020), S.4.

133 Vgl. mFUND (2019), S.11; OKF (2018), S.4.

134 Vgl. mFUND (2019), S.11.

135 Vgl. Antipattern (2021).

136 Vgl. Open Legal Data (2021), S.2f.

137 Vgl. Open Legal Data (2021), S.2.

138 Vgl. Wiebe (2020).

139 Vgl. OKF (2018), S.1.

140 Vgl. Wiebe (2020), S.16.

Das Thema Lizenzierung wurde durch die OKF auch im Hinblick auf die PSI-Richtlinie thematisiert.¹⁴¹ In dem entsprechenden Statement wurde vor allem die Befürchtung geäußert, dass unterschiedliche Lizenzrecht zur Weitergabe von Daten in den Mitgliedsstaaten der EU könne die Kombination von Daten aus unterschiedlichen Staaten unmöglich machen und die Bildung von „Datensilos“ befördern. Beides sei dem Ziel eines einheitlichen europäischen Datenmarktes abträglich, daher sei auch europaweit die Einigung auf Standardlizenzen anzustreben.¹⁴²

3.5.4.2.3 Anmerkungen zum Zweiten Open-Data-Gesetz

Einige der betrachteten Statements zum Gesetzesentwurf für das Zweite Open-Data-Gesetz äußerten Lob für die Bemühungen der Bundesregierung, das bestehende Open-Data-Gesetz zu reformieren.¹⁴³ Besonderen Anklang fanden dabei die Einführung von Open-Data-Koordinatorinnen und Koordinatoren, sowie die Grundsätze Open by Default und Open by Design.¹⁴⁴ Ein Umdenken der Verwaltung in Richtung Open-Data-Kultur sei dringend erforderlich und die Einführung entsprechender Koordinatorinnen und Koordinatoren könne hierzu beitragen, sofern die entsprechenden Stellen mit einem ausreichend großen Budget ausgestattet werden.¹⁴⁵

Allerdings wurde der Gesetzesentwurf von einigen Akteuren auch als eine Minimallösung zur Umsetzung der PSI-Richtlinie gesehen.¹⁴⁶ Bemängelt wurde vor allem der Ausschluss des Rechtsanspruchs.¹⁴⁷ Darüber hinaus wurde in mehreren Statements kritisch angemerkt, dass der Gesetzesentwurf die bereitzustellenden Daten zu wagen beschrieben seien, so sei unter anderem der verwendete Begriff der „hochwertigen Daten“ unklar.¹⁴⁸ Somit obliege es den bereitstellenden Behörden weiterhin zu entscheiden, welche der vorliegenden Daten zu veröffentlichen seien und welche nicht.¹⁴⁹ Die Konrad Adenauer Stiftung äußerte als Vorschlag, um diesem Problem entgegen zu wirken, ein Recht auf Auskunft darüber, warum bestimmte Daten nach einem bestimmten Zeitraum nach Einführung des Gesetzes noch nicht veröffentlicht wurden.¹⁵⁰ In dem Statement der Bitkom wurde hingegen vorgeschlagen, das Problem zu lösen, indem eine zentrale Stelle mit der Befugnis ausgestattet wird, zu definieren, welche Daten in den unterschiedlichen Behörden nach dem Gesetz veröffentlicht werden müssen und die entsprechende Veröffentlichung einzufordern.¹⁵¹

141 Vgl. OKF (2018), S.4.

142 Vgl. OKF (2018), S.4.

143 Vgl. D64 (2021), S.1; Die Familienunternehmer e. V. (2021), S.1; Konrad-Adenauer-Stiftung (2021), S.1; RELX (2021), S.1.

144 Vgl. D64 (2021), S.1; Bundesverbands Digitale Wirtschaft e. V. (2020), S.2; OKF (2021), S.1.

145 Vgl. Litta (2021).

146 Vgl. Max-Planck-Institut für Innovation und Wettbewerb (2021), S.1; Konrad Adenauer Stiftung (2021), S.4.

147 Vgl. D64 (2021), S.1.

148 Vgl. Max-Planck-Institut für Innovation und Wettbewerb (2021), S.9.

149 Vgl. Max-Planck-Institut für Innovation und Wettbewerb (2021), S.6.

150 Vgl. Konrad Adenauer Stiftung (2021), S.3.

151 Vgl. Bitkom (2021), S.6.

3.5.5 Zusammenfassung

Die drei durchgeführten Fokusgruppen richteten sich jeweils an eine zuvor identifizierte Nutzergruppe von offenen Verwaltungsdaten, diese waren die Nutzung innerhalb der öffentlichen Verwaltung, der privaten Wirtschaft und der Zivilgesellschaft. Als inhaltliche Struktur wurde in allen drei Fokusgruppen eine Unterteilung in vier Blöcke vorgegeben: Inhaltliche Aspekte, allgemeine Nutzungsaspekte, technische Aspekte und rechtliche Aspekte. Zusätzlich wurden veröffentlichte Statements von Nutzerinnen und Nutzern zu deutschen Open-Data-Angeboten nach dem gleichen Schema ausgewertet.

Im Block zu den inhaltlichen Aspekten wurde in den Fokusgruppen hauptsächlich über Geodaten, Sozialstrukturdaten und Mobilitätsdaten diskutiert. Weiterhin wurden Wetterdaten, Kulturdaten und Gesundheitsdaten in einzelnen Fokusgruppen angesprochen. Zu den Geodaten benannten die Teilnehmenden vielfältige Nutzungszusammenhänge, bemängelten aber Lücken in den vorhandenen Daten. Bezogen auf Sozialstrukturdaten wurden vor allem Nutzungszusammenhänge im Bereich der Forschung beschrieben, typische Problemlagen hierbei seien vor allem mangelnde Aktualität sowie fehlende Indikatoren. Als Mobilitätsdaten wurden Informationen zum öffentlichen Nahverkehr sowie Informationen zu Verkehrshindernissen und Unfällen als besonders relevant benannt. Neben Mobilitätsdaten wurden in den betrachteten Statements vor allem Gesundheitsdaten und wissenschaftliche Basisdaten als relevante Datenarten benannt, wobei neben dem positiven Nutzen der Bereitstellung dieser Daten als Open Data auch mögliche Problemstellungen aus Nutzenden- und gesamtgesellschaftlicher Sicht angesprochen wurden.

