

Fraunhofer FOKUS
im Auftrag der Senatsverwaltung für
Wirtschaft, Technologie und Forschung in Berlin

Dr. Wolfgang Both (Hrsg.)
Prof. Dr. Ina Schieferdecker (Hrsg.)

Berliner Open Data-Strategie

Organisatorische, rechtliche und
technische Aspekte offener Daten in Berlin.

Konzept, Pilotsystem und Handlungsempfehlungen.



Fraunhofer FOKUS
im Auftrag der Senatsverwaltung für
Wirtschaft, Technologie und Forschung in Berlin

Dr. Wolfgang Both (Hrsg.)
Prof. Dr. Ina Schieferdecker (Hrsg.)

Berliner Open Data-Strategie

Organisatorische, rechtliche und
technische Aspekte offener Daten in Berlin.

Konzept, Pilotsystem und Handlungsempfehlungen

Autoren:

Dr. Nils Barnickel	Fraunhofer-Institut für Offene Kommunikationssysteme (FOKUS)
Dr. Matthias Flügge	Fraunhofer FOKUS
Daniel Hanke	Fraunhofer FOKUS
Dr. Edzard Höfig	Fraunhofer FOKUS
Jens Klessmann	Fraunhofer FOKUS
Florian Marienfeld	Fraunhofer FOKUS
Prof. Dr. Ina Schieferdecker	Fraunhofer FOKUS
Jan Ziesing	Fraunhofer FOKUS

Weitere Bearbeiter:

Hartmut Bömermann	Amt für Statistik Berlin-Brandenburg
Dr. Wolfgang Both (Gesamtprojektleitung)	Senatsverwaltung für Wirtschaft, Technologie und Forschung in Berlin
Julia Franke	Praktikantin der Senatsverwaltung für Wirtschaft, Technologie und Forschung in Berlin
Olaf Franke	Senatsverwaltung für Inneres und Sport
Andreas Mängel	BerlinOnline Stadtportal GmbH & Co. KG

Berlin, Januar 2012

Kontaktadresse:

Fraunhofer-Institut für Offene Kommunikationssysteme (FOKUS)
Kaiserin-Augusta-Allee 31, 10589 Berlin

Telefon +49 30 3463 7200
Telefax +49 30 3463 8000
E-Mail info@fokus.fraunhofer.de
URL www.fokus.fraunhofer.de

Bibliografische Information der Deutschen Nationalbibliothek:

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.
ISBN 978-3-8396-0368-0

Druck und Weiterverarbeitung:
IRB Mediendienstleistungen
Fraunhofer-Informationszentrum Raum und Bau IRB, Stuttgart

Für den Druck des Buches wurde chlor- und säurefreies Papier verwendet.

© by FRAUNHOFER VERLAG, 2012
Fraunhofer-Informationszentrum Raum und Bau IRB
Postfach 800469, 70504 Stuttgart
Nobelstraße 12, 70569 Stuttgart
Telefon +49 711 970-2500
Telefax +49 711 970-2508
E-Mail verlag@fraunhofer.de
URL <http://verlag.fraunhofer.de>

Der Inhalt dieses Dokuments steht unter der CC BY-Lizenz (Creative Commons Namensnennung 3.0 Deutschland). Als Quelle muss <http://s.fhg.de/BerlinerOpenDataStrategie> genannt werden. Informationen zur Lizenz finden sich unter <http://creativecommons.org/licenses/by/3.0/de>.

Inhaltsverzeichnis

Zusammenfassung	11
1 Einleitung	15
1.1 Hintergrund und Motivation	15
1.2 Zielsetzung und Vorgehen	16
1.3 Ergebnisse, Handlungsempfehlungen und Stufenplan	17
2 Bestands- und Potenzialanalyse	19
2.1 Open Data: Grundlagen und Begriffe	19
2.1.1 Definition und Potenziale von Open Data und Open Government Data	20
2.1.2 Verständnis von Open Data und Open Government Data	21
2.1.3 Open Government Data, Public Sector Information und Informa- tionsfreiheit	23
2.1.4 Potenziale offener Daten	24
2.2 Situation offener Verwaltungsdaten in Deutschland und im Ausland	27
2.2.1 Situation in Deutschland	28
2.2.2 Bereitstellung von Verwaltungsdaten im Land Berlin	32
2.2.3 Internationale Beispiele für die Bereitstellung offener Verwal- tungsdaten	37
2.2.4 Beispiele für die wirtschaftliche Weiterverwendung von Verwal- tungsdaten	39
2.3 Zusammenfassende Beurteilung	40
3 Organisationsanalyse	45
3.1 Senatsverwaltung für Bildung, Wissenschaft und Forschung	46
3.2 Amt für Statistik Berlin-Brandenburg	49
3.3 Zentral- und Landesbibliothek Berlin	53
3.4 Web-AG der Umweltämter von Land und Bezirken	55
3.5 Fazit	57
4 Architektur	59
4.1 Generischer Soll-Prozess mit Abbildung auf Systemkomponenten	60
4.2 Anwendungsfälle	62
4.2.1 Akteur: Datenbereitsteller	63
4.2.2 Akteur: Plattformbetreiber	65

4.2.3	Akteur: direkter Datennutzer	66
4.2.4	Akteur: indirekter Datennutzer	68
4.2.5	Akteur: Datenaufbereiter	69
4.3	Ausbaustufen der Architektur	71
4.3.1	Domäne Behörden	72
4.3.2	Berliner Open Data-Plattform	72
4.3.3	Datennutzer und externe App Stores	73
4.3.4	Umsetzung der Architektur	74
5	Geschäftsmodelle	75
5.1	Produkte und Kunden	75
5.2	Monetarisierungspotentiale	76
5.3	Mögliche Geschäftsmodelle für Berlin	78
5.3.1	Option I: Betrieb als Costcenter der Stadt Berlin	78
5.3.2	Option II: Private Nutzung kostenfrei, kommerzielle Nutzung kostenpflichtig	79
5.3.3	Option III: Marktplatzmodell – Plattform mit Transaktionsgebühren	81
5.3.4	Zusatzgeschäft	83
5.4	Fazit	83
6	Rechtliche Rahmenbedingungen	85
6.1	Rohdatenbeschaffung aus Sicht einer natürlichen oder juristischen Person	86
6.1.1	Informationszugang und Akteneinsicht in Berlin gemäß BerlIFG .	86
6.1.2	Die Informationsweiterverwendung gemäß IWG	91
6.1.3	Zugang nach dem Umweltinformationsgesetz	93
6.1.4	Zugang nach dem Verbraucherinformationsgesetz	94
6.1.5	Geodatenzugangsgesetz Berlin	94
6.1.6	Ausblick auf das eGovernment- und Organisationsgesetz in Berlin	95
6.2	Behördliche Veröffentlichung und Lizenzierung von PSI	95
6.2.1	Datenschutzrechtliche Vorgaben für die Veröffentlichung von PSI .	95
6.2.2	Daten- und Datenbanklizenzierung	97
6.2.3	Zusätzliche Nutzungsbedingungen	102
6.3	Datenschutz der Nutzer der Open Data-Plattform	103
6.3.1	Anwendbarkeit des TMG auf die Open Data Internet-Plattform . .	103
6.3.2	Vertragsverhältnis zwischen Nutzer und der ODP	103
6.3.3	Personenbezogene Daten	103
6.3.4	Personenbezogene Daten der Datenanbieter	104
6.4	Zusammenfassung und Fazit	104
7	Pilotsystem	107
7.1	Datenportal	108
7.2	Datenregister	112
7.3	Metadaten für den Berliner Datenkatalog	112
7.3.1	Analyse existierender Datenkataloge	113

7.3.2	Metadatenschema, Kategorien und APIs für Berlin	114
7.4	Umfang und Entwicklung bei Datensätzen und Apps	115
7.5	Datenregister im Einsatz, Handreichung, Rückmeldung	116
7.6	Nutzerzahlen	116
7.7	Medienresonanz	118
7.8	Wartungsaufgaben während des Pilotbetriebs	118
7.9	Fazit und nächste Schritte	119
8	Handlungsempfehlungen	121
8.1	Politische Handlungsempfehlungen	122
8.1.1	Politisches Bekenntnis zu offenen Daten in und aus Berlin	122
8.1.2	Generelle Informationsstrategie	122
8.1.3	Politisches Bekenntnis zum Berliner Datenportal	123
8.1.4	Gesamtverantwortlicher für offene Daten und Ansprechpartner zu den Daten	123
8.1.5	Stabile und langfristige Bereitstellung der Daten	123
8.1.6	Ausbau des Datenportals	124
8.1.7	Föderation des Berliner Datenportals mit anderen Portalen	124
8.2	Organisatorische Handlungsempfehlungen	124
8.2.1	Anschlussfähigkeit der Open Data-Strategie	124
8.2.2	Richtlinien zur Auswahl der Daten	125
8.2.3	Wohldefinierte Prozesse für das Datenportal	125
8.2.4	Weiterbildungsmaßnahmen zu offenen Daten	125
8.2.5	Vielfältige Interaktionen mit den Nutzern des Berliner Datenportals	126
8.2.6	Vorschlagsmöglichkeiten für Datensätze	126
8.2.7	Publizierte Daten in das Datenportal stellen	126
8.2.8	Beschlüsse und Protokolle in das Datenportal stellen	127
8.2.9	Daten landeseigener Unternehmen in das Datenportal stellen	127
8.3	Rechtliche Handlungsempfehlungen	127
8.3.1	Einheitliche und einfache Lizenz- und Nutzungsbestimmungen	127
8.3.2	Direkter Zugang zu den Lizenz- und Nutzungsbestimmungen	128
8.3.3	Nachvollziehbarkeit der Datenquellen	128
8.4	Technische Handlungsempfehlungen	128
8.4.1	Standards und Sicherheit des Berliner Datenportals	128
8.4.2	Kompatibilität und Integrierbarkeit der Daten	129
8.4.3	Qualitätssicherung der Daten	129
8.4.4	Dienste und Werkzeuge für den Umgang mit den Daten	130
8.5	Stufenplan	130
	Abkürzungen	131
	Glossar	135
	Quellenverzeichnis	141

Anhang	147
A Interviewleitfaden der Organisationsanalyse	147
B FAQ – Häufig gestellte Fragen	151
C Linksammlung	154
D Metadaten-Schema und API	156
E Metadaten-Formular	164
F Handreichung Datenregister	167

Abbildungsverzeichnis

3.1	Schema des Ist-Veröffentlichungsprozesses bei der SenBWF	46
3.2	Schema des Soll-Veröffentlichungsprozesses bei der SenBWF	47
3.3	Schema des Ist-Veröffentlichungsprozesses beim AfS	50
3.4	Schema des Soll-Veröffentlichungsprozesses beim AfS	51
3.5	Schema des Ist-Veröffentlichungsprozesses bei der ZLB	54
3.6	Schema des Soll-Veröffentlichungsprozesses bei der Web-AG	56
4.1	Generischer Sollprozess mit Abbildung auf die Plattformkomponenten . .	60
4.2	Grobarchitektur und Schnittstellen	71
7.1	Startseite des Datenportals daten.berlin.de	108
7.2	Katalogseite des Datenportals daten.berlin.de	109
7.3	Detailseite des Datenportals daten.berlin.de	110
7.4	Nutzungsstatistik des Datenportals daten.berlin.de	117

Tabellenverzeichnis

2.1	Abgrenzung der Begriffe Open Government Data	24
2.2	SWOT-Analyse der Rahmenbedingungen im Land Berlin	43
3.1	Groberfassung des Istprozesses bei der SenBWF	47
3.2	Groberfassung des Istprozesses beim AfS	50
3.3	Groberfassung des Istprozesses bei der ZLB	54
4.1	Anwendungsfälle für den Akteur Datenbereitsteller	64
4.2	Anwendungsfälle für den Akteur Plattformbetreiber	65
4.3	Anwendungsfälle für den Akteur direkter Datennutzer	67
4.4	Anwendungsfälle für den Akteur indirekter Datennutzer	69
4.5	Anwendungsfälle für den Akteur Datenaufbereiter	70
5.1	Produkte und Kunden offener Daten	76
5.2	Monetarisierungspotenziale offener Daten	77
7.1	Nutzungsstatistik des Datenportals daten.berlin.de nach Bereichen	118
7.2	Nutzungsstatistik des Datenportals daten.berlin.de nach Datensätzen . . .	119

Zusammenfassung

Nicht zuletzt der unerwartete Erfolg der Piratenpartei bei den Wahlen zum Berliner Abgeordnetenhaus und zu den Bezirksparlamenten hat deutlich gemacht, dass Forderungen nach offenem Regierungs- und Verwaltungshandeln, Öffnung von Datenbeständen und aktivem Bereitstellen von Regierungs- und Verwaltungsinformationen über Gesetzestexte und Verordnungen hinaus nicht irgendwo vom Rand der Gesellschaft erhoben werden. Bürgerbegehren, Bürgerbeteiligung, Bürgerproteste, Bürgerhaushalte zeigen, dass diese Themen in der Mitte der Gesellschaft angekommen sind. Dass diese Erwartungen mehrheitsfähig sind, hat eine SAS-Umfrage¹ vom Sommer 2010 eindrücklich nachgewiesen, in der sich 88 % der Befragten für eine Veröffentlichung nicht personenbezogener Informationen durch die Behörden aussprachen und 81 % darin eine Chance für mehr Partizipation sahen.

Alle demokratischen Parteien haben die Themen Transparenz und Open Government in ihre Wahlprogramme aufgenommen, die aus den letzten Wahlen hervorgegangene Berliner Regierungskoalition aus SPD und CDU kündigt in ihrem Koalitionsvertrag die Fortsetzung und den Ausbau der Open Data-Initiative des Landes an.

Die Berliner Open Data-Strategie

Aufbauend auf Vorarbeiten, in Reflexion einer Reihe von Interviews mit Akteuren im Land Berlin und in Weiterentwicklung existierender Open Data-Konzepte wird in diesem Dokument die Berliner Open Data-Strategie formuliert. Sie umfasst folgende Elemente:

- Bestands- und Potenzialanalyse: Zur Fundierung der Strategie werden der aktuelle Stand und Potenziale offener Daten im Land Berlin untersucht. Es wird eine wissenschaftlich hergeleitete Einschätzung des Status quo innerhalb Berlins und im Vergleich zu anderen Gebietskörperschaften im In- und Ausland gegeben.
- Organisationsanalyse: Es werden ausgewählte organisatorische Prozesse in der Verwaltung bezüglich des Umgangs mit öffentlichen Daten beleuchtet. Darüber hinaus werden die notwendigen Änderungen im Hinblick auf die organisatorischen Gegebenheiten, die Qualitätssicherung, eventuelle zusätzliche Aufwände und die Ausbildung der Mitarbeiter dargestellt. Zudem werden organisatorische Barrieren und Risiken identifiziert, die mit der Umsetzung einer Strategie zur Freigabe öffentlicher Daten einhergehen können.
- Architekturvorschlag für eine Open Data-Plattform: Es werden Architekturempfehlungen und eine technische Konzeption für eine Berliner Open Data-Plattform erarbeitet. Es wird eine mit möglichen Datenbereitstellern, -bearbeitern,

¹<http://www.sas.de/open>.

- verarbeitern und -nutzern abgestimmte Grobspezifikation beschrieben, die als Grundlage für die Pilotierung des Berliner Open Data-Portals genutzt wurde.
- Analyse möglicher Geschäftsmodelle: Es werden unterschiedliche Betriebskonzepte des Datenportals definiert und bewertet. Zudem werden Finanzierungsmöglichkeiten für offene Daten untersucht, die zusätzliche Erlöse und Finanzierungsbeiträge für die Erstellung und den Betrieb der Berliner Open Data-Plattform liefern können, sodass ein rascher Ausbau des Berliner Open Data-Angebots ermöglicht wird.
 - Rechtliche Rahmenbedingungen: Basierend auf einer gründlichen rechtlichen Prüfung der relevanten Gesetze, wird der rechtliche Rahmen für Open Data in Berlin bestimmt und ein aktueller Überblick über die Rechtslage gegeben. Es werden geeignete Lösungs- und Novellierungsvorschläge unterbreitet, die eine rechtliche Umsetzung von Open Data in Berlin ermöglichen sollen.
 - Open Data-Pilotsystem mit Pilotdaten: Aufbauend auf den Architekturempfehlungen für eine Berliner Open Data-Plattform, wurde ein Datenportal für Berlin errichtet, das trotz kurzer Laufzeit schon jetzt den Nachweis der Relevanz und der Potenziale von Open Data für die Stadt führt. Es werden praktische Erfahrungen bei der Bereitstellung und Weiterverarbeitung von öffentlichen Daten gesammelt und dargestellt.
 - Handlungsempfehlungen und Stufenplan: Die Handlungsempfehlungen fassen die Ergebnisse des Projekts in Form von politischen, organisatorischen, rechtlichen und technischen Handlungsempfehlungen zusammen und bestimmen einen Stufenplan für den weiteren Ausbau des Datenangebots in Berlin.

Das Berliner Datenportal

Im Rahmen des Projekts wurde gemeinsam mit der Senatsverwaltung für Wirtschaft, Technologie und Forschung, der Senatsverwaltung für Inneres und Sport, dem Amt für Statistik Berlin-Brandenburg und der BerlinOnline Stadtportal GmbH ein Pilotsystem für ein Berliner Datenportal entwickelt. Das Datenportal bildet eine zentrale Anlaufstelle für Daten des Landes Berlin. Die hier gesammelten Datensätze sind maschinenverarbeitbar und frei lizenziert und ermöglichen so eine größere Transparenz bei Verwaltungsvorgängen, bieten u. a. Informationen zu Infrastruktur- und Umweltaspekten in der Stadt und laden zur Weiternutzung und Weiterverarbeitung durch Dritte ein.

Das Portal ist technisch zweigeteilt: Die Oberfläche bildet ein Drupal Content Management-System und wird von der BerlinOnline Stadtportal GmbH umgesetzt und betrieben. Als Backend kommt das Comprehensive Knowledge Archive Network (CKAN) als Datenkatalog zum Einsatz. Es hält die Metadaten über die Datensätze bereit. Das Berliner CKAN wurde von Fraunhofer FOKUS konzipiert und eingerichtet und wird zunächst auch von FOKUS unterhalten.

Das separate Backend bietet verschiedene Möglichkeiten, deren wichtigste die Föderation ist. Im Zusammenhang mit Open Data bedeutet das: Das Berliner CKAN kann einerseits Inhalte untergeordneter Instanzen konzentrieren, beispielsweise CKAN-Datenportale der Bezirke oder sogar solche, die in einzelnen Institutionen eingesetzt werden. Zudem kann Fraunhofer FOKUS die Berliner Daten durch CKAN dem Bund bereitstellen, sobald dieser ein vergleichbares Portal betreibt. Schon heute können die Daten außerdem auf europäischer Ebene z.B. auf publicdata.eu veröffentlicht werden. Damit diese Föderation sinnvoll funktionieren kann, ist ein durchdachtes Metadatenschema notwendig. Das heißt, alle Beteiligten müssen die gleichen Informationen über Datensätze pflegen, und zwar in kompatiblen Formaten. Fraunhofer FOKUS hat daher auf der Basis von Erfahrungen existierender Portale aus Großbritannien sowie den Städten London und Wien sowie einer Umfrage unter weiteren europäischen Großstädten ein Metadatenschema entwickelt, das im Rahmen des EU-Projektes »Open Cities« zudem in den Metropolen Barcelona, Amsterdam und Paris zum Einsatz kommen wird.

Derzeit bietet das Portal den Zugriff auf rund 60 Datensätze aus fast 15 verschiedenen Datenkategorien, nicht nur der öffentlichen Hand, sondern auch bereits von anderen Organisationen. Die Nutzer von <http://daten.berlin.de> dürfen die Daten für alle Zwecke verwenden, solange nicht gegen die Nutzungsbedingungen und gegen Rechte Dritter verstoßen wird. Wo möglich werden die Datenbestände unter die Lizenz »Creative Commons Namensnennung 3.0 Deutschland« (CC BY) gestellt. Des Weiteren werden Applikationen angeboten, die auf den Daten aufbauen, diese auswerten und/oder visualisieren. Gerade bei den Applikationen zeigt sich die Stärke des Ansatzes, in einer Stadt Daten offen zur Verfügung zu stellen: Bereits kurz nach der Öffnung des Portals wurden Applikationen hochgeladen, die neue Einsichten ermöglichen – wie z.B. über den aktuellen Stand der Ozonbelastung in Berlin. Für Hinweise, Vorschläge und Wünsche zu den Daten, Applikationen und generell um in Interaktion mit Nutzern zu treten und mögliche Weiterentwicklungen des Portals zu besprechen, wurde auf dem Berliner Datenportal eine Feedbackmöglichkeit realisiert.

Verwaltungsintern werden Informationsveranstaltungen und Softwarewerkzeuge angeboten, damit eigene Datensätze zukünftig ebenfalls auf dem Datenportal eingestellt werden können. Beiträge aus Wirtschaft, Wissenschaft und Zivilgesellschaft sind willkommen und können über eine spezielle Schnittstelle des Datenregisters bekannt gemacht werden.

Bis zum Ende des Jahres 2011 läuft der Pilotbetrieb, dann werden die Vorbereitungen für einen weiteren Ausbau und den Regelbetrieb getroffen.

Handlungsempfehlungen

Die strategischen Handlungsempfehlungen zu Offenen Daten in Berlin umfassen folgende Bereiche:

- Politisch das Bekenntnis zu offenen Daten in und aus Berlin und für das Berliner Datenportal. Es müssen ein Gesamtverantwortlicher für offene Daten und jeweils Ansprechpartner zu den Daten bestellt werden. Vermittelt über einen Ausbau des Datenportals, sind die Daten stabil und langfristig bereitzustellen. Die Föderation des Berliner Datenportals mit anderen Portalen ist vorzubereiten.
- Organisatorisch besteht die Notwendigkeit, Richtlinien zur Auswahl der Daten und wohldefinierte Prozesse für das Datenportal zu entwickeln. Es sind Weiterbildungsmaßnahmen für die Datenbearbeiter, Datenbereitsteller und Redakteure zu erarbeiten. Die Nutzergemeinde ist über vielfältige Interaktionsmöglichkeiten einzubinden, die auch Vorschlagsmöglichkeiten für Datensätze, Applikationen und Werkzeuge umfassen. Von der Senatsverwaltung publizierte Daten sollten automatisch in das Datenportal gestellt werden.
- Zu den rechtlichen Handlungsempfehlungen gehören einheitliche und einfache Lizenz- und Nutzungsbestimmungen und der direkte Zugang zu diesen Bestimmungen, um Barrieren bei der Nutzung der Daten zu minimieren. Wesentlich bei der Nutzung der Daten ist die Nachvollziehbarkeit der Daten.
- Technisch wird empfohlen, klare Standards für das Datenportal selbst und seine Sicherheit festzulegen. Es ist dabei insbesondere auf Kompatibilität, Integrierbarkeit und Qualitätssicherung der Daten zu achten.

Stufenplan

Auf Basis der Handlungsempfehlungen und Vorschläge zum weiteren Ausbau des Berliner Datenportals wird ein Stufenplan für die Berliner Open Data-Strategie entwickelt, der kurzfristige, mittelfristige und langfristige Maßnahmen umfasst: Kurzfristig ist »Open Data« in den Verwaltungsvorschriften des Landes Berlin zu verankern und das Berliner Datenportal in den Regelbetrieb zu überführen. Mittelfristig ist der nachhaltige Ausbau der Datenangebote des Landes Berlin und ihre Ergänzung um Werkzeuge, höherwertige Dienste, Schulungen etc. sicherzustellen. Langfristig ist an der überregionalen Abstimmung und der Integration der Berliner Datenangebote in Angebote in Deutschland, im deutschsprachigen Raum und in Europa zu arbeiten.

We are sending a strong signal to administrations today. Your data is worth more if you give it away. So start releasing it now: use this framework to join the other smart leaders who are already gaining from embracing open data. Taxpayers have already paid for this information, the least we can do is give it back to those who want to use it in new ways that help people and create jobs and growth.

Neelie Kroes, Vizepräsidentin der EU-Kommission,
12. Dezember 2011

1

Einleitung

1.1 Hintergrund und Motivation

Mit der zunehmenden Durchdringung verschiedenster Bereiche der Wirtschaft und Gesellschaft durch die Informationstechnik treten Informationen als Gut in das Zentrum der Betrachtung. Daten und Informationen sind wesentliche Ressourcen einer Informationsgesellschaft.

Fraunhofer FOKUS hat im Rahmen des Auftrags »Offene Daten für Bürger, Wirtschaft und Verwaltung« u. a. Konzepte, eine Architektur und ein Pilotsystem für den offenen und verlässlichen Zugriff auf aktuelle öffentliche Daten aus Berlin erarbeitet. Die Verwaltung, Unternehmen, Organisationen sowie Bürgerinnen und Bürger der Stadt erhalten so Informationen zur Erhöhung der Transparenz, zur gemeinsamen Gestaltung der städtischen Prozesse und Abläufe sowie zur Erhöhung der wirtschaftlichen Attraktivität der Stadt. Das entwickelte Konzept ermöglicht eine effiziente Bereitstellung, Analyse und Weiterverarbeitung der offenen Daten in Diensten und Applikationen durch Dritte. So werden Transparenz, Partizipation und Zusammenarbeit in der Stadt gestärkt.

Die Berliner Open Data-Initiative wird gestützt von zwei Initiativen zur Öffnung von Daten beliebiger Couleur – Open Data in den USA, Großbritannien und weiteren Staaten – und zur Verfügbarmachung von öffentlichen Daten – EU-Richtlinie über die Weiterverwendung von Informationen des öffentlichen Sektors (Public Sector Information – PSI) und weitere Vorschriften.