Zu den Nutzungserfahrungen wurden durch die Teilnehmenden der Fokusgruppen sowohl aktuell wahrgenommene Probleme bei der Nutzung genannt, als auch bereits durchgeführte erfolgreiche Nutzungszusammenhänge von offenen Daten beschrieben und Mehrwerte für die verschiedenen Anwendungsbereiche erläutert. Als grundlegendes Problem bei der Nutzung wurde durch die Teilnehmenden aus allen Fokusgruppen eine mangelnde Bereitstellung offener Verwaltungsdaten beschrieben. Solange Daten nur sporadisch oder erst nach sehr langer Zeit veröffentlicht würden, seien die bestehenden Angebote kaum zu gebrauchen. Auch in den betrachteten Statements wurde mehrfach die Umsetzung von Open Data in Deutschland als bisher unzureichend beschrieben. Weiterhin wurde in den Fokusgruppen kritisiert, dass häufig die bestehenden Angebote potenziellen Nutzerinnen und Nutzern unbekannt seien, da die Bereitstellung über viele verschiedene Portale erfolge und die Potenziale offener Daten zu wenig in den gängigen Studiengängen behandelt würden. Als Ursachen für diese Probleme wurden durch die Teilnehmenden vor allem eine mangelnde Ressourcenausstattung sowie eine unzureichende Kommunikation zwischen Nutzerinnen und Nutzern und Bereitstellerinnen und Bereitstellern identifiziert. Als konkrete Lösungsansätze wurden daher der Ausbau von Ressourcen innerhalb der öffentlichen Verwaltung z. B. in Form von Kompetenzzentren sowie Austauschforen und der systematische Ausbau von Feedbacktools vorgeschlagen. Auch mehrere der betrachteten Statements schlugen den Auf- und Ausbau derartiger Austauschzentren vor.

Als Mehrwert der Nutzung offener Daten im Bereich der öffentlichen Verwaltung wurde sowohl eine Vereinfachung der Kommunikation innerhalb der Verwaltung als auch die Möglichkeit zur Entwicklung von verbesserten Angeboten für die Bürgerinnen und Bürger genannt. Ein weiterer zentraler Vorteil sei der Abbau von Doppelbearbeitung durch verschiedene Behörden. In der Fokusgruppe Wirtschaft wurden vielfältige kommerzielle Nutzungszusammenhänge offener Daten beschrieben, in vielen Bereichen sei man jedoch nach Einschätzung der Teilnehmenden aktuell noch in einer Orientierungsphase. Aus zivilgesellschaftlicher Perspektive könnten offene Verwaltungsdaten nach Einschätzung der Teilnehmenden in vielfältigen gesellschaftlichen Bereichen zur Entwicklung innovativer Lösungen beitragen. Genannt wurden hierzu Beispiele aus den Bereichen

Pandemiebekämpfung, Verkehrs- und Energiewende, Zusammenleben in urbanen Räumen und die Bekämpfung von Fake News.

Im dritten inhaltlichen Block der Fokusgruppen wurde über technische Eigenschaften von offenen Verwaltungsdaten gesprochen, die einen Einfluss auf die Nutzungserfahrungen der Teilnehmenden haben. Angesprochen wurden dabei Aspekte zu Dateiformaten, Strukturierung der Daten sowie die Bedarfe in Hinblick auf die Versionierung und Dokumentation von Daten. Für die Auswertung wurden die getätigten Äußerungen zur Plattform GovData.de ebenfalls diesem Block zugeordnet. Als bevorzugte Dateiformate wurden insbesondere nicht proprietäre maschinenlesbare Formate wie CSV, JSON und XML genannt. Allerdings habe sich nach der Erfahrung der Teilnehmenden die Bereitstellung in geeigneten Formaten inzwischen weitgehend durchgesetzt. Als problematischer wurde hingegen durch die Teilnehmende die uneinheitliche Strukturierung inhaltlich ähnlicher Daten aus verschiedenen Quellen wahrgenommen z. B. eine uneinheitliche Benennung der verwendeten Variablen. Gewünscht wurde daher eine stärkere Standardisierung der offenen Daten in den unterschiedlichen inhaltlichen Bereichen, insbesondere im Bereich der Geodaten. Ebenfalls gewünscht wurden die Bereitstellung von dynamischen Daten mittels einer API und die verstärkte Dokumentation der bereitgestellten Daten und Schnittstellen. Hierzu wurde auch die Beteiligung der Nutzerinnen und Nutzer bei der Priorisierung der Dokumentation angeregt. Die nationale Metadatenplattform GovData.de war nicht allen Teilnehmenden der Fokusgruppen bekannt. Diejenigen, die bereits Erfahrungen mit der Plattform gesammelt hatten, forderten insbesondere eine Vereinheitlichung der dort vorhandenen Metadaten. Bezogen auf die technischen Eigenschaften offener Daten stimmten die ausgewerteten Statements in nahezu allen Punkten mit den in den Fokusgruppen geäußerten Positionen überein und es fanden sich kaum neue Aspekte, die nicht schon in den Fokusgruppen angesprochen wurden. Während in den Fokusgruppen die Bereitstellung offener Daten über Schnittstellen eher am Rande angesprochen wurde, fanden sich in den Statements häufiger Forderungen nach dem Auf- und Ausbau von Schnittstellen zum Abruf offener Daten. Zudem wurden in den betrachteten Statements Qualitätskriterien derartiger Schnittstellen unter dem Begriff „Open API“ diskutiert.

In Hinblick auf rechtliche Aspekte der Open-Data-Nutzung wurde in den Fokusgruppen vor allem über Lizenzierung und Datenschutz gesprochen. Die Änderung des Open-Data-Gesetzes wurde ebenfalls erwähnt, aber nicht tiefergehend diskutiert. Zum Thema Datenschutz wurde durch die Teilnehmenden insbesondere Unsicherheit im Hinblick auf die datenschutzkonforme Nutzung von Daten beschrieben. Zudem wurde die Befürchtung geäußert, dass die Anforderungen an die Anonymisierung die Bereitstellerinnen und Bereitsteller von der Veröffentlichung personenbezogener Daten in anonymisierter Form abhalten könnten. Insgesamt wurden mehr Informationsangebote zum Datenschutz gewünscht sowie die vermehrte Bereitstellung von Mitteln für die Anonymisierung der Daten bei den bereitstellenden Organisationen angeregt. In der Fokusgruppe Zivilgesellschaft wurde über das Thema Lizenzen am ausgiebigsten diskutiert. Dort waren sich die Teilnehmenden einig, dass derzeit zu viele unterschiedliche Lizenztypen die Nutzung offener Daten erschweren. Besonders problematisch sei es, wenn bestehende Standardlizenzen durch Bereitstellerinnen und Bereitsteller von Daten durch individuelle Zusätze ergänzt würden. Die in den Fokusgruppen geäußerten Aspekte zum Datenschutz und zur Wichtigkeit von einheitlichen Lizenzbedingungen fanden sich in ähnlicher Form auch in den betrachteten Statements. Im Hinblick auf das Zweite Open Data Gesetz kritisierten viele der betrachteten Statements, die gesetzlichen Regelungen seien zu weich formuliert. Insbesondere wurde die Einführung eines Rechtsanspruchs auf Open Data häufig gefordert.