Fraunhofer FOKUS hat im Vorfeld dieser Arbeiten mit der Senatsverwaltung für Wirtschaft, Technologie und Frauen eine Vorstudie zum Stand der öffentlichen Daten in

Berlin erarbeitet. Dieser Studie zufolge stehen zwar relativ umfangreiche Berliner Daten inklusive Such- und Filterfunktionen zur Verfügung, jedoch handelt es sich bei diesen überwiegend um statische Informationen mit eingeschränkten Möglichkeiten zur Weiterverarbeitung. Sie werden mehrheitlich als Portable Document Format (PDF) angeboten und stellen oftmals bereits akkumulierte, interpretierte und bewertete Daten dar. Abgesehen von statistischen Daten und Geodaten fehlt es in Berlin an Basisdaten, die in wohldefinierten Formaten und an offenen Schnittstellen für die Weiterverarbeitung angeboten werden. Zudem werden einschränkende Lizenzbestimmungen verwendet, die die Entwicklung neuer Geschäftsmodelle auf der Grundlage öffentlicher Daten nachhaltig behindern. Ähnlich wie die seit mehr als einem Jahrzehnt verfügbaren Global Positioning System (GPS)-Daten, bieten auch andere Daten aus öffentlichen Quellen die Möglichkeit einer erfolgreichen wirtschaftlichen Nutzung; dies sollte im Interesse der Stadt gefördert werden.

1.2 Zielsetzung und Vorgehen

Berlin bereitet sich seit dem Frühjahr 2010 auf die Öffnung seiner Datenbestände vor. Technische, rechtliche und organisatorische Fragen wurden im Rahmen einer Vorstudie durch Fraunhofer FOKUS für die Landesregierung beleuchtet. Im Herbst 2010 brachte ein Onlinevoting zu den wichtigen und interessanten Datenkategorien mehr Klarheit über die Interessenlage der Bürgerinnen und Bürger.

Somit konnte im Frühjahr 2011 durch den Staatssekretärsausschuss für Verwaltungsmodernisierung im Rahmen des Programms »ServiceStadt Berlin« ein Projekt »Von der öffentlichen zur offenen Verwaltung« gestartet werden. Das Ziel des Auftrags »Offene Daten für Bürger, Wirtschaft und Verwaltung« ist es, über einen zentralen Zugang (Datenportal) einen strukturierten Zugriff (Datenkataloge) auf die weiterhin dezentral gepflegten Datenbestände im Land zu erhalten.

Wirtschaftsverwaltung, Innenverwaltung und das Amt für Statistik Berlin-Brandenburg bildeten das engere Projektteam. Die Bearbeitung erfolgte durch Fraunhofer FOKUS. Auf dem Berlin Open Data Day (BODDy) im Mai 2011 konnten bereits erste Ergebnisse wie die Berliner Open Data-Agenda vorgestellt werden. Dieses Forum bot auch die Gelegenheit, zivilgesellschaftliche Akteure mit einzubeziehen. Netzwerke, wie das Government 2.0 Netzwerk und das Open Data Network, sowie Wikimedia unterstützten die Aktivitäten und wollen die Datensätze für eigene Entwicklungen einsetzen. Auf der anderen Seite wurden Wünsche formuliert, die bei der Entwicklung und Erprobung berücksichtigt werden können. Durch die Einbeziehung der BerlinOnline Stadtportal GmbH als Betreiber des offiziellen Hauptstadtportals des Landes Berlin Berlin.de beschleunigte sich die gemeinsame Entwicklungsarbeit, da dort bereits Lösungen für ein internes Data Warehouse-Konzept vorlagen. Somit konnte Berlin zum 14. September 2011 als erste deutsche Stadt ein Datenportal freischalten.

1.3 Ergebnisse, Handlungsempfehlungen und Stufenplan

Fraunhofer FOKUS entwickelte in enger Abstimmung mit dem Auftraggeber und den Berliner Akteuren ein Konzept und Vorgehensmodell für eine Berliner Open Data-Plattform. Dafür wurden Stand und die Potenziale (siehe Kapitel 2) wie auch die organisatorischen (siehe Kapitel 3), rechtlichen (siehe Kapitel 6) und technischen Rahmenbedingungen (siehe Kapitel 4) für offene Daten in Berlin bewertet. Es wurden zudem Finanzierungsmöglichkeiten für offene Daten in der Stadt untersucht (siehe Kapitel 5).

Des Weiteren wurde gemeinsam mit der Senatsverwaltung für Wirtschaft, Technologie und Frauen, der Senatsverwaltung für Inneres, dem Amt für Statistik Berlin-Brandenburg und BerlinOnline ein Pilotsystem für ein Berliner Datenportal entwickelt (siehe Kapitel 7). Das Datenportal bildet eine zentrale Anlaufstelle für Daten des Landes Berlin. Die hier gesammelten Datensätze sind maschinenverarbeitbar und frei lizenziert. Sie ermöglichen so eine größere Transparenz bei Verwaltungsvorgängen, bieten u. a. Informationen zu Infrastruktur- und Umweltaspekten in der Stadt und laden zur Weiternutzung und Weiterverarbeitung durch Dritte ein.

Im Ergebnis der Studie und in Zusammenfassung der Erfahrungen mit dem Pilotsystem wurden kurz-, mittel- und langfristige Handlungsempfehlungen zum politischen, organisatorischen, rechtlichen und technischen Umgang mit offenen Daten in Berlin für die kommenden fünf Jahre formuliert. Ein Stufenplan gibt Empfehlungen für nächste Schritte zur Verstetigung und zum Ausbau des Datenangebots in Berlin.

2

Bestands- und Potenzialanalyse

Im Rahmen der Untersuchung zu offenen Daten in Berlin werden bestehende Aktivitäten sowohl auf lokaler Ebene als auch darüber hinaus in anderen Bundesländern, auf Bundesebene und beispielhaft auf internationaler Ebene betrachtet.

Eingangs werden dazu die Grundlagen des Open Data-Ansatzes und die gebräuchlichsten Begriffe erläutert. Dabei erfolgt auch ein Vergleich und eine Abgrenzung der Ziele der Ansätze von Open Data, Public Sector Information und Informationsfreiheit. Ferner werden die Potenziale von Open Data dargestellt.

2.1 Open Data: Grundlagen und Begriffe

Daten und Informationen spielen in nahezu allen gesellschaftlichen Bereichen eine wichtige Rolle. Wissensbasierte Gesellschaften sind angewiesen auf Daten, die als ein wesentlicher Rohstoff im Informationszeitalter gesehen werden müssen. Über Faktoren wie Verfügbarkeit, Korrektheit, Aktualität oder Preis dieser Daten haben Akteure wie Ersteller, Wiederverwerter oder Endnutzer von Daten häufig differierende Ansichten. Alle Gruppen haben jedoch ein hohes Interesse an Daten.

Werden Daten als Antriebsmittel unserer Gesellschaft verstanden, so kann die Bereitstellung oder das Fehlen von Daten bedeutende Auswirkungen auf Abläufe und Entscheidungen einer Gesellschaft und die Leistungsfähigkeit der Wirtschaft und des öffentlichen Sektors haben.

Öffentliche Verwaltungen erzeugen und verarbeiten täglich große Mengen von Daten und Informationen. Einen Teil davon stellen sie für Dritte häufig in Form von Dokumenten

wie Gesetzen, Registern, Karten, Tabellen, Zeitreihen oder Übersichten zur Verfügung. Viele der in diesen Dokumenten enthaltenen Informationen sind bereits aggregiert, die zugrunde liegenden Basisdaten nur organisationsintern verfügbar.

Dies hat verschiedene Gründe, so etwa dass es in vielen Einrichtungen des öffentlichen Sektors als eine wesentliche Aufgabe aufgefasst wird, Informationen in entsprechend verdichteter Form zu erstellen und zu kommunizieren. In historischer Perspektive war es bis vor Kurzem gar nicht oder nur mit hohem technischem und organisatorischem Aufwand möglich, einer Vielzahl von Personen die zugrunde liegenden Basisdaten von Verwaltungs- und Regierungsentscheidungen zur Verfügung zu stellen. Nicht zuletzt sind aktuelle Informationen und deren Besitz die Grundlage für Herrschaftswissen zur Steuerung der Gesellschaft.

Wie derzeit in verschiedenen Kontexten zu beobachten ist, bewirken technologische und ethisch-kulturelle Entwicklungen moderner Gesellschaften jedoch eine Entwicklung hin zu einem Verlangen nach mehr Mitbestimmung und Transparenz staatlicher Prozesse und Entscheidungen. Die Bereitstellung von Verwaltungsdaten zur Nutzung durch Gruppen außerhalb des öffentlichen Sektors kann hier ein wesentlicher Baustein sein. Informationsverarbeitung, semantische Technologien etc. bieten die Möglichkeit, diese Daten erstmalig in größerem Umfang und teilweise automatisiert weiterzuverarbeiten. Dies erlaubt es vergleichsweise kleinen zivilgesellschaftlichen Organisationen oder mittelständischen Unternehmen, leichter mit großen Datenmengen umzugehen und diese für ihre eigenen Ziele einzusetzen.

Auch die Repräsentanten öffentlicher Einrichtungen versprechen sich vom Open Data-Ansatz positive Effekte für ihre Organisationen. Zum einen dürfte eine leichtere Verfügbarkeit eigener Daten die Effizienz der eigenen Abläufe erhöhen. Zum anderen, so die Hoffnung, wird die Input- und Outputlegitimität des politisch-administrativen Systems verbessert: Transparenz und Mitbestimmung sollen besseren Input für politische Prozesse bewirken, was wiederum zu einem verbesserten Output in Form politischer Entscheidungen führt. Die Öffnung der Verwaltung hat also ein Potenzial für eine Transparenzdividende in Form von höherer Effizienz und gewachsener Legitimität und Glaubwürdigkeit.

Im folgenden Text wird zu Beginn Open Data dargelegt, das der vorliegenden Studie zugrunde liegt, um dann seine Relevanz im öffentlichen Sektor vorzustellen. Der aktuelle Stand in Deutschland und speziell Berlin wird umrissen. Einige internationale Beispiele für erfolgreiche Open Data-Portale werden benannt.

2.1.1 Definition und Potenziale von Open Data und Open Government Data

Der öffentliche Sektor ist ein stark informations- und wissensbasierter Sektor. Prozesse und Entscheidungen von Verwaltung und Politik werden fast immer auf Grundlage von

Informationen durchgeführt. Die Entwicklung einer Strategie zur Förderung der regionalen Wirtschaft erfordert die Zusammenführung und Auswertung ganz unterschiedlicher Datensätze. Für die Durchführung eines Planfeststellungsverfahrens in der Stadt- und Regionalplanung werden beispielsweise Informationen über die betreffende bauliche Veränderung, ihre Umwelt und die Belange der Anwohner benötigt. Öffentliche Verwaltungen handhaben täglich viele verschiedene Datenbestände. Ein Großteil dieser Daten steht verwaltungsintern zur Verfügung. Dritte außerhalb der jeweiligen Behörden haben keinen oder nur eingeschränkten Zugriff darauf. In Fällen wie der Entwicklung einer Wirtschaftsförderungsstrategie werden ausgewählte Dokumente zugänglich gemacht, etwa das finale Dokument der Strategie. Dem Dokument zugrunde liegende, feinkörnige Daten bleiben jedoch häufig allein in der Hand der Behörden bzw. der beauftragten Ersteller. In Vorgängen wie einem Planfeststellungsverfahren werden bereits vergleichsweise viele (Zwischen-)Informationen der Öffentlichkeit zugänglich gemacht.

Bei der Umsetzung eines Open Data-Ansatzes geht es darum: Welche Informationen werden in welcher Form wem unter welchen Bedingungen zur Verfügung gestellt? Im Folgenden wird dargestellt, was unter offenen Daten und speziell offenen Daten der Verwaltung zu verstehen ist und worin die Potenziale dieses Ansatzes gesehen werden.

2.1.2 Verständnis von Open Data und Open Government Data

Laut der Definition der Open Knowledge Foundation ist freies Wissen als ein Gegenstand oder Werk zu verstehen, mit dem Wissen transferiert wird und das verschiedene Kriterien erfüllt. Das Werk sollte u. a. im Ganzen zugänglich sein, einer diskriminierungsfreien Lizenz unterliegen, die eine Weiterverteilung und Wiederverwendung erlaubt, ohne dass der Nutzer dabei technologischen Restriktionen zu unterliegt. Die Verpflichtung, bei der Wiederverwendung die Urheber zu nennen und ein verändertes Werk als solches zu kennzeichnen, kann Bestandteil dieser Lizenz sein. Zudem muss die Lizenz des Werkes mit diesem weiterverteilt werden, dabei Gültigkeit behalten und gleichzeitig nicht die Weiterverteilung anderer Werke behindern (Open Knowledge Foundation, 2010). Im Open Data-Ansatz werden diese Punkte für Datenbestände aus allen Bereichen aufgegriffen und teilweise konkretisiert, wie z. B. offene Daten der Wissenschaft oder offene Daten der Wirtschaft. Open Government Data ist einer dieser Bereiche mit teilweise besonderen Anforderungen.

Bei der Bereitstellung von Open Government Data geht es um Daten der öffentlichen Verwaltung, die Dritten zur Weiterverwendung zur Verfügung gestellt werden. Ob die bereitgestellten Daten als offen bezeichnet werden können, ist abhängig von verschiedenen Faktoren wie der Zugänglichkeit, den Formaten und den rechtlichen Bedingungen, unter denen die Daten genutzt werden dürfen (von Lucke und Geiger, 2010, S. 6). Obwohl bislang keine einheitliche Definition des Begriffs existiert, besteht weitgehend Einigkeit über die Grundsätze, die Einfluss auf den Grad der Offenheit haben. Dabei werden die Prinzipien von Open Data auf Daten des öffentlichen Sektors angewendet.