Insgesamt kam in allen drei Fokusgruppen eine konstruktive Debatte zu den Nutzungserfahrungen und wahrgenommenen Problemstellungen in Hinblick auf offene Verwaltungsdaten zustande. Positiv fiel

dabei insbesondere auf, dass die Teilnehmenden häufig Verständnis für die Herausforderungen im Zusammenhang mit der Bereitstellung offener Verwaltungsdaten zeigten und in der Lage waren, die von Ihnen wahrgenommenen Problemstellungen sachlich und anhand von anschaulichen Beispielen zu schildern. Durch die Auswertung der Statements konnte eine größere Zahl an Positionen von Nutzerinnen und Nutzern berücksichtigt werden. In den meisten Punkten stimmten die Aussagen und Forderungen in den Statements mit den in den Fokusgruppen geäußerten Positionen überein.

4 Abschließende Betrachtungen

Die vorliegende Studie zu offenen Verwaltungsdaten in Deutschland kann durch eine Reihe von Untersuchungen relevante Ergebnisse für die weitere Ausgestaltung der Struktur offener Verwaltungsdaten in Deutschland liefern, hierbei sind jedoch einige Limitationen der Studie zu beachten, diese werden im Folgenden kurz dargestellt. Weiterhin werden als Ergänzung zu der kapitelweise erfolgten Darstellung der einzelnen Untersuchungsergebnisse im Folgenden die Ergebnisse der Analysen noch einmal nach inhaltlichen Aspekten strukturiert und die Erkenntnisse unabhängig der Untersuchungsmethode zusammengefasst. Abschließend wird in einem Ausblick mögliche anschließende Forschungsfragen und Optimierungsansätze für die Verwaltungspraxis erörtert.

4.1 Limitationen der Studie

Empirische Untersuchungen wie die vorliegende unterliegen grundsätzlich Limitationen, sowohl in der Datenerhebung als auch in der Analyse und Aussagekraft der Ergebnisse. Einige dieser Limitationen wurden bereits in den jeweiligen Kapiteln der einzelnen Erhebungen diskutiert. Für die Zusammenführung der Ergebnisse ergeben sich jedoch weitere Limitationen.

Erstens beziehen sich die einzelnen Erhebungen bzw. Untersuchungen jeweils auf unterschiedliche begrenzte Zeiträume. Zweitens bezieht sich der erste Teil der Erhebung (Logfiles, Metadaten, Matomo-Daten und Onlineumfrage) lediglich auf die zentrale Plattform offener Verwaltungsdaten in Deutschland, GovData.de. Die Fokusgruppen sowie die untersuchten Statements hingegen befassten sich mit den Positionen von Nutzerinnen und Nutzern aller offenen Verwaltungsdaten in Deutschland. Insgesamt ist die Studie eher auf die Erfahrungen auf Bundesebene angelegt; spezifische Nutzungs- oder Bereitstellungserfahrungen auf Landes- oder Kommunalebene wurden hingegen nicht im Speziellen behandelt. Drittens sind die vorgenommenen empirischen Erhebungen (Onlineumfrage und Fokusgruppen) nicht repräsentativ. Die Erkenntnisse aus diesen Untersuchungen spiegeln primär die Erfahrungen der jeweiligen teilnehmenden Personen wieder und repräsentieren nicht zwangsläufig die Gesamtheit der Nutzerinnen und Nutzer offener Verwaltungsdaten in Deutschland.

Unabhängig von den Limitationen der jeweiligen Erhebungs- bzw. Auswertungsmethoden birgt auch das Untersuchungsfeld an sich Herausforderungen für Forschende. Zum einen ist das Angebot an offenen Daten in Deutschland dezentral organisiert. So sind beispielsweise bisher nicht alle Bundesländer auf der Metadatenplattform GovData.de vertreten und es ist auch nicht davon auszugehen, dass alle kommunalen Open-Data-Angebote auf den Landesplattformen für offene Verwaltungsdaten eingestellt werden. Die in den Fokusgruppen als unübersichtlich angesprochene Vielfalt an Angeboten erschwert auch die Erforschung der Nutzung dieser Angebote. Die Wahl der Plattform GovData.de erscheint zwar aufgrund der Ebenen übergreifenden Gestaltung dieses Angebots sinnvoll, die betrachtete Nutzung von GovData.de spiegelt jedoch nur einen Teil der Nutzung offener Verwaltungsdaten in Deutschland wieder. Eine Betrachtung der Nutzung aller Open Data Plattformen in Deutschland war unter denen gegebenen Umständen bzw. aus Kapazitätsgründen jedoch nicht möglich.

Da es bisher kaum empirische Untersuchungen zu Nutzerinnen und Nutzern offener Verwaltungsdaten in Deutschland gibt, ist der Feldzugang aufgrund fehlender Informationen zu betroffenen Nutzerinnen und Nutzern deutlich erschwert. Zudem ergibt sich aus dem Anspruch der Diskriminierungsfreiheit, dass jede Person potenziell Nutzerin oder Nutzer offener Daten werden kann, ohne sich dabei in irgendeiner Form zu erkennen geben zu müssen, was die Identifizierung weiter erschwert. So blieb

dem Forschungsteam in diesem Fall nur die öffentliche Einladung zu Fokusgruppen oder die öffentliche Schaltung der Onlineumfrage auf GovData.de. Diese Studie hat hierdurch (lediglich) den Anspruch, Einblicke in die Nutzergruppen sowie in unterschiedliche Aspekte von Anwendungsfällen offener Verwaltungsdaten zu bieten. Für weitere Untersuchungen in diesem Bereich wäre die Etablierung von Austauschkanälen zwischen Nutzerinnen und Nutzern und Anbietenden von offenen Daten, wie sie auch in den Fokusgruppen und den betrachteten Statements vorgeschlagen wurde, sehr hilfreich.

Zudem ist insbesondere bei der Interpretation der Ergebnisse dieser Untersuchung zu beachten, dass sich das Forschungsvorhaben auf die direkte Nachnutzung offener Daten konzentrierte. Also auf die Erfahrungen der Nutzerinnen und Nutzer, die die bereitgestellten offenen Verwaltungsdaten selbst auswerten oder weiterverwenden. Diejenigen, die die aus diesen Nutzungsprozessen entstandenen Anwendungen nutzen, wurden hingegen nicht angesprochen.

Ferner sind abhängig vom Zeitraum der Untersuchungen (insbesondere in den Logfiles) die Effekte der Covid-19-Pandemie erkennbar, die sich unter anderem in Schlagwörter-Suchen etc. wiederfinden. Dies spricht dafür, dass der Bedarf an offenen Daten von aktuellen gesellschaftlichen Problemfeldern abhängig ist. Wie lange derartige Effekte über die Zeit anhalten und welche Problemstellungen in Zukunft einen besonderen Bedarf an bestimmten Daten auslösen, ist jedoch bisher unklar und bedarf einer fortlaufenden Auswertung von Daten zur Nutzung der entsprechenden Angebote.