In einer deutschsprachigen Übersetzung des Begriffs Open Government Data wird dieser Begriff als »offene Verwaltungsdaten« bezeichnet und wie folgt definiert: »Offene Verwaltungsdaten sind jene Datenbestände des öffentlichen Sektors, die von Staat und Verwaltung im Interesse der Allgemeinheit ohne jedwede Einschränkung zur freien Nutzung, zur Weiterverbreitung und zur freien Weiterverwendung frei zugänglich gemacht werden« (von Lucke und Geiger, 2010, S. 6).

Ob Daten, speziell Verwaltungsdaten, tatsächlich offen sind, lässt sich an der Erfüllung verschiedener Kriterien festmachen (Sebastopol Group, 2007; Sunlight Foundation, 2010):

- **Vollständigkeit:** Datensätze öffentlicher Einrichtungen sollten so vollständig wie möglich veröffentlicht werden, soweit dies die Regelungen zum Datenschutz zulassen. Die Daten sollen dabei möglichst im Rohformat inklusive beschreibender Metadaten veröffentlicht werden.
- **Primärquellen:** Daten, die als offene Verwaltungsdaten bereitgestellt werden, sollten direkt aus den ursprünglichen Quellen veröffentlicht und mit Informationen zum Ablauf der Sammlung und Erstellung der Daten angereichert werden. Dies erlaubt es Dritten, nachzuvollziehen, wie die Datenerstellung bewerkstelligt wurde.
- **Zeitliche Nähe:** Geschieht die Veröffentlichung der Daten möglichst zeitnah zur Entstehung oder Aktualisierung des Datensatzes? Werden Daten erst mit großem Abstand zu ihrer Kreierung zur Nutzung durch externe Personen bereitgestellt, so können sie unter Umständen nur noch eingeschränkt Verwendung finden. Welcher Zeitraum als zeitnah zu verstehen ist, ist dabei abhängig vom Einzelfall, statistische Daten sind hier beispielsweise anders einzuordnen als Daten zur Auslastung von Verkehrssystemen. Die zeitnahe Bereitstellung kann große Aufwände bei der Datenquelle erzeugen.
- **Leichter Zugang:** Welche Aufwände sind notwendig, um Zugang zu einem oder mehreren Datensätzen zu erhalten? Hindernisse hierbei reichen von technischen Problemen, wie unterschiedlichen Formaten oder fehlenden Metadaten, über organisatorische Barrieren (z. B. Pflichtregistrierung) bis hin zu rechtlichen Vorgaben.
- **Maschinenlesbarkeit:** Erst durch die Möglichkeit zur automatisierten Verarbeitung lassen sich viele Potenziale offener Daten ausschöpfen. Maschinenlesbarkeit eines Datensatzes erlaubt die Einbindung in Softwareanwendungen. Große Mengen an Daten lassen sich dementsprechend in neuen Applikationen nutzbar machen. Dies setzt u. a. eine solide Auszeichnung mit Metadaten sowie die Verwendung einheitlicher Formate und Terminologien voraus.
- **Diskriminierungsfreiheit:** ein Zugriff auf Daten ohne Ansehen der Person, zeitliche Restriktionen sowie die Anforderung, die eigene Identität nachzuweisen und eine Begründung für den Zugriff vorzulegen. Beispiele für eine Diskriminierung

beim Zugriff auf Daten wären eine notwendige Registrierung, Mitgliedschaft oder Einschränkung auf ausgewählte Softwareanwendungen.

- **Verwendung offener Standards:** Mit dem Einsatz offener Standards wird die Verwendung von Formaten angestrebt, die nicht nur von ausgewählten Programmen gelesen und verarbeitet werden können. Offene Standards gewährleisten die Freiheit, mit verschiedenen Programmen auf die Daten zugreifen zu können, ohne dass dafür Lizenzkosten an einzelne Hersteller abgeführt werden müssen.
- **Lizenzierung:** Eine restriktive Lizenzierung stellt ein Hindernis für die Nutzung von Daten dar. Offene Verwaltungsdaten sollten eindeutig sichtbar mit einer Lizenz versehen und gemeinfrei nutzbar sein.
- **Dauerhaftigkeit:** Daten, die im Internet veröffentlicht werden, können einfach verändert werden, ohne dass dies für Dritte transparent wird. Offene Verwaltungsdaten sollten permanent verfügbar sein und Änderungen, Aktualisierungen und Löschungen mit Versionskontrolle und Archivierung nachvollziehbar gestaltet werden.
- **Nutzungskosten:** Die Bereitstellung von Verwaltungsdaten gegen Entgelt behindert die Weiterverwendung dieser Daten. Ein Großteil der Daten der öffentlichen Hand wird unabhängig von späteren Nutzungsgebühren gesammelt und erstellt. Die Erhebung von Gebühren beschränkt generell die Gruppe der Nutzer und auch den Einsatz zu wirtschaftlichen Zwecken, was wiederum einen Effekt auf das Steuereinkommen haben kann.

2.1.3 Open Government Data, Public Sector Information und Informationsfreiheit

Neben dem noch relativ neuen Ansatz der offenen Verwaltungsdaten existieren – teilweise schon seit Jahrzehnten – Aktivitäten für mehr Transparenz und leichtere Weiterverwendung von Verwaltungsdaten. Diese Bestrebungen werden kurz vorgestellt, um die Ähnlichkeiten und Unterschiede der verschiedenen Begriffe und Konzepte zu umreißen (siehe hierzu auch Tabelle 2.1).

Bereits deutlich älter als der Open Data-Ansatz sind die Diskussionen um die Weiterverwendung von Informationen des öffentlichen Sektors sowie die Veröffentlichung administrativer Unterlagen im Zusammenhang mit Anfragen auf Grundlage von Informationsfreiheitsgesetzen. Zu beiden Bereichen existieren in vielen Ländern und auch auf EU-Ebene gesetzliche Regelungen.

Beim Ansatz der Informationsfreiheit wird davon ausgegangen, dass Bürger eines Landes einen Anspruch auf Informationen des öffentlichen Sektors haben. Mit der regulativen Umsetzung dieses Anspruchs in Form rechtlicher Normen zur Informationsfreiheit wird Bürgern die Möglichkeit eingeräumt, Anfragen an öffentliche Einrichtungen zu stellen, in denen sie um die Bereitstellung spezifischer Informationen bitten (Kubicek, 2008, S. 9).

Solange die gewünschten Informationen nicht unter verschiedene Ausnahmetatbestände fallen, ist die jeweilige Einrichtung gezwungen, die entsprechenden Unterlagen zur Verfügung zu stellen. Die Erfüllung dieses individuellen Interesses soll kostendeckend erfolgen.

Mit der Weiterverwendung von Informationen öffentlicher Einrichtungen werden bislang vor allem Ziele der wirtschaftlichen Entwicklung verfolgt. Diese Zielsetzung basiert auf der Annahme, ausgewählte Verwaltungsdaten besäßen einen Wert, der in der primären Verwendung, aber vor allem in der erneuten Verwendung durch Dritte zum Tragen kommt. Verschiedene Untersuchungen legen diese volkswirtschaftliche Bedeutung nahe (Dekkers, Polman, te Velde und de Vries, 2006; Fornefeld, Boele-Keimer, Recher und Fanning, 2009; Pollock, 2009).

	Open Government Data (Offene Verwaltungsdaten)	Public Sector Information	Informationsfreiheit
Ziele (in absteigender Bedeutung)	Transparenz, zivilgesellschaftliches Engagement, Wirtschaftsförderung, Effizienz	Wirtschaftsförderung, Transparenz	Transparenz
Bereitstellung	proaktiv	proaktiv	bislang häufig nur auf Nachfrage
Kosten für die Nutzer ¹	nach Möglichkeit ohne Kosten	Marginalkosten – Profitmaximierung	üblicherweise Marginalkosten
Umfang der Datensätze	große Mengen	große Mengen	eher kleine bis mittlere Mengen
Digital/analog	Digital	Digital	beides
Weiterverwendung	ja	ja	derzeit häufig nicht vorgesehen
Art der Daten	Überwiegend quantitativ	Überwiegend quantitativ	Überwiegend qualitativ

Tabelle 2.1: Abgrenzung der Begriffe Open Government Data (Offene Verwaltungsdaten), Public Sector Information, Informationsfreiheit anhand verschiedener Kategorien. Eigene Darstellung

2.1.4 Potenziale offener Daten

Open Data bieten gerade in und um den öffentlichen Sektor enorme Potenziale. Durch die Bereitstellung von Daten kann eine größere Transparenz für Außenstehende über

¹ Vgl. Newbery PSI Provision 2008, S. 11 f.

die Abläufe und Entscheidungen in Einrichtungen der öffentlichen Verwaltung erreicht werden. In Ergänzung finanzieller Subventionen für die lokale Wirtschaft kann der öffentliche Sektor Unterstützung in Form relevanter, aktueller Daten bieten. Diese können dazu beitragen, vorhandene Geschäftsmodelle zu verfeinern oder neue zu entwickeln. Eine strukturierte Veröffentlichung von Daten und Informationen kann auch die Arbeit von Einrichtungen des öffentlichen Sektors befördern. Die Möglichkeit, nach Arbeitsergebnissen anderer Organisationseinheiten zu suchen und darauf zuzugreifen, dies eventuell sogar organisations- und ebenenübergreifend, kann helfen, Doppelarbeiten zu vermeiden und die Ergebnisse der eigenen Arbeiten zu verbessern.

Im Folgenden werden die drei genannten Dimensionen – Transparenz, Wirtschaftsförderung und Produktivitätssteigerung – kurz im Einzelnen betrachtet.

2.1.4.1 Transparenz

Transparenz wird meist verstanden als die Möglichkeit für Dritte, außerhalb von Einrichtungen Einsicht in Prozesse und Entscheidungen zu nehmen, um sich so selbst eine Meinung dazu bilden zu können (Florini, 2007, S. 5; Piotrowski und Van Ryzin, 2007, S. 308). Diese recht breit angelegte Definition von Transparenz lässt sich auch auf das Verständnis von Transparenz im öffentlichen Sektor übertragen.

Mit der Bereitstellung von Daten und Informationen durch öffentliche Einrichtungen erhalten Außenstehende die Möglichkeit, ein genaueres Bild über bestimmte Vorgänge in Behörden zu bekommen. Werden diese Daten so zur Verfügung gestellt, dass sie einfach weiterverarbeitet werden können, vereinfacht dies die Aufbereitung der Daten für unterschiedliche Fragestellungen, die Organisationen oder interessierte Bürger haben.

Verschiedene Untersuchungen deuten auf einen gewissen positiven Zusammenhang zwischen der Durchsichtigkeit der Handlungen und Ergebnisse öffentlicher Organisationen und der allgemeinen Qualität politisch-administrativer Steuerung hin (Relly und Sabharwal, 2009, S. 149). Insofern kann davon ausgegangen werden, dass die proaktive Veröffentlichung staatlicher Daten auch einen positiven Einfluss auf die Handlungen der Akteure in Politik und Verwaltung haben dürfte.

Ebenso kann die größere Informiertheit zivilgesellschaftlicher Akteure dazu beitragen, dass diese sich stärker in politisch-administrative Vorgänge einbringen, ein aus demokratietheoretischer Sicht sicher erwünschter Effekt (Schmidt, 2008, S. 236). Allerdings ist durch die bloße Bereitstellung von Daten und Informationen nicht automatisch eine enorme allgemeine Steigerung der Bürgerpartizipation zu erwarten. Dies lässt sich eindeutig aus diversen Untersuchungen zu teilweise aufwendig organisierten partizipativen Angeboten in der öffentlichen Verwaltung ableiten (Holtkamp, 2006, S. 201; Kubicek, 2007, s. 4 f.). Unter Informationsbereitstellung als solcher ist auch noch keine Bürgerbeteiligung zu verstehen, offene Verwaltungsdaten können jedoch die Grundlage für eine aktive Teilhabe am politischen Geschehen bilden.

2.1.4.2 Wirtschaftsförderung

Eine übliche Methode zur Unterstützung von Wirtschaftsunternehmen ist die Bereitstellung finanzieller Mittel in Form von Subventionen. Neben anderen Bedenken bei diesen Werkzeugen bedeutet die direkte finanzielle Unterstützung in Zeiten knapper Mittel eine enorme Belastung der öffentlichen Haushalte. Als eine ergänzende Möglichkeit zur Beförderung des Wirtschaftswachstums kann die Bereitstellung von Verwaltungsdaten verstanden werden.

Mithilfe offener Daten der Verwaltung können Unternehmen bestehende Geschäftsmodelle verfeinern oder neue entwickeln. Für eine Firma im Bereich der Immobilienvermittlung etwa können präzise und aktuelle Daten zur sozialen Infrastruktur von Interesse sein, da diese Informationen eine weitere Verfeinerung der Einschätzung des Werts einer Immobilie erlauben.

Untersuchungen deuten darauf hin, dass die Bereitstellung offener Verwaltungsdaten zu Grenzkostenpreisen für eine Volkswirtschaft insgesamt vorteilhaft ist (Newbery, Bently und Pollock, 2008, S. 2). Im Falle einer digitalen Bereitstellung der Daten tendieren die Grenzkosten gegen null. Als Argumente für eine beinahe kostenlose oder sogar kostenfreie Bereitstellung von Verwaltungsdaten wird u. a. angeführt: Um die Durchschnittskosten der Datenerstellung tatsächlich zu decken, wäre ein deutlicher Preisaufschlag notwendig, der zu einer geringeren Verwendung der Daten und damit einer niedrigeren Wirkung auf die Gesamtwirtschaft führen würde (Newbery u. a., 2008, S. 123 f.). Zumindest sollten Basisdaten nicht preismaximierend, sondern höchstens kostendeckend angeboten werden (Office of Fair Trading, 2006, S. 134).

Am Beispiel des Bereichs der Geodaten lässt sich bereits heute ablesen, welchen Effekt eine Bereitstellung von Daten zu Grenzkostenpreisen haben kann. Das österreichische Bundesamt für Eich- und Vermessungswesen (BEV) hat im Jahr 2006 von einer Durchschnittskostenbepreisung auf die Bereitstellung zu Grenzkostenpreisen umgestellt. Dadurch sanken die Preise für ausgewählte Datentypen um bis zu 97 %. Gleichzeitig wuchs der Absatz dieser Daten um teilweise bis zu 7000 %. Im Ergebnis konnte das BEV so die Weiterverwendung seiner Daten bei gleichbleibendem Umsatz steigern. Wesentlichen Anteil an der gestiegenen Nachfrage hatten kleine und mittlere Unternehmen (Fornefeld u. A., 2009, S. 30).

2.1.4.3 Größere Produktivität in der Verwaltung

Wesentliche Teile der Aufgaben der öffentlichen Verwaltung lassen sich als wissensintensive Dienstleistungen beschreiben. Die Erarbeitung strategischer Dokumente zur wirtschaftlichen Entwicklung einer Metropolregion oder die Forschungsförderung eines Bundeslandes, aber auch die Konzentration verschiedener Aktivitäten auf sogenannte Sozialräume erfordert die Zusammenführung, Aufbereitung und Interpretation von Daten und Informationen aus unterschiedlichen Quellen. Wie generell bei großen Organisationen

gilt auch hier, dass die einzelnen Einheiten der Organisation meist nicht ausreichend informiert sind, welche Daten bereits von anderen Abteilungen erhoben wurden und somit der Organisation prinzipiell zur Verfügung stehen.

Bereits eine Übersicht über die in einer Organisation und ihren Teilen verfügbaren Daten und Informationen kann insofern helfen, die Arbeit der Mitarbeiter der Einrichtung zu erleichtern, und zu besseren Arbeitsergebnissen führen (Glock und Broens, 2008, S. 274). Ein Portal mit offenen Verwaltungsdaten kann somit ein Baustein eines organisationsweiten Informationsmanagements sein. Eine Auszeichnung der in einem Datenportal verzeichneten oder bereitgehaltenen Daten mit Metainformationen unterstützt das zielgerichtete Auffinden und Verwenden der eigentlichen Daten. Bei Rückfragen können die Datenbereitsteller einfach identifiziert und kontaktiert werden.

Mit den richtigen Mechanismen bietet sich auch die Möglichkeit, über ein Datenportal die Qualität der Informationen zu verbessern. Eingebaute Feedbackmöglichkeiten erlauben den Datennutzern, Rückmeldungen zu Problemen mit einzelnen Datensätzen zu geben oder sogar selbst die Fehler in den bereitgestellten Informationen zu beheben oder die Daten zu ergänzen.

2.2 Situation offener Verwaltungsdaten in Deutschland und im Ausland

Von der Entwicklung des Konzepts und der damit verbundenen Prinzipien offener Verwaltungsdaten bis zur praktischen Umsetzung in öffentlichen Einrichtungen und Gebietskörperschaften unterschiedlicher Länder ist es ein weiter Weg. Im Ausland wurde der Ansatz teilweise wesentlich zügiger aufgegriffen, und entsprechend ist die Realisierung weiter fortgeschritten. Neben dem zeitlichen Faktor lassen sich sicher auch kulturelle und verwaltungsorganisatorische Gründe für ein schnelleres Aufgreifen des Ansatzes benennen. Transparenz hat in einigen Nationen eine lange Tradition, auf entsprechend weniger Barrieren stößt hier die Umsetzung von Transparenz im öffentlichen Sektor mithilfe von Technologien des 21. Jahrhunderts. Daneben spielt sicher auch die Ausgestaltung des staatlichen Aufbaus eine Rolle. In stärker zentralistisch orientierten Ländern, wie etwa Großbritannien, können neue, transformative Konzepte im öffentlichen Sektor schneller implementiert werden, als in eher föderalistisch geprägten Nationen.

Im Folgenden wird übersichtsartig die gegenwärtige Situation in Deutschland und speziell in Berlin im Hinblick auf die Bereitstellung von Daten durch die öffentliche Verwaltung dargestellt; außerdem werden beispielhaft entsprechende internationale Ansätze skizziert. Dabei erfolgt eine Konzentration auf Aktivitäten, die bewusst dem Ansatz offener Daten folgen bzw. diesem aufgrund ihres Ziels, den Zugang zu Informationen oder deren Weiterarbeit zu vereinfachen, zugeordnet werden können.

2.2.1 Situation in Deutschland

In Deutschland setzen sich neben zivilgesellschaftlichen Organisationen wie dem Open Data Network² oder dem Government 2.0 Netzwerk³ erste Personen und Organisationseinheiten in Politik und Verwaltung mit der Bedeutung offener Daten und ihrer Bereitstellung für die öffentliche Verwaltung auseinander. Open Data haben als Pilotvorhaben Eingang in die Strategie des Bundes zur Verwaltungsmodernisierung gefunden. In den Landtagswahlen 2010 in Nordrhein-Westfalen und 2011 in Baden-Württemberg wurden Open Data-Positionen geäußert und fanden Eingang in die entsprechenden Regierungsprogramme. Ferner entwickeln neben Berlin der Stadtstaat Bremen sowie die bayerische Landeshauptstadt München erste Aktivitäten.

2.2.1.1 Aktivitäten auf Bundesebene

Verschiedene Bundesbehörden und ihre nachgeordneten Einrichtungen stellen seit Jahren proaktiv Verwaltungsdaten online bereit. Speziell in den Bereichen statistischer und umweltspezifischer Daten sowie im Bereich der Informationen mit Raumbezug haben entsprechende verwaltungspolitische Aufgabenstellungen und die nationale Umsetzung von EU-Vorgaben zu ausgeprägten Initiativen geführt. Zum Beispiel sind beim Bundesamt für Statistik⁴ bereits heute vielfältige Informationen zu ökonomischen, gesellschaftlichen und weiteren Entwicklungen zum Download erhältlich. Diese Informationen sind in ihrem Umfang und in ihrer Struktur häufig bereits kompatibel zu vergleichbaren statistischen Daten der europäischen Statistik sowie der deutschen Landes- und Kommunalstatistikämter. Im Umweltbereich wird spätestens seit der Neufassung des Umweltinformationsgesetzes (UIG) von 2004 der freie Zugang zu Informationen über die Umwelt auf Bundesebene vorgeschrieben (Umweltinformationsgesetz (UIG), 2004), entsprechende Regelungen finden sich auch auf Länderebene. Als technologisches Werkzeug zur Umsetzung des UIG bietet das PortalU⁵ des Bundesumweltamtes Zugang zu umfangreichen Datenbanken mit Umweltinformationen. Das Portal ist Ergebnis einer Bund-Länder-Kooperation und wurde mit dem Ziel eines einheitlichen Einstiegspunktes für Umweltdaten entwickelt. Nutzer können über die Suchmaschine in den angeschlossenen Datenbeständen mit Unterstützung eines Fachwörterbuches, dem Semantischen Networkservice recherchieren (Schenk, Kruse, und Klenke, 2010) (zu den rechtlichen Details des Umweltinformationsrechts siehe Abschnitt 6.1.3).

Im Bereich räumlicher Informationen, also geobasierter Daten, existiert als nationale Umsetzung der EU-Richtlinie zur Schaffung einer Geodateninfrastruktur in der Europäischen Gemeinschaft mit dem Gesetz über den Zugang zu digitalen Geodaten (GeoZG, 2009) eine rechtliche Vorgabe zum Aufbau einer digitalen Infrastruktur für Deutschland

²<http://opendata-network.org>.

³<http://www.gov20.de>.

⁴<http://www.destatis.de>.

⁵<http://www.portalu.de>.

(GDI-DE), die den Zugang zu Geodaten für unterschiedliche Zielgruppen ermöglichen soll. Entsprechende Normen finden sich ebenfalls auf Ebene der Länder, in denen Regelungen für Geodaten im Zuständigkeitsbereich von Ländern und Kommunen getroffen werden (zu den rechtlichen Details der Geodatengesetzgebung siehe Abschnitt 6.1.5).

Mit dem organisatorisch-technischen Aufbau der bundesweiten Geodateninfrastruktur GDI-DE wurde bereits im Jahr 2003 begonnen, vor dem Erlass des GeoZG. Ein technisches Architekturkonzept liegt seit 2007 vor und wurde im Jahr 2010 in einer überarbeiteten Version veröffentlicht. Begleitet wurde die technische Umsetzung mit Informationsmaßnahmen für die involvierten datenhaltenden Stellen der öffentlichen Verwaltung sowie der Abfrage existierender Geodatenansätze nach Anhang 1 der EU-Richtlinie INSPIRE bei den entsprechenden Stellen bei Bund, Ländern und Kommunen. Im Zusammenhang mit dem Aufbau der GDI-DE werden vorrangig die Geodaten aus den Bereichen bereitgestellt, die in den Anhängen 1 bis 3 der INSPIRE-Richtlinie beschrieben sind (INSPIRE, 2007). Einer der nächsten geplanten Schritte ist die Umsetzung eines Geoportals für Deutschland, über das ein zentraler Zugang sowohl zu den Geodaten als auch zu den Geodatendiensten realisiert wird.

Bei den Daten in den beschriebenen Bereichen mit geregelter Informationszugang handelt es sich nicht notwendigerweise um offene Daten entsprechend der Open Data-Prinzipien. So sieht das Geodatenrecht bislang eine eingeschränkte, häufig kostenpflichtige Weiterverwendung raumbezogener Daten vor. Ein über die Umweltdaten oder Geodaten hinausgehender einheitlicher Zugriff auf Verwaltungsdaten über einen gemeinsamen Einstiegspunkt, eine durchgängig abgestimmte Auszeichnung mit Metadaten sowie die Bereitstellung als Linked Data steht dagegen vielfach noch aus. Die Entwicklung einer organisationsübergreifenden Strategie für offene Verwaltungsdaten ist in diesem Zusammenhang mittel- bis langfristig sicher unerlässlich, wenn die Zielsetzung von mehr Transparenz und Wirtschaftsförderung mittels einfach zugänglicher Verwaltungsinformationen angenommen wird.

Zu unterscheiden von Initiativen zur proaktiven Bereitstellung von Verwaltungsdaten zur Weiterverwendung durch Dritte ist bislang meistens der Ansatz des Informationsfreiheitsrechts. Die meisten einschlägigen Gesetze in Deutschland sehen einen Zugang zu Verwaltungsinformationen auf Anfrage einer Person vor. Eine Ausnahme bildet die gesetzliche Regelung auf Landesebene im Fall des Stadtstaates Bremen; dort ist eine proaktive Bereitstellung von Informationen vorgesehen (siehe Abschnitt 2.2.1.2). Das Gesetz zur Regelung des Zugangs zu Informationen des Bundes schreibt seit dem Jahr 2006 einen Anspruch auf Informationszugang bei Verwaltungsinformationen im Bereich der Bundeseinrichtungen vor.

Als zivilgesellschaftliche Initiative setzt das Projekt »FragdenStaat.de«⁶ an dem Anspruch auf Informationszugang auf Bundesebene an. Mithilfe eines Internetportals gleichen Namens sollen entsprechende Anfragen zum einen vereinfacht und zum anderen

⁶<https://fragdenstaat.de>.

öffentlich gemacht werden. Auf diese Weise erhofft man sich eine Steigerung der Bekanntheit des Rechts auf Informationszugang sowie eine Zunahme positiver Antworten seitens der angesprochenen Verwaltungseinheiten bei gleichzeitiger Abnahme von Doppelanfragen aufgrund der öffentlichen Dokumentation der Anfragen.

Im Herbst 2010 hat die Bundesregierung ihr aktuelles Programm zur Verwaltungsmodernisierung »Vernetzte und transparente Verwaltung« veröffentlicht. Als wesentlichen Ziele, die bei der Reform der öffentlichen Verwaltung berücksichtigt werden sollen, werden hier die stärkere Kooperation mit Akteuren aus Zivilgesellschaft und Wirtschaft sowie eine generell höhere Transparenz der eigenen Handlungen und Entscheidungen festgestellt.

Im Regierungsprogramm werden 19 verschiedene Modernisierungsprojekte mit kurzen Projektsteckbriefen aufgeführt. Im Projekt »Open Government« soll »wo es rechtlich möglich und zweckmäßig ist – ein freier Zugang zu Informationen der Verwaltung geschaffen werden« (Bundesministerium des Innern, 2010, S. 57 f.). Ziel ist es, Pilotvorhaben für die Bereitstellung von Verwaltungsdaten zu realisieren und bis zum Jahr 2013 eine mit den Ländern abgestimmte Open Government-Strategie entwickelt zu haben, um dann an die Umsetzung zu gehen.