4.2 Zentrale Erkenntnisse

Anwendungsgebiete und Zwecke der Nutzung für unterschiedliche Nutzerinnen und Nutzer

Die Untersuchungen haben gezeigt, dass es viele Anwendungsfälle und vielfältigen Nutzen offener Verwaltungsdaten in Deutschland für unterschiedliche Nutzerinnen und Nutzer gibt. Die Plattform GovData.de wird mit konstanten Besuchszahlen genutzt, sowohl von Besucherinnen und Besuchern, die GovData.de zum ersten Mal besuchen, als auch solchen, die das Portal regelmäßig nutzen. Allerdings war keine deutliche Zunahme an Interaktionen im betrachteten Zeitraum zu erkennen. Ein deutlicher Trend hin zu einer stärkeren Nutzung offener Daten war somit anhand der ausgewerteten Nutzungsdaten nicht erkennbar.

Aus allen drei identifizierten Nutzergruppen (Wirtschaft, Zivilgesellschaft und öffentliche Verwaltung) wurde durch die Teilnehmenden in den Fokusgruppen bereits bestehende Anwendungsgebiete offener Daten beschrieben. Auch die Ergebnisse der Onlineumfrage und der Logfileauswertung lassen auf vielfältige Nutzungszusammenhänge schließen. Es bestätigt sich, dass insbesondere auch die Verwaltung offene Daten im Sinne einer Kommunikation bzw. zum Austausch unter den Behörden nutzt, besonders auch auf Landes- und Kommunalebene. Aus wirtschaftlicher Sicht bieten die offenen Verwaltungsdaten eine Grundlage für innovative Prozesse; so gab knapp die Hälfte der Teilnehmenden an der Umfrage aus dem Anwendungsbereich der privatwirtschaftlichen Nutzung an, dass sie GovData.de zur Entwicklung oder Weiterentwicklung einer Anwendung nutzten. Aus zivilgesellschaftlicher Sicht können offene Verwaltungsdaten die Transparenz staatlichen Handelns erhöhen und damit zu einem positiven Demokratieverständnis beitragen, als weitere zivilgesellschaftliche Anwendungsgebiete konnte die Verkehrs- und Energiewende sowie die Verbesserung des Zusammenlebens in urbanen Räumen identifiziert werden. Die Untersuchungen zeigen somit, dass bereits ein Nutzen von offenen Verwaltungsdaten in Deutschland generiert wurde, der sich in vielfältiger Form bei den identifizierten Nutzergruppen offener Daten manifestiert hat und auf dessen Basis weitere Nutzungsprozesse angestoßen werden können.

Allerdings wurde auch in den Fokusgruppen sowie in den betrachteten Statements mehrfach erläutert, dass die als sehr hoch eingeschätzten Potenziale offener Daten zurzeit noch nicht vollumfänglich verwirklicht werden könnten. Es wurden häufig Einschätzungen geäußert, die darauf hindeuten, dass sich Deutschland noch in der Anfangsphase eines Prozesses in Richtung Datenoffenheit befinde und dass für viele potenzielle Anwendungsgebiete heute noch die passenden Angebote fehlen bzw. die Qualität der vorhandenen Daten noch nicht ausreichend ist.

Herausforderungen bei der Nutzung offener Daten und Qualitätsanforderungen

Herausforderungen im Rahmen der Nutzungsprozesse wurden an verschiedenen Stellen der Studie benannt. Die unzureichende Sichtbarkeit offener Verwaltungsdaten, die uneinheitliche Strukturierung von Daten im Vergleich der unterschiedlichen bereitstellenden Behörden sowie die uneinheitliche Lizenzierung von Daten insbesondere im Vergleich der EU-Länder konnten als zentrale Problemstellung bei der Nutzung offener Daten identifiziert werden. Diese Herausforderungen wurden im Austausch mit den Nutzerinnen und Nutzern identifiziert – die Kommunikation zwischen Nutzerinnen und Nutzern auf der einen Seite und Bereitstellerinnen und Bereitsteller bzw. Portalbetreiberinnen und -betreiber auf der anderen Seite ist dahingehend unerlässlich, um Nutzungsprozesse weiter zu optimieren und Hindernisse zu beseitigen.

Als besonders schwerwiegendes Problem bei der Nutzung offener Daten wurde grundsätzlich das Fehlen passender Angebote benannt, demgegenüber erscheinen die genannten Probleme zur Qualität der bereitgestellten Daten (Formatierung, Dokumentation etc.) sekundär. Auch von den abgefragten Aspekten zur Seite GovData.de wurde das Datenangebot am schlechtesten bewertet, gravierende Probleme mit der Benutzerfreundlichkeit der Seite konnten hingegen nicht festgestellt werden. Einige angemerkte Punkte in Hinblick auf die Nutzerfreundlichkeit wurden von Seiten der Betreibenden bereits überarbeitet, so wurde die Filterung nach regionalem Bezug vereinfacht.

Das in verschiedenen Bereichen beschriebene Problem der fehlenden Standardisierung erfordere nach Einschätzungen der Nutzerinnen und Nutzer verstärkte Abstimmungsprozesse der Datenbereitstellerinnen und -bereitsteller in den unterschiedlichen Behörden. Um eine wirklich bedarfsgerechte Bereitstellung zu ermöglichen, sollten auch die Nutzerinnen und Nutzer in diese Austauschprozesse einbezogen werden. Ein Einbezug der Nutzerinnen und Nutzer könnte auch zu einer zielgenaueren Priorisierung bei der oftmals aufwändigen Aufbereitung von Daten beitragen, insbesondere im Bereich der Anonymisierung von vormals personenbezogenen Daten.

In bestimmten Bereichen habe zumindest punktuell vorhandener Austausch zwischen Nutzerinnen und Nutzern und Bereitstellerinnen und Bereitsteller offener Daten bereits zu einer Verbesserung der Qualität der bereitgestellten Daten beigetragen. Derartige bisher erfolgte sporadische Austauschmöglichkeiten sollten verstetigt und um eine dauerhafte Feedback- und Austauschkultur zwischen Nutzerinnen und Nutzern und den Bereitstellerinnen und Bereitstellern ergänzt werden.

Neben dem fehlenden Austausch wurde mangelndes Wissen bzw. fehlende Motivation bei den bereitstellenden Behörden als Ursache für die genannten Problemstellungen angesehen, jedoch auch fehlende technische und finanzielle Ressourcen für die Bereitstellung, dementsprechend wurde ein Ausbau dieser Ressourcen auf allen Verwaltungsebenen angeregt.

Offene Bedarfe

Durch die unterschiedlichen Erhebungsmethoden ist ein Abgleich zwischen Datenangebot (in diesem Fall das Angebot auf GovData.de) und den Bedarfen der Nutzerinnen und Nutzer, die sich im Rahmen

der Untersuchungen äußern konnten, möglich (eingeschränkt durch die zuvor genannten Limitationen).