Ebenfalls von Bedeutung ist die Realisierung eines nationalen Wettbewerbs für Softwareanwendungen und Ideen auf Basis offener Verwaltungsdaten. Dieser »Apps für Deutschland« genannte Wettbewerb läuft von Anfang November 2011 bis Anfang März 2012 unter der Schirmherrschaft des Bundesministers des Innern, Herrn Dr. Friedrich. Koordiniert wird der Wettbewerb von den zivilgesellschaftlichen Organisationen Government 2.0 Netzwerk Deutschland e. V., Open Data Network e. V. und Open Knowledge Foundation Deutschland. Als Unterstützer sind der Bundesverband Informationswirtschaft, Telekommunikation und neue Medien (Bitkom e. V.) sowie verschiedene Gebietskörperschaften und Fachbehörden tätig. Das Land Berlin sowie die Freie Hansestadt Bremen fördern den nationalen Wettbewerb mit der Bereitstellung eigener offener Verwaltungsdaten. Ziele des Wettbewerbs sind es, den Ansatz offener Verwaltungsdaten bekannter zu machen, innovative Anwendungen und Ideen für die Weiterverwendung von Verwaltungsinformationen zu erhalten und weitere Daten zur Verfügung zu stellen. Verwaltungseinheiten aller föderalen Ebenen sind im Rahmen des Wettbewerbs aufgerufen eigene Datensätze als offene Daten bereitzustellen. Preise werden in den drei Kategorien Ideen, Anwendungen und Daten vergeben. Die Preise werden durch Sponsoren aus der Privatwirtschaft ermöglicht (Government 2.0 Netzwerk Deutschland e. V. et al., 2011).

Ebenfalls von deutschlandweiter Bedeutung ist eine mögliche Teilnahme Deutschlands an der internationalen Initiative »Open Government Partnership«⁷. Dieses Vorhaben wurde im Herbst 2011 von verschiedenen Staaten wie den USA, Brasilien und Norwegen und Nicht-Regierungsorganisationen (NGOs) angestoßen. Ziel ist die Förderung der Umsetzung und Fortentwicklung des Open Government-Ansatzes in den beteiligten Ländern. Diese setzen sich eigene Verpflichtungen, die öffentlich bekannt gegeben werden.

⁷<http://www.opengovpartnership.org>.

Mithilfe eines gemeinsam mit den zivilgesellschaftlichen Organisationen durchzuführenden Monitorings sollen entsprechende Fortschritte dokumentiert werden. In Deutschland formieren sich verschiedene NGOs vor allem aus den Bereichen Bürgerbeteiligung und Transparenz, um die Bundesregierung zu einer Teilnahme an der Initiative zu bewegen (Drescher, 2011).

2.2.1.2 Aktivitäten auf regionaler und lokaler Ebene

Entsprechend den nationalen Gesetzgebungen in den Bereichen Umwelt-, Verwaltungs- und Informationsrecht finden sich vergleichbare Normen auf der regionalen Ebene der Länder in Deutschland. Hier weisen die Gesetzgebungen entsprechend der föderalen Konkurrenz unterschiedliche Ausprägungen auf oder fehlen teilweise auch gänzlich. Speziell auf das Thema Bereitstellung von Verwaltungsdaten als offene Daten konzentrieren sich derzeit neben Berlin wenige Länder und Kommunen. Nachfolgend werden einzelne Initiativen skizziert, die unterschiedlich weit fortgeschritten sind. Im darauf folgenden Abschnitt werden die spezifischen Berliner Aktivitäten dargestellt.

Auf regionaler Ebene in Deutschland findet sich z. B. in Nordrhein-Westfalen (NRW) die Ankündigung, sich dem Thema Open Government und damit auch offenen Verwaltungsdaten widmen zu wollen. In der Koalitionsvereinbarung der aktuellen Regierungsparteien in NRW wird eine stärkere Öffnung von Verwaltung angekündigt. Mit der genannten Initiative soll eine »kosten- und diskriminierungsfreie Bereitstellung öffentlicher Daten und Informationen« erfolgen. Weiterhin sollen regelmäßig Wettbewerbe durchgeführt werden, bei denen »Entwickler unter Verwendung öffentlich bereitgestellter Daten Applikationen erstellen, die zu Transparenz, Partizipation und Zusammenarbeit beitragen« (NRW SPD, 2010, S. 81).

Das zweite Flächenbundesland, in dem mehr Transparenz über die Bereitstellung von Verwaltungsinformationen nach dem Open Data-Ansatz angekündigt wurde, ist Baden-Württemberg. Die neue grün-rote Landesregierung ist bestrebt, zuerst ein Landesinformationsfreiheitsgesetz auf den Weg zu bringen, analog zu bereits existierenden Normen in den meisten anderen Bundesländern. Hierbei soll jedoch explizit eine Ausrichtung an den Open Data-Prinzipien erfolgen (Bündnis 90/Die Grünen und SPD Baden-Württemberg, 2011, S. 79).

Im Gesetz des Bundeslandes Bremen zur Informationsfreiheit sind Regelungen zur proaktiven Bereitstellung einzelner Informationen des öffentlichen Sektors und zur Erstellung eines entsprechenden Informationsregisters festgelegt. Hierin unterscheidet sich das bremische Informationsfreiheitsgesetz von vergleichbaren Normen in anderen Bundesländern und auf Bundesebene. Um eine strategische Herangehensweise an offene Verwaltungsdaten zu entwickeln, wurde von lokalen Akteuren aus Verwaltung, Politik und Wissenschaft eine »Bremer Empfehlung zu Open Government Data« verfasst und veröffentlicht (Linnert, Kubicek und Sommer, 2011).

In dieser Empfehlung werden zum einen ausgewählte Open Data-Prinzipien für die lokalen Gegebenheiten interpretiert und zum anderen eigene Anliegen formuliert. In dem Dokument wird die Notwendigkeit einer an den Bedürfnissen verschiedener Nutzergruppen orientierten Erschließung als wichtig angesehen. Technische Standardisierung wird als wesentliche Voraussetzung für die großflächige Umsetzung offener Verwaltungsdaten verstanden; diese soll mit Arbeiten kombiniert werden, die ohnehin in der Verwaltung notwendig werden, um etwa den Einsatz von Dokumentenmanagementsystemen sinnvoll zu gestalten. Ebenso behandelt werden die Themen Kosten und Maschinenlesbarkeit. Für Ersteres wird eine differenzierte Betrachtung empfohlen, abhängig von den späteren Nutzergruppen. Maschinenlesbarkeit wird als grundsätzlich richtig erachtet, jedoch sollte die Umwandlung in maschinenlesbare Formate möglichst keine Mehrkosten aufseiten der Verwaltung erfordern.

Die Bremer Empfehlung enthält mehrere neue Ergänzungen zu den existierenden Open Data-Prinzipien. Es wird die Schaffung organisatorischer Verfahren und Strukturen in der Verwaltung gefordert, um offene Verwaltungsdaten nachhaltig zu ermöglichen. Zur Realisierung einer föderalen Plattform für offene Verwaltungsdaten wird ein verteiltes Redaktionsmodell empfohlen, bei dem von Beginn an auf Qualität der Inhalte und Beschreibungen geachtet werden soll. Abschließend wird der Aufbau einer kooperativen Dateninfrastruktur von den Urhebern der Empfehlung als notwendig erachtet.

Auf kommunaler Ebene sind in Deutschland neben den Stadtstaaten Berlin und Bremen bislang vor allem Aktivitäten in München feststellbar. München hat im Rahmen des lokalen Open Government-Pilotprojektes »Munich Open Government Day« (MOGDy) je einen Anwendungs- und Ideenwettbewerb durchgeführt.⁸ Dabei ging es darum, Anwendungen oder Vorschläge mit dem Ziel zu entwickeln, Daten des öffentlichen Sektors kreativ zur Lösung städtischer Probleme, zur Verbesserung existierender Geschäftsmodelle oder für mehr Transparenz in Entscheidungen und Prozessen der öffentlichen Verwaltung einzusetzen. Das Pilotprojekt MODGy hatte eine Laufzeit von Dezember 2010 bis Juni 2011. Im Herbst 2011 wurde das Projekt im Rahmen des European Public Sector Awards des European Institute for Public Administration mit einem Best-Practice-Zertifikat ausgezeichnet (European Institute for Public Administration, 2011).

2.2.2 Bereitstellung von Verwaltungsdaten im Land Berlin

In diesem Abschnitt werden verschiedene Berliner Initiativen mit Bezug zur Bereitstellung von Daten und Informationen aus dem öffentlichen Sektor auf Anknüpfungspunkte für den Open Data-Ansatz hin untersucht. Neben der allgemeinen Initiative zur Modernisierung der Berliner Verwaltung »ServiceStadt Berlin« werden die Anstrengungen zum Aufbau einer Geodateninfrastruktur auf Grundlage der EU-Verordnung INSPIRE sowie die konkreten verwaltungsinternen Projekte »Datenpool« und »Prisma« betrachtet. Weiter werden die existierenden Aktivitäten zur Umsetzung des Open Data-Ansatzes in

⁸ MOGDy-Programmierwettbewerb:
<http://www.muenchen.de/Rathaus/dir/limux/mogdy/Programmierwettbewerb>.

Berlin beschrieben. Abschließend erfolgt eine Bewertung der beschriebenen Initiativen im Hinblick auf ihre Bedeutung für die Fortentwicklung offener Daten im Land Berlin.

Auf die relevanten rechtlichen Rahmenbedingungen im Land Berlin wird in Kapitel 6 detailliert eingegangen. Technische Aspekte der Datenbereitstellung inklusive einer entsprechenden IKT-Architektur werden in Kapitel 4 beschrieben.

Mit der Initiative »ServiceStadt Berlin 2016« wurde durch das Deutsche Institut für Urbanistik im Jahr 2009 im Auftrag des Landes Berlin eine Analyse der Servicelandschaft Berlins durchgeführt und diese mit verschiedenen Trends abgeglichen, um so Anforderungen an die zukünftige Ausgestaltung der Dienstleistungen im Land Berlin zu identifizieren und Empfehlungen zur strategischen Fortentwicklung der Berliner Verwaltung und ihrer Dienstleistungen zu formulieren. In der Studie werden aus unterschiedlichen Perspektiven auch Aspekte der Informationsbereitstellung für die verwaltungsinterne und die -externe Verwendung angesprochen.

Als ein relevantes Leitbild zukünftiger Entwicklungen, das auch bei der Modernisierung der öffentlichen Verwaltung zu berücksichtigen ist, wird das Bild der Bürgergesellschaft beschrieben. In dieser Vision versuchen Politik und Verwaltung, Bürgerinnen und Bürger dafür zu gewinnen, sich an der Erstellung öffentlicher Leistungen zu beteiligen bzw. diese zu unterstützen. Dem Leitbild zufolge müssen staatliche Einrichtungen die Bürger aktivieren. Ein Mittel hierzu liegt in der Verringerung existierender Hürden zu bürgerschaftlichem Engagement (Grabow, Knipp und Schneider, 2009, S. 27 f). Im Zusammenhang mit einer stärkeren Teilhabe der Bürger an politischen und administrativen Prozessen wird in der Untersuchung die Notwendigkeit verstärkter Informationsangebote unterstrichen. Im Hinblick auf die Erreichung von mehr Transparenz und Beteiligung wird die Bereitstellung von Informationen als ein wichtiger Baustein gerade für informelle Formen der Bürgerbeteiligung erachtet (Grabow et al., 2009, S. 117f).

An der Bereitstellung verbesserter Informationsangebote machen die Autoren der Untersuchung verschiedene Vorteile fest. Für Verwaltungsmitarbeiter kann teilweise die Beantwortung von Fragen erleichtert werden, da auf online zugängliche Informationen zugegriffen oder verwiesen werden kann. Weiter kann es für Bürgerinnen und Bürger einfacher werden, Verwaltungshandeln nachzuvollziehen, was der Studie zufolge zu einer erhöhten Handlungsakzeptanz führt. Darüber hinaus bilden Informationen eine wichtige Grundlage, um sich in politische Prozesse einzubringen. Verbesserte Informationsangebote helfen also, Beteiligungsmöglichkeiten zu stärken (Grabow et al., 2009, S. 130 f).

In der Studie werden existierende Dienstleistungen zu allgemeinen Servicetypen zusammengefasst. Ein Typ öffentlicher Dienstleistungen wird als »amtliche Dokumente und Register« bezeichnet. Hierbei geht es vor allem um die organisationsinterne Bereitstellung von Informationen zur besseren Aufgabenerledigung durch die verschiedenen Verwaltungseinheiten. Als Ergebnis für diesen Bereich wird in der Untersuchung empfohlen, die Organisation der Informationen und die Zugriffssteuerung weiterhin in den Händen öffentlicher Einrichtungen zu belassen, um sicherzustellen, dass Belange des Datenschutzes erfüllt werden. Um den Umgang mit Daten und Registern zu verbessern, werden

eine komfortablere Bereitstellung der Informationen sowie eine stärkere Orientierung an Qualitätsstandards sowie einfache »Zugriffsmöglichkeiten für andere Behörden und gegebenenfalls andere Adressaten« (Grabow et al., 2009, S. 146) vorgeschlagen.

Grabow, Knipp und Schneider (2009, S. 77) stellen fest, dass »multifunktionale und multimediale Geräte den Bedarf nach öffentlichen Informationen und Dienstleistungen an jedem Ort und zu jeder Zeit (z. B. unmittelbare Ortsinformationen auf das Handy)« auf der Seite der Adressaten von Verwaltungsinformationen verstärken. Um für diese Adressaten von Verwaltungsleistungen den Zugang zu vereinfachen, sind einheitliche Zugänge von wesentlicher Bedeutung; hier können sich »deutliche Verbesserungen hinsichtlich der Erreichbarkeit« ergeben (Grabow, Knipp und Schneider 2009, S. 79f).

In der Berliner Verwaltung werden unterschiedliche Anstrengungen zur Verbesserung der Bereitstellung, des Zugangs und der Weiterverwendung von Verwaltungsdaten vor dem Hintergrund verschiedener fachlicher Anforderungen unternommen.

Mit dem Aufbau der Geodateninfrastruktur Berlin-Brandenburg (GDI-BE/BB) verfolgen die Verwaltungen beider Bundesländer die Realisierung einer digitalen Infrastruktur, mit der Geodaten bereitgestellt werden können. Diese Arbeiten erfolgen entsprechend den Anforderungen der EU-Richtlinie INSPIRE, des nationalen Geodatenzugangsgesetzes und der Landesgeodatenzugangsgesetze. Als Basis dieser Entwicklung erarbeiteten Vertreter beider Bundesländer bis Ende 2008 den »Masterplan für den Aufbau der Geodateninfrastruktur-Berlin/Brandenburg«. Eine zentrale Maßnahme hier ist die Umsetzung eines länderübergreifenden Geoportals als Werkzeug zur Erschließung und Integration der dezentral vorliegenden Geodaten. Zur Vereinheitlichung der Daten mit Raumbezug wird ein Geodatenmanagement entwickelt, sodass die Daten interoperabel zwischen den verschiedenen Verwaltungseinrichtungen in Berlin und Brandenburg und allen weiteren Teilnehmern der GDI-DE sowie entsprechender Einrichtungen der Mitgliedstaaten der EU genutzt werden können. Dafür werden verschiedene Werkzeuge entwickelt und bereitgestellt, mit denen Geodaten und ihre Metadaten betrachtet und bearbeitet werden können. Organisatorisch wird als zentrale Kontaktstelle rund um die Geodateninfrastruktur ein gemeinsames Geoservicecenter errichtet. Diese Einrichtung hat neben Aufbau und Betrieb der technischen Infrastruktur Controlling- und Kommunikationsfunktionen. Daneben sind die wichtigsten Anbieter von Verwaltungsdaten in einem länderübergreifenden Lenkungsausschuss vertreten (Land Brandenburg und Land Berlin, 2008).

Im Rahmen der verstärkten Sozialraumorientierung der Berliner Sozial- und Stadtentwicklungsplanung erfolgt seit einigen Jahren eine Fokussierung auf sogenannte Lebensweltlich orientierte Räume (LoR). In diesem Konzept, das im Jahr 2006 durch den Berliner Senat beschlossen wurde, wird eine teilträumliche Gliederung als bezirksübergreifende Vergleichsebene festgelegt, die zwischen den relevanten Akteuren im Land abgestimmt ist. Um den involvierten Verwaltungseinheiten ein effektiveres Eingreifen in die Entwicklung der Sozialräume zu ermöglichen, wurde der Aufbau eines Monitoringsystems beschlossen. Einer der Bausteine dieses Monitoringsystems ist der gemeinsame Datenpool, der unter Federführung des Amts für Statistik Berlin-Brandenburg und

in Zusammenarbeit mit der Senatsverwaltung für Stadtentwicklung sowie verschiedenen Bezirksverwaltungen entwickelt wird. Dazu werden u. a. entsprechende Datensätze mit einheitlichen Merkmalen und Schlüsselindikatoren identifiziert, abgestimmt und auf Ebene der LoR zur Weiterverwendung durch die Senats- und Bezirksverwaltungen bereitgestellt (Senatsverwaltung für Stadtentwicklung Berlin, 2009, S. 35 f.). In der finalen Ausbaustufe soll der Datenpool ungefähr 60 Datenbestände mit circa 500 Variablen enthalten. Ein weiterer Ausbau speziell mit Daten der Bezirksverwaltungen ist vorgesehen.

Aufbauend u. a. auf den digitalen Informationen des Datenpools, wird im Projekt »Planungsraumbezogenes Informationssystem für Monitoring und Analyse« (PRISMA) im Rahmen des Senatsprogramms »ServiceStadt Berlin« ein System zur Verwaltung und Auswertung von Daten mit Raumbezug für die Berliner Verwaltung geschaffen. Mit diesem System wird die Entwicklung und Steuerung sozialraumorientierter Planungsmaßnahmen und Handlungen unterstützt. Gleichzeitig wird das Datenmanagement stärker vereinheitlicht und ein System zur Handhabung von Metadaten geschaffen. Verwaltungsmitarbeiter sollen zukünftig in die Lage versetzt werden, selbstständig Auswertungen der zugrunde liegenden Datenbestände vorzunehmen. Einerseits werden Fachanwender ohne Spezialwissen in Informationstechnologie und zum anderen beispielsweise GIS-Experten dabei unterstützt, einfache bis komplexe Analysen vorzunehmen. Diese Untersuchungen können für spontan auftretende Fragestellungen genauso eingesetzt werden, wie für regelmäßig wiederkehrende Aufgaben bspw. im Berichtswesen. In dem Pilotvorhaben sind die Senatsverwaltung für Stadtentwicklung und der Bezirk Marzahn-Hellersdorf federführend aktiv. Weitere Bezirke werden bereits in das Pilotvorhaben aktiv einbezogen. Der Prototyp des PRISMA-Systems soll bis Ende des Jahres 2011 als skalierbares System aufgebaut sein, das im Anschluss auf alle Berliner Bezirke ausgeweitet werden kann (Bangert, 2010, S. 107 ff.; Senatsverwaltung für Inneres und Sport, 2010, S. 135 ff.).

Der Berliner Senat ist darüber hinaus direkt mit dem Aufbau von Strukturen für die Online-Bereitstellung seiner Daten entsprechend dem Open Data-Ansatz beschäftigt. Gemeinsam mit den Städten Amsterdam, Barcelona, Helsinki und Paris ist Berlin am EU-Projekt »Open Innovation Mechanisms in Smart Cities« (Open Cities) beteiligt. Eines der wesentlichen Ziele dieses Projektes ist die Entwicklung und Umsetzung einer prototypischen Open Data-Plattform. Im Anschluss an die Implementierung der Plattform in den beteiligten Städten werden Softwareentwicklungswettbewerbe durchgeführt, um die Bekanntheit und Nutzung der offenen Verwaltungsdaten zu erhöhen (Almirall, 2010).

In Berlin werden die Öffnung von Verwaltungsdaten und deren strukturierte Bereitstellung seit einiger Zeit von verschiedenen Akteuren vorangetrieben. Die Senatsverwaltungen für Wirtschaft, Technologie und Frauen sowie für Inneres und Sport und das Amt für Statistik Berlin-Brandenburg sind wesentliche Akteure der öffentlichen Verwaltung, die das Thema offene Verwaltungsdaten verfolgen. Das IT-Dienstleistungszentrum

Berlin und die BerlinOnline Stadtportal GmbH befassen sich ebenfalls mit den Potenzialen offener Daten. Vorangetrieben haben die stärkere Öffnung von Verwaltungsdaten von Anfang an verschiedene zivilgesellschaftliche Akteure in Berlin wie das Government 2.0 Netzwerk Deutschland, das Open Data Network, die deutsche Abteilung der Open Knowledge Foundation und Wikimedia Deutschland. Vertreter lokaler Forschungseinrichtungen, wie des Fraunhofer-Instituts für offene Kommunikationssysteme FOKUS, und verschiedener Unternehmen der IT- und Politik-/Verwaltungsberatung sind ebenfalls einbezogen.

Die verschiedenen genannten Akteure kommen bei einer regelmäßig stattfindenden, informellen Veranstaltung, dem Berliner Open Data-Stammtisch, zusammen und bilden mittlerweile eine lose Aktionsplattform. Aus diesen Treffen heraus werden weitere lokale Aktivitäten zur Bereitstellung von Verwaltungsdaten organisiert.

Im Mai 2011 konnte der Berliner Open Data Day (BODDy)⁹ von Open Data-Interessierten gemeinsam organisiert und durchgeführt werden. Mit dieser Veranstaltung wurden speziell Mitarbeiter und Mitarbeiterinnen der Berliner Verwaltung auf Landes- und Bezirksebene angesprochen. Zielsetzung der Veranstaltung war es, die Mitarbeiter über das Thema zu informieren, für die Potenziale offener Verwaltungsdaten zu werben, Akteure zu vernetzen und den Status quo offener Verwaltungsdaten in einer lokalen Leistungsschau aufzuzeigen. Für die halbtägige Veranstaltung mit circa 120 Teilnehmern konnten der Berliner Wirtschaftssenator und der IT-Staatssekretär der Senatsverwaltung für Inneres sowie diverse engagierte verwaltungsinterne und -externe Einrichtungen und Einzelpersonen gewonnen werden, die mit Verwaltungsdaten unterschiedliche Themen bearbeiten.

Zum Berliner Open Data Day entwickelten die verschiedenen Stakeholder des Berliner Open Data-Stammtisches eine gemeinsame Agenda, in der die aus ihrer Sicht notwendigen Schritte hin zu offenen Daten in der Berliner Verwaltung stichpunktartig beschrieben wurden. Dieses Papier wird von der Aktionsplattform als Werkzeug für die Information über offene Verwaltungsdaten, für die Gewinnung von Unterstützung und für Marketingzwecke eingesetzt¹⁰.

Bereits im Herbst 2010 wurde von der Berliner Senatsverwaltung für Wirtschaft, Technologie und Frauen eine Onlinebefragung zu Verwaltungsdaten durchgeführt. Hierdurch sollte ermittelt werden, an welchen Verwaltungsdaten Bürgerinnen und Bürger überhaupt interessiert sind und ob sie sich vorstellen können, solche Daten für eigene Zwecke einzusetzen. An der Befragung nahmen ungefähr 1300 Personen teil. Im Ergebnis besteht das größte Interesse an Daten der Stadtplanung, allgemeinen Verwaltungsinformationen und Umweltdaten (Berliner Senatsverwaltung für Wirtschaft, Technologie und Frauen, 2010).

⁹<http://berlin.opendataday.de>.

¹⁰Aktionsplattform Open Government Data, 2011: Berliner Open Data-Agenda, <http://berlin.opendataday.de/agenda>.

Ebenfalls im Herbst 2010 startete die Berliner Senatsverwaltung für Wirtschaft, Technologie und Frauen einen Wettbewerb zur Entwicklung von mobilen Softwareanwendungen (Apps4Berlin), die u. a. Daten der öffentlichen Verwaltung einsetzen. Ausgelobt wurden Preise in den Kategorien Studierende, Unternehmen, Forschungseinrichtungen und Vereine. Eine mehrköpfige Jury wählte auf Basis eines zuvor festgelegten Kriterienkatalogs die Sieger in den jeweiligen Kategorien aus. Als Gewinner bestimmt wurden Ideen oder Anwendungen zur Stadt- und Umweltentwicklung, ein multilinguales Lern- und Informationssystem sowie ein Informationssystem zur barrierefreien Stadt (Berliner Senatsverwaltung für Wirtschaft, Technologie und Frauen, 2011).

Im Wahlkampf zur Berliner Abgeordnetenhauswahl 2011 äußerten sich mehrere Parteien in ihren Programmen positiv zu mehr Transparenz mittels offener Daten in der Berliner Verwaltung. Unter anderem forderte die Berliner SPD, dass Politik und Verwaltung sich den Kriterien des offenen Regierens und Verwaltens verschrieben. Dabei sollten »Informationen und Dienste der Verwaltung umfassend, aktuell, vollständig und barrierefrei online verfügbar gemacht werden« sowie öffentliche Daten über Schnittstellen für Bürgerinnen und Bürgern bereitgestellt werden (SPD Berlin, 2011, S. 28 f).

Die aus den letzten Wahlen hervorgegangene Regierungskoalition aus SPD und CDU hat in ihrem Koalitionsvertrag angekündigt, die Open Data-Initiative des Landes fortzusetzen und auszubauen. Hierzu soll die »weitgehende Offenlegung von öffentlichen Daten (z. B. Geoinformationsdaten) unter Wahrung des persönlichen Datenschutzes« geprüft werden (SPD Berlin und CDU Berlin, 2011, S. 94).