Ein besonderer Bedarf an Daten war im Bereich der Gesundheitsdaten, insbesondere in Bezug auf die Covid-19-Pandemie feststellbar. Sowohl die Logfiles und die Ergebnisse der Onlineumfrage als auch die Auswertung der Fokusgruppen und der betrachteten Statements wiesen hierauf hin. Die Auswertung der Metadaten ergab in diesem Bereich ein kleines Angebot auf der Seite GovData.de, sodass in Bezug auf diese Daten von aktuell offenen Bedarfen ausgegangen werden kann. Seit dem Ende der Erhebung der Logfiles wurde das entsprechende Datenangebot in diesem Bereich jedoch bereits erweitert. Die betrachteten Logfiles wiesen ebenfalls auf die besondere Relevanz von Geodaten und Daten zur Bebauungs- und Stadtplanung hin. In Hinblick auf die Bereitstellung von Geodaten wurden in den Fokusgruppen, der Onlineumfrage sowie in den ausgewerteten Statements ebenfalls offene Bedarfe angegeben. Um dies zu ändern, wurde von mehreren Akteuren die Einführung eines Rechtsanspruchs auf offene Daten gefordert.

In anderen Bereichen zeigte sich ein ausgewogeneres Bild in Hinblick auf Angebot und Nachfrage an offenen Daten. So wurden die Detailseiten zu Sozialstrukturdaten vielfach auf GovData.de aufgerufen und sind auf der Seite im vergleichsweise großen Umfang vorhanden.

In verschiedenen Kontexten wurde aus Sicht der Nutzerinnen und Nutzer der Bedarf geäußert, die Kommunikation mit den bereitstellenden Behörden bzw. den Betreiberinnen und Betreibern der Datenportale zu verstärken, auch wurden Schulungs- und Beratungsangebote zur rechtskonformen Nutzung offener Daten gewünscht. Unsicherheiten bestanden dabei in Hinblick auf die geltenden Lizenzbedingungen und die Anforderungen des Datenschutzes.

4.3 Ausblick und Optimierungsempfehlungen

Aufgrund der zentralen Untersuchungsergebnisse können an dieser Stelle zum Abschluss des Berichts einige Empfehlungen für folgende Forschungsvorhaben, zur Optimierung von Nutzungsprozessen und der Ausgestaltung des „Systems“ offener Verwaltungsdaten in Deutschland benannt werden. In der Open-Data-Strategie der Bundesregierung (Stand Juni 2021) werden die im Rahmen der vorliegenden Studie festgestellten Optimierungsbedarfe aus Sicht der Nutzerinnen und Nutzer stellenweise adressiert: Die Potentiale von offenen Verwaltungsdaten zu steigern, die stärkere Vernetzung zwischen Akteuren offener Verwaltungsdaten und die Bereitstellung der Daten zu verbessern, werden in den Zielen der Open-Data-Strategie formuliert und finden sich in den Handlungsfeldern in konkreten Vorhaben zum Teil wieder. Im Folgenden sollen die Optimierungsempfehlungen, die sich aus den Ergebnissen der vorgenommenen Untersuchungen ergeben, auch anhand der bereits formulierten Handlungsfelder der Bundesregierung eingeordnet werden.

Der Wert der offenen Daten ist zum einen von der **Menge und Inhalte der Daten**, zum anderen durch **die Strukturierung und die Bedingungen der Veröffentlichung der Daten**, insbesondere die **zeitnahe** (ausgehend von der Erhebung der Daten) und **kontinuierliche Bereitstellung**, bedingt. Für Innovations- bzw. Wertschöpfungsprozesse sind diese Bedingungen der Bereitstellung in erheblichem Maße relevant. Die Ausweitung der Datenbereitstellung auf die mittelbare Bundesverwaltung und die offenen Forschungsdaten, die im zweiten Open-Data-Gesetz geregelt sind, ist hierbei ein wichtiger Schritt, ebenso wie die geplante Datenangebotsenerweiterung der in der Open-Data-Strategie

benannten Stellen¹⁵². Ob diese Ausweitung auch konsequent umgesetzt wird, bleibt abzuwarten. Die Fortschritte in diesem Bereich sollten in jedem Fall kontinuierlich dokumentiert und die entsprechende Dokumentation veröffentlicht werden. Die Kontinuität und Aktualität der Daten sollte ferner durch die Standardisierung von Veröffentlichungsprozessen in den Behörden gestärkt werden, sodass insbesondere die kombinierte Nutzung verschiedener Angebote vereinfacht wird.

Die Steigerung des Angebots offener Verwaltungsdaten in den beschriebenen Dimensionen setzt **finanzielle, technische und personelle Ressourcen** sowie entsprechendes **Wissen über die Prozesse der effektiven und effizienten Bereitstellung offener Verwaltungsdaten** in den Behörden voraus. Die Unterstützungsleistungen, die in der Open-Data-Strategie der Bundesregierung formuliert und geplant sind, können dazu einen Teil beitragen: Die Einrichtung eines IT-Unterstützungstools¹⁵³ und das Angebot von Trainings- und Schulungsmaßnahmen für die Bundesverwaltung¹⁵⁴ können für den Ausbau der Kompetenzen der zuständigen Behördenmitarbeiterinnen und -mitarbeiter hilfreich sein. Ferner können die Stärkung des CCOD¹⁵⁵, die Einführung von Open-Data-Koordinatorinnen und -Koordinatoren in den Behörden des Bundes (§ 12a Abs. 9 EGovG) oder die Erstellung eines Themenkatalogs zur standardisierten Zuordnung von Datenkategorien¹⁵⁶ die Prozesse der Bereitstellung unterstützen bzw. effektiver zu gestalten und dazu beitragen Hindernisse abzubauen.

Der **beidseitige Austausch zwischen Nutzerinnen und Nutzern und den bereitstellenden Behörden bzw. Portalbetreibenden** ist eine zentrale Empfehlung des Forschungsteams. Die Kommunikation kann dabei in vielfältiger Hinsicht ausgestaltet werden: Einerseits in Form eines Feedbacksystems zu konkreten Datensätzen auf Portalen offener Verwaltungsdaten, Kommunikationsplattform zum Austausch über Datenbedarfe (inhaltlicher Aspekte) von Nutzerinnen und Nutzern oder ähnliche digitale Angebote und andererseits auch in Form der direkten gegenseitigen Kommunikation, sodass bereitstellende Behörden Nutzungsprozesse besser verstehen lernen oder ihrerseits über Reglementierungen und Barrieren der Bereitstellung transparent berichten können. Die geplante Intensivierung des jährlichen Bund-Länder-Austausches („Round Table Open Data“)¹⁵⁷ mit Bund, Ländern sowie relevanten Akteuren aus der Zivilgesellschaft ist dahingehend ein wichtiger Schritt. Eine weitere Möglichkeit zur Beförderung des Austausches ist die Einrichtung eines Open-Data-Instituts, einer Stiftung oder einer ähnlichen Organisation, die eine Teilhabe sowohl für Akteure aus der Verwaltung als auch für Akteure aus anderen Bereichen ermöglicht. Eine derartige Einrichtung könnte sowohl als Plattform für den weiteren Austausch dienen als auch Schulungs- und Beratungsangebote im Bereich Open-Data entwickeln.¹⁵⁸

Die vorgenommenen Untersuchungen und die daraus gewonnenen Ergebnisse stellen nur einen Ausschnitt der Nutzungsprozesse und Bedarfe von Nutzerinnen und Nutzern offener

152 Vgl. Bundesregierung (2021a), S. 20ff.

153 Vgl. Bundesregierung (2021a), S. 19.

154 Vgl. Bundesregierung (2021a), S. 27.

155 Vgl. Bundesregierung (2021a), S. 24.

156 Vgl. Bundesregierung (2021a), S. 27.

157 Vgl. Bundesregierung (2021a), S. 27.

158 Die Prüfung einer Einrichtung von Open Data Instituten ist ebenfalls als Vorhaben in der Open Data Strategie der Bundesregierung formuliert (vgl. Bundesregierung (2021a), S. 27). Dieses könnte unterschiedliche Akteure sowohl bei Prozessen der Aufbereitung als auch der Nutzung unterstützen (vgl. ebd.).

Verwaltungsdaten in Deutschland dar. Eine **kontinuierliche, möglichst umfassende Erhebung von Daten der Nutzungsprozesse und Bedarfe** wird in jedem Fall empfohlen, um weiterhin Erkenntnisse über Nutzungsverhalten und die Bedarfe der Nutzerinnen und Nutzer zu generieren. Die Ergebnisse der Matomo-Analyse sowie der Logfiles können bei einer dauerhaften Auswertung der berücksichtigten Daten besser eingeordnet werden, da in diesem Fall Vergleichszeiträume vorliegen und Trends besser erkennbar werden. Trotz der Limitationen kann die vorliegende Studie einen Beitrag zur Erkennung von Anwendungsfällen und den Zwecken der Nutzung von offenen Verwaltungsdaten in Deutschland liefern.

Literaturverzeichnis

Abella, Alberto / Ortiz-de-Urbina-Criado, Marta / De-Pablos-Heredero, Carmen, The Process of Open Data Publication and Reuse, *Journal of the Association for Information Science and Technology*, 70(3), 2019, S.296-300.

Allianz der Wissenschaftsorganisationen, Stellungnahme der Allianz der Wissenschaftsorganisationen zum Entwurf der EU-Kommission zur Novellierung der Richtlinie über die „Weiterverwendung von Informationen des öffentlichen Sektors“, 2018, online: https://www.leopoldina.org/uploads/tx_leopublication/2018_08_Allianz_RR_PSI-Direktive.pdf (Abrufdatum: 25. Februar 2022).

Antipatterns, Transportkollektiv: Opendata Antipatterns, 2021, online: <https://github.com/transportkollektiv/opendata-antipatterns> (Abrufdatum: 25. Februar 2022).

Bitkom, 10-Punkte für Open Government Data- Open Government Data als ein Schlüssel der Datenökonomie, 2020, online: https://www.bitkom.org/sites/default/files/2020-03/20200331_10_punkte_fur_open_government_data_final.pdf (Abrufdatum: 25. Februar 2022).

Bitkom, Stellungnahme Gesetz zur Änderung des E-Government-Gesetzes und zur Einführung des Gesetzes für die Nutzung von Daten des öffentlichen Sektors, 2021, online: https://www.bmwi.de/Redaktion/DE/Downloads/Stellungnahmen/Stellungnahmen-ODG/bitkom.pdf?__blob=publicationFile&v=2 (Abrufdatum: 25. Februar 2022).

bundDEV, 5-Punkte-Plan „Sofortmaßnahmen Zweites Open Data Gesetz“, 2021, online: <https://github.com/bundesAPI/sofortmassnahmen/blob/main/Vorschlag.md> (Abrufdatum: 25. Februar 2022).

Bundesregierung, Erster Nationaler Aktionsplan 2017 -2019: Grundsteine für offenes Regierungs- und Verwaltungshandeln, 2017, online: <https://www.open-government-deutschland.de/resource/blob/1567548/1591018/e31a83e82c377fad6d60f739b9d03a7c/erster-nap-data.pdf?download=1> (Abrufdatum: 13. April 2022).

Bundesregierung, Zweiter nationaler Aktionsplan 2019 – 2021, 2019a, online: <https://www.open-government-deutschland.de/resource/blob/1567548/1667952/76e3507032b45db327d7427d67e70f66/zweiter-nap-data.pdf?download=1> (Abrufdatum: 13. April 2022).

Bundesregierung, Unterrichtung durch die Bundesregierung: Erster Bericht der Bundesregierung über die Fortschritte der Bereitstellung von Daten (1. Open-Data-Fortschrittsbericht), Drucksache 19/14140, 2019b, online: <https://dserver.bundestag.de/btd/19/141/1914140.pdf> (Abrufdatum: 13. April 2022).

Bundesregierung, Dritter Nationaler Aktionsplan 2021 – 2023, 2021a, online: <https://www.open-government-deutschland.de/resource/blob/1567548/1936828/e887b96e1bf8c85c48bcb1b24d0894cc/dritter-nap-data.pdf?download=1> (Abrufdatum: 13. April 2022).

Bundesregierung, Gesetzentwurf der Bundesregierung: Gesetz zur Änderung des E-Government-Gesetzes und zur Einführung des Gesetzes für die Nutzung von Daten des öffentlichen Sektors, 2021b, online: <https://www.bmwi.de/Redaktion/DE/Downloads/G/gesetzentwurf-aenderung-des-e->

government-gesetzes-und-%20Gesetz-fuer-die-nutzung-von-daten-des-oeffentlichen-sektors.pdf?__blob=publicationFile&v=8 (Abrufdatum: 13. April 2022).

Bundesverband Digitale Wirtschaft e. V., Stellungnahme des Bundesverbands Digitale Wirtschaft e. V. zum gemeinsamen Referentenentwurf des Bundesministeriums für Wirtschaft und Energie und des Bundesministeriums des Inneren, für Bau und Heimat für ein Gesetz zur Änderung des E-Government-Gesetzes und zur Einführung des Gesetzes für die Nutzung von Daten des öffentlichen Sektors, 2020, online:

https://www.bmwi.de/Redaktion/DE/Downloads/Stellungnahmen/Stellungnahmen-ODG/bvdw.pdf?__blob=publicationFile&v=4 (Abrufdatum: 25. Februar 2022).

Bundesvereinigung der kommunalen Spitzenverbände, Stellungnahme zum Entwurf eines Gesetzes zur Änderung des E-Government-Gesetzes und zur Einführung des Gesetzes für die Nutzung von Daten des öffentlichen Sektors, 2021, online:

https://www.bmwi.de/Redaktion/DE/Downloads/Stellungnahmen/Stellungnahmen-ODG/staedtetag.pdf?__blob=publicationFile&v=4 (Abrufdatum: 25. Februar 2022).

Charalabidis, Yannis / Loukis, Euripides / Alexopoulos, Charalampos, Evaluating Second Generation Open Government Data Infrastructure Using Value Models, 47th Hawaii International Conference on System Science, 2014.

Conradie, Peter / Choenni, Sunil, On the barriers for local government releasing open data, *Government Information Quarterly*, 31(1), 2014, S.10-17.

D64 Zentrum für digitalen Fortschritt e. V., Stellungnahme von D64 Zentrum für digitalen Fortschritt e. V. zu dem Entwurf eines Gesetzes zur Änderung des E-Government-Gesetzes und zur Einführung des Gesetzes für die Nutzung von Daten des öffentlichen Sektors, 2021, online:

https://www.bmwi.de/Redaktion/DE/Downloads/Stellungnahmen/StellungnahmenODG/d64.pdf__blob=publicationFile&v=4 (Abrufdatum: 25. Februar 2022).

Danneels, Lieselot / Viaene, Stijn / Van den Bergh, Joachim, Open data platforms: Discussing alternative knowledge epistemologies, *Government Information Quarterly*, 34(3), 2017, S.365-378.

Dbb, Betreff: Referentenentwurf Gesetz zur Änderung des E-Government-Gesetzes und zur Einführung des Gesetzes für die Nutzung von Daten des öffentlichen Sektors, 2021, online:

https://www.bmwi.de/Redaktion/DE/Downloads/Stellungnahmen/Stellungnahmen-ODG/dbb.pdf?__blob=publicationFile&v=4 (Abrufdatum: 25. Februar 2022).

Deutscher EDV-Gerichtstag e. V., DEUTSCHER EDV-GERICHTSTAG e. V. Stellungnahme zum Entwurf eines Gesetzes zur Änderung des E-Government-Gesetzes und zur Einführung des Gesetzes für die Nutzung von Daten des öffentlichen Sektors, 2021, online:

https://www.bmwi.de/Redaktion/DE/Downloads/Stellungnahmen/Stellungnahmen-ODG/edv.pdf?__blob=publicationFile&v=4 (Abrufdatum: 25. Februar 2022).

Deutsches Jugendinstitut, Stellungnahme des DJI zum Referentenentwurf „Gesetz zur Änderung des E-Government-Gesetzes und zur Einführung des Gesetzes für die Nutzung von Daten des öffentlichen Sektors“ vom 17.12.2020, 2021, online:

https://www.bmwi.de/Redaktion/DE/Downloads/Stellungnahmen/Stellungnahmen-ODG/dji.pdf?__blob=publicationFile&v=4 (Abrufdatum: 25. Februar 2022).

EDV-Gerichtstag e. V., Stellungnahme zum Entwurf eines Gesetzes zur Änderung des E-Government-Gesetzes und zur Einführung des Gesetzes für die Nutzung von Daten des öffentlichen Sektors, 2021, online: https://www.bmwi.de/Redaktion/DE/Downloads/Stellungnahmen/Stellungnahmen-ODG/edv.pdf?__blob=publicationFile&v=4 (Abrufdatum: 25. Februar 2022).

European Data Portal, Open Data Maturity Report 2021, Publications Office of the European Union. 2021a, online: https://data.europa.eu/sites/default/files/landscaping_insight_report_n7_2021_0.pdf (Abrufdatum: 12. Januar 2022).

European Data Portal, Open Data Maturity Report 2021: Germany, 2021b, online: https://data.europa.eu/sites/default/files/country-factsheet_germany_2021.pdf (Abrufdatum: 12. Januar 2022).

Gásco-Hérendez, Mila / Martin, Erika G. / Reggi, Luigi / Pyo, Sunyoung / Luna-Reyes, Luis F., Promoting the use of open government data: Cases of training and engagement, *Government Information Quarterly*, 35(2), 2018, S.233-242.

Harrison, Teresa M. / Pardo, Theresa A. / Cook, Meghan, Creating Open Government Ecosystem: A Research and Development Agenda, *Future Internet*, 4(4), 2012, S.900-928.

Höchtel, Johann / Kaltenböck, Martin / Schössböck, Judith / Thurner, Thomas, Open Government Data: Potentiale, Risiken und Hürden, 41. Jahrestagung der Gesellschaft für Informatik, 4.-7. Oktober 2011, Berlin 2011.

Janssen, Marijn / Charalabidis, Yannis / Zuiderwijk, Anneke, Benefits, Adoption Barriers and Myths of Open Data and Open Government, *Information System Management*, 29(4), 2012, S.258-268.

Johnson, Peter A. / Sieber, Renee / Scassa, Teresa / Stephens, Monica / Robinson, Pamela, The Cost(s) of Geospatial Open Data, *Transaction in GIS*, 21(3), 2017, S.434-445.

Konrad-Adenauer-Stiftung, Btr.: Entwurf eines Gesetzes zur Änderung des E-Government-Gesetzes und zur Einführung des Gesetzes für die Nutzung von Daten des öffentlichen Sektors- Datennutzungsgesetz (DNG), 2021, online: https://www.bmwi.de/Redaktion/DE/Downloads/Stellungnahmen/Stellungnahmen-ODG/kas.pdf?__blob=publicationFile&v=4 (Abrufdatum: 25. Februar 2022).

Ksoll, Wolfgang / Schildhauer, Thomas / Beck, Annalies, Open Data – Wertschöpfung im digitalen Zeitalter. Bertelsmann Stiftung, 2017, online: https://www.bertelsmann-stiftung.de/fileadmin/files/Projekte/Smart_Country/OpenData_2017_final.pdf (Abrufdatum: 13. April 2022).

Kucera, Jan / Chlapek, Dusan, Benefits and Risks of Open Government Data, *Journal of System Integration*, 5(1), 2014, S.30-41.

Lassinantti, Josefin / Ståhlbröst, Anna / Runardotter, Mari, Relevant social groups for open data use and engagement. *Government Information Quarterly*, 36(1), 2019, S.98-111.

Litta, Henriette, 2. Open Data-Gesetz: Die größte Chance wurde nicht ergriffen. *Netzpolitik.org*, 2021, online: <https://netzpolitik.org/2021/2-open-data-gesetz-die-groesste-chance-wurde-nicht-ergriffen/> (Abrufdatum: 25. Februar 2022).

Martin, Sébastien / Foulonneau, Muriel / Turki, Slim / Ihadjadene, Madjid, Risk Analysis to Overcome Barriers to Open Data, *Electronic Journal of e-Government*, 11(1), 2013, S.348-359.

Matomo, How is a ‚visit‘ defined in Matomo?, 2020a, online:
https://matomo.org/faq/general/faq_36/ (Abrufdatum: 13. April 2022).

Matomo, How is a ‚unique visitor‘ counted in Matomo?, 2020b, online:
https://matomo.org/faq/general/faq_43/ (Abrufdatum: 13. April 2022).

Matomo, Metrics: Bounce Rate, 2020c, online: <https://glossary.matomo.org/> (Abrufdatum: 13. April 2022).

Max-Planck-Institut für Innovation und Wettbewerb, Stellungnahme zum Entwurf eines Gesetzes zur Änderung des E-Government- Gesetzes und zur Einführung des Gesetzes für die Nutzung von Daten des öffentlichen Sektors vom 17.12.2020, 2021, online:
https://www.bmwi.de/Redaktion/DE/Downloads/Stellungnahmen/Stellungnahmen-ODG/ip-mpg.pdf?__blob=publicationFile&v=4 (Abrufdatum: 25. Februar 2022).

mFUND, Zugang zu Behördendaten für Digitalisierungsprojekte des mFUND- Eine Veröffentlichung der mFUND-Begleitforschung des WIK, 2019, online:
https://www.wik.org/fileadmin/mFUND_VF/mFUND-WIK-Studie-Zugang_Behoerdendaten_AF_Open_Data_und_Behoerden.pdf (Abrufdatum: 25. Februar 2022).

Open Knowledge Foundation Deutschland e. V., Stellungnahme zur vorgelegten Änderung des E-Government-Gesetzes und zur Einführung des Gesetzes für die Nutzung von Daten des öffentlichen Sektors, 2021, online:
https://www.bmwi.de/Redaktion/DE/Downloads/Stellungnahmen/Stellungnahmen-ODG/okf.pdf?__blob=publicationFile&v=6 (Abrufdatum: 25. Februar 2022).

Open Legal Data, Open Legal Data Stellungnahme zum Referentenentwurf eines Gesetzes zur Änderung des E-Government-Gesetzes und zur Einführung des Gesetzes für die Nutzung von Daten des öffentlichen Sektors, 2021, online:
https://www.bmwi.de/Redaktion/DE/Downloads/Stellungnahmen/Stellungnahmen-ODG/old.pdf?__blob=publicationFile&v=4 (Abrufdatum: 25. Februar 2022).

Rat für Sozial- und Wirtschaftsdaten, Stellungnahme des RatSWD zum gemeinsamen Referentenentwurf des Bundesministeriums für Wirtschaft und Energie und des Bundesministeriums des Innern, für Bau und Heimat für ein Gesetz zur Änderung des E-Government-Gesetzes und zur Einführung des Gesetzes für die Nutzung von Daten des öffentlichen Sektors, 2021, online:
https://www.bmwi.de/Redaktion/DE/Downloads/Stellungnahmen/Stellungnahmen-ODG/ratswd.pdf?__blob=publicationFile&v=2 (Abrufdatum: 25. Februar 2022).

RELX, Beitrag zur Diskussion über den Referentenentwurf zum zweiten Open-Data-Gesetz und Datennutzungsgesetz des Bundesministeriums für Wirtschaft und Energie und des Bundesministeriums des Inneren, für Bau und Heimat, 2021, online:
https://www.bmwi.de/Redaktion/DE/Downloads/Stellungnahmen/Stellungnahmen-ODG/relx.pdf?__blob=publicationFile&v=4 (Abrufdatum: 25. Februar 2022).

Riley, Jenn, Understanding Metadata: What is metadata and what is it for, National Information Standards Organization, Baltimore 2017.

Ruijter, Erna / Grimmelikhuijsen, Stephan / Hogan, Michael / Enzerink, Sem / Ojo, Adegboyega / Meijer, Albert, Connecting societal issues, users and data: Scenario-based design of open data platforms, *Government Information Quarterly*, 34(3), 2017, S.470-480.

Sadiq, Shazia / Indulska, Marta, Open data: Quality over quantity. *International Journal of Information Management*, 37(3), 2017, S.150-154.

Seibel, Benjamin, *Open Data in der Praxis: Bereitsteller und Anwender offener Daten in Berlin*, Technologiestiftung, Berlin 2016.

Smart City Dialog, Positionen und Diskussionsbedarf zur Umsetzung der Richtlinie über offene Daten und die Weiterverwendung von Informationen des öffentlichen Sektors (EU) 2019/1024, 2020, online: https://www.smart-city-dialog.de/wp-content/uploads/2021/03/PSI-Richtlinie_Themenpapier_bf_n.pdf (Abrufdatum: 25. Februar 2022).

Smith, Göran / Sandberg, Johan, Barriers to innovating with open government data: Exploring experiences across service phases and user types, *Information Polity*, 23(3), 2018, S.249-265.

Susha, Iryna / Grönlund, Åke / Janssen, Marijn, Organizational measures to stimulate user engagement with open data, *Transforming Government: People, Process and Policy*, 9(2), 2015, S.181-206.

Verband kommunaler Unternehmen, Stellungnahme: Zum Entwurf eines Gesetzes zur Änderung des E-Government-Gesetzes und zur Einführung des Gesetzes für die Nutzung von Daten des öffentlichen Sektors, 2021, online: https://www.vku.de/fileadmin/user_upload/Verbandsseite/Themen/Digitalisierung/210112_VKU-SN_DNG_RefE.pdf (Abrufdatum: 25. Februar 2022).

Verdi, Stellungnahme zum Entwurf eines Gesetzes zur Änderung des E-Government-Gesetzes und zur Einführung des Gesetzes für die Nutzung von Daten des öffentlichen Sektors, 2021, online: https://www.bmwi.de/Redaktion/DE/Downloads/Stellungnahmen/Stellungnahmen-ODG/verdi.pdf?__blob=publicationFile&v=4 (Abrufdatum: 25. Februar 2022).

Weber, Nicholas, Integrating User Feedback with Open Data Quality Models, 80th Annual Meeting of the Association for Information Science & Technology, Washington, DC 2017, S.824-826.

Welle Donker, Frederika / van Loenen, Bastiaan, Sustainable Business Models for Public Sector Open Data Providers, *eJournal of eDemocracy and Open Government*, 8(1), 2016, S.28-61.

Wiebe, Andreas, *Open Data in Deutschland und Europa*. Konrad-Adenauer-Stiftung, 2020, online: <https://openall.info/open-data-in-deutschland-und-europa/> (Abrufdatum: 25. Februar 2022).

Zeleti, Fatemeh A. / Ojo, Adegboyega, Open data value capability architecture, *Information System Frontiers*, 19(2), 2017, S.337-360.

Zuiderwijk, Anneke / Janssen, Marijn, Participation and Data Quality in Open Data use: Open Data Infrastructures Evaluated, 2015, S.1-9.

Zuiderwijk, Anneke / Janssen, Marijn / Davis, Chris, Innovation with open data: Essential elements of open data ecosystem, *Information Polity*, 19(1), 2014, S.17-33.

Zuiderwijk, Anneke / Janssen, Marijn / Choenni, Sunil / Meijer, Ronald / Alibaks, Roexsana Sheikh, Socio-technical Impediments of Open Data, Electronic Journal of e-Government,10(2), 2012, S.156-172.