2.2.3 Internationale Beispiele für die Bereitstellung offener Verwaltungsdaten

Auf Ebene der Europäischen Union existieren verschiedene Aktivitäten zur Bereitstellung offener Daten. Neben den bereits genannten EU-Richtlinien zur Weiterverwendung von Verwaltungsdaten und verschiedenen EU-Projekten (Open Cities, ENGAGE, LOD2, LAPSI etc.) hat die Europäische Union offene Daten in verschiedenen strategischen Politikinitiativen aufgegriffen:

- der Ministerdeklaration zu eGovernment von Malmö, in der die vermehrte Bereitstellung von Verwaltungsdaten zur Weiterverwendung und eine Erhöhung der Transparenz administrativer Vorgänge und Entscheidungen angekündigt wird (EU Presidency, 2009, S. 3);
- der Deklaration zur Digitalen Agenda mit der Forderung, Verwaltungsdienstleistungen zu entwickeln, die ein offenes und transparentes Handeln von Regierung und Verwaltung unterstützen (Spanish EU Presidency, 2010, S. 3);
- der Digital Agenda Assembly der EU-Kommission, in deren Rahmen offene Daten als eine der sogenannten »big ideas« verfolgt werden.

Datenportale, über die der Zugang zu offenen Verwaltungsdaten realisiert wird, funktionieren wie Kristallisationspunkte für die verschiedenen Bemühungen, eine stärkere Öffnung staatlicher Datenbestände zu bewirken. Daher werden hier Beispiele für Portalinitiativen auf nationaler und lokaler Ebene sowie für entsprechende Fachportale kurz vorgestellt.

Das Datenportal der britischen Regierung `data.gov.uk` wurde im Frühjahr 2010 für die Öffentlichkeit freigeschaltet. Vorausgegangen waren ein Forschungsprojekt und eine Planungsphase. In dem Forschungsprojekt AKTive PSI wurde die allgemeine Tauglichkeit von Semantic Web-Ansätzen für die Verwendung im öffentlichen Sektor geprüft (Shadbolt, O'Hara, Salvadores und Alani, 2011, s. 866). Aufbauend auf den Erkenntnissen, wurde die Idee des einheitlichen Einstiegspunkts für offene Verwaltungsdaten in Großbritannien weiterverfolgt.

Heute werden über das Portal circa 5700 Datensätze bereitgestellt. Dem Ansatz folgend, erst Basisdaten online zur Verfügung zu stellen und sich anschließend um die Aufbereitung in das bestmögliche Format zu kümmern, werden viele dieser Daten in Formaten wie XML, CSV, XLS, PDF oder DOC zur Verfügung gestellt. Der geringere Prozentsatz der Datensätze ist in Form von aufbereiteten Linked Data verfügbar.

Das Portal bietet semantische und nichtsemantische Funktionen (Shadbolt et al., 2011, S. 876):

semantic Web-Funktionen

- SPARQL-Endpunkte
- Datensätze im Linked Data-Format
- Richtlinien zur Erzeugung und Bereitstellung von Linked Data

nicht-semantische Funktionen

- Diskussionsforum und ein Wiki zur Unterstützung der Zusammenarbeit der Datennutzer und -bereitsteller
- Funktionen zu Suche und Browsen von Datensätzen
- Verzeichnis von Ideen und Anwendungen, die auf offenen Verwaltungsdaten aus `data.gov.uk` aufbauen

Beispiele für lokale Aktivitäten gibt es mittlerweile zahlreiche. So wurde etwa das Datenportal »DataSF«¹¹ der kalifornischen Stadt San Francisco im Jahr 2009 gestartet. Es umfasste Anfang 2011 etwa 200 Datensätze in unterschiedlichen Formaten. Vorhandene Datensätze und Anwendungen, die auf den Verwaltungsdaten aufbauen, können von Nutzern in ihrer Qualität mit einem einfachen Sternesystem bewertet werden.

Im deutschsprachigen Ausland nehmen die Aktivitäten der Stadt Wien hinsichtlich Open Government Data eine wichtige Rolle ein. Am 17. Mai 2011 hat die Stadt Wien als erste

¹¹<http://datasf.org/>.

deutschsprachige Kommune überhaupt ein Portal für offene Verwaltungsdaten gestartet.¹² Über das Portal können Daten aus Bereichen wie Bevölkerung, Verkehr, Budget und Wirtschaft in unterschiedlichen Formaten wie CSV, RSS oder JSON abgerufen und in Abhängigkeit von der Lizenz weiterverwendet werden. Zu Beginn ging das Portal mit etwa 30 Datensätzen an den Start und enthielt im Oktober 2011 über 70 Datensätze. Beispiele einzelner Datensätze sind ein Baumkataster, Standorte von Fahrradabstellanlagen und die Möglichkeit, auf den offiziellen Stadtplan über eine Programmierschnittstelle zuzugreifen. Die Datensätze wurden unter der Lizenz Creative Commons Namensnennung 3.0 Österreich veröffentlicht. Diese international anerkannte Lizenz der gemeinnützigen Creative Commons-Organisation erlaubt die freie Weiterverwendung des jeweiligen Inhalts unter der Voraussetzung, dass die Quelle der Inhalte genannt wird.

Ein Beispiel für ein Portal für offene Verwaltungsdaten einer Fachbehörde bietet die Weltbank. Diese Einrichtung stellt seit Mitte des Jahres 2010 große Mengen an Daten über ein spezielles Portal zur Verfügung. Unter data.worldbank.org macht die Weltbank ihre statistischen Indikatoren verfügbar, mit denen die Mitarbeiter der Einrichtung arbeiten, um die Entwicklung einzelner Länder oder ganzer Regionen zu beurteilen und miteinander zu vergleichen. Gegenwärtig sind über die Programmierschnittstelle der Weltbank ungefähr 2000 Datensätze in Formaten wie XML, CSV, KML als RSS oder JSON verfügbar.

2.2.4 Beispiele für die wirtschaftliche Weiterverwendung von Verwaltungsdaten

Auf der Seite der Datennutzer finden sich verschiedene Beispiele der erfolgreichen Nutzung offener Daten. Gerade der Privatsektor hat in einigen Fällen von Daten des öffentlichen Sektors profitiert.

Die Bereitstellung von Informationen des Global Positioning Systems hat eine Fülle von zivilen Applikationen ermöglicht, die ohne diese Geoinformationen kaum oder nicht möglich wären. Im Zusammenhang mit den frei nutzbaren Geodaten des Communityprojektes OpenStreetMap (OSM) können diese Geoinformationen mittlerweile von jedermann ohne Berücksichtigung restriktiver Lizenzen für eigene Projekte und Produkte eingesetzt werden. So entwickeln beispielsweise Unternehmen Navigationssoftware für aktuelle Smartphones auf Basis der OSM. Diese Software kann einfach über existierende Vertriebsplattformen der Smartphonebetriebssysteme bereitgestellt werden. Auf Basis offener Verwaltungsdaten versuchen Unternehmen, offene und nicht offene Daten zu integrieren, Daten zu veredeln und einen einheitlichen Zugriff auf Datensätze über verschiedene Datenbereitsteller hinweg anzubieten. Beispiele für diesen Ansatz sind das französische Unternehmen Data Publica¹³ und die isländische Firma DataMarket¹⁴. Beide Unternehmen kombinieren offene Verwaltungsdaten mit Daten und Informationen aus

¹²<http://data.wien.gv.at>.

¹³<http://www.data-publica.com>.

¹⁴<http://www.datamarket.com>.

anderen Quellen und bieten sie über eine einheitliche Schnittstelle an. In Abhängigkeit von ihrer Herkunft sind die bereitgestellten Daten teilweise kostenfrei nutzbar.

Ein weiteres Beispiel für die positiven Wirkungen von Daten öffentlicher Einrichtungen ist die Bereitstellung von Verkehrs- und Wetterdaten. Gerade Letztere finden vielfache Verwendung durch private Anbieter von Wetterinformationen. Die Bereitstellung von Wetterdaten als offene Daten dürfte diesen Markt verändern. Etablierte Akteure, die bislang in der Lage waren, kostenpflichtige Wetterdaten zu erwerben, sehen sich zukünftig neuer Konkurrenz durch Kleinunternehmen gegenüber, für die der Erwerb meteorologischer Informationen bislang nicht in Frage kam. Diese Veränderungen im Markt der wetterbasierten Dienstleistungen (Etablierung neuer Akteure) sind kennzeichnend für die Effekte der Bereitstellung offener Daten (siehe hierzu auch Abschnitt 2.1.4).

Ähnlich verhält es sich mit der Bereitstellung von Verkehrsdaten wie Fahrplandaten, aktuellen Stauinformationen oder Qualitätsdaten zu einzelnen Verkehrsträgern. In Fällen, in denen kommunale Verkehrsbetriebe Zugang zu ihren Datensätzen gewähren, entwickelt sich ein Ökosystem verschiedener Dienstleister und Anwendungen auf Basis dieser Daten. Dies reicht von Webanwendungen mit Pünktlichkeitsinformationen einer Buslinie bis zu dynamischen Fahrplanauskünften für Mobiltelefone über verschiedene Gebietskörperschaften hinweg.

Neben privaten Unternehmen profitieren auch öffentliche Organisationen von der Bereitstellung maschinenlesbarer Verwaltungsdaten. Die Informationen zu den Standorten der Glascontainer, die von verschiedenen Unternehmen betrieben werden, helfen der Berliner Stadtreinigung bei der Identifikation und anschließenden Reinigung ebendieser Standorte.

Nachrichtenorganisationen wie die britische Tageszeitung *The Guardian*, aber auch deutsche Lokalzeitungen verwenden offene Daten im Rahmen ihrer Berichterstattung. Gleichzeitig entwickeln sich erste Unternehmen, die auf die neue Form des Datenjournalismus spezialisiert sind und ausgefeilte Analysen und Visualisierungen auf Basis frei verfügbarer Daten anbieten.

Die britische Regierung sieht noch großes Potenzial in der Weiterverwendung offener Verwaltungsdaten. Entsprechend hat der britische Schatzkanzler George Osborne im Herbst 2011 die Bereitstellung weiterer Datensätze und die Unterstützung von Unternehmen bei der Verwendung von Daten des öffentlichen Sektors angekündigt (HM Treasury, 2011, s. 40).

2.3 Zusammenfassende Beurteilung

Der Ansatz offener Daten hat seit seiner ersten Diskussion in Fachkreisen wie z. B. der Sebastopol-Gruppe im Jahre 2007 (Sebastopol Group, 2007) zu einer weiten Verbreitung gefunden. Nicht zuletzt durch die Initiativen für mehr Transparenz in den USA

oder Großbritannien in den letzten Jahren stieß der Ansatz bei immer mehr zivilgesellschaftlichen Organisationen und Verwaltungseinrichtungen auf Interesse.

Gleichzeitig hat die technische Fortentwicklung gerade von Webtechnologien dazu geführt, dass auch kleinere Organisationen und einzelne Entwickler in der Lage sind mit vergleichsweise geringem Know-how, maschinenlesbare Daten weiterzuverarbeiten und ihren eigenen Zwecken entsprechend einzusetzen. Aus diesen zivilgesellschaftlichen Gruppen entstand in den letzten Jahren eine verstärkte Nachfrage nach Daten des öffentlichen Sektors unter den Prinzipien von Open Data.

Unternehmen können ebenfalls von dem Ansatz offener Daten profitieren, und zwar gerade kleine und mittlere Unternehmen, für die sich der Kauf von Public Sector Information eventuell bislang nicht gelohnt hat, die aber bei Daten zu Marginalkosten in der Lage sind, diese zu erwerben und weiterzuverwenden.

Die mit dem Ansatz offener Daten vielfach einhergehende Veröffentlichung über ein zentrales Datenregister ist auch in der Lage, die Arbeit der Verwaltungsmitarbeiter zu vereinfachen und die Zusammenarbeit zu intensivieren. Datensätze einer Fachbehörde lassen sich einfacher von Mitarbeitern anderer Behörden auffinden und für das eigene Aufgabenfeld einsetzen, als dies bislang der Fall war.

Der Ansatz, Daten speziell des öffentlichen Sektors aus rechtlicher, finanzieller und technologischer Sicht so einfach wie möglich weiterverwendbar zu offerieren, bietet grundsätzlich Potenzial, dazu beizutragen, Transparenz im öffentlichen Sektor und eine größere Verbreitung der Daten zu befördern. Zur Unterstützung von Wirtschaftsunternehmen mit Daten zur Fortentwicklung ihrer bestehenden Geschäftsmodelle oder zur Entwicklung neuer Ideen bietet angesichts knapper öffentlicher Kassen eine zusätzliche Möglichkeit der Wirtschaftsförderung.

Für die Berliner Verwaltung ergibt sich mit den weiter oben geschilderten Initiativen zur strukturierten Aufbereitung, Bereitstellung und Weiterverwendung von Daten öffentlicher Einrichtungen folgendes Bild (siehe auch Abschnitt 6.1.3):

Die beschriebenen Initiativen in Berlin zur Modernisierung der Verwaltung und zu einer verbesserten Handhabung und Analyse sozialraumorientierter Daten und geographischer Informationen leisten wertvolle Arbeit. Der Ansatz, dass Daten, die an einer Stelle der Verwaltung entstehen, sinnvoll auch an anderer Stelle oder von Unternehmen oder Bürgerinnen und Bürgern eingesetzt werden können, findet sich hier immer wieder und sollte, soweit möglich und sinnvoll, gebündelt und ausgebaut werden.

Eine Verbreitung von Informationen zu Open Data wurde mit verschiedenen Aktivitäten erreicht. Mit der Durchführung des BODDy, den regelmäßigen Treffen des Stammtisches und dem Entwicklungswettbewerb konnte die Auseinandersetzung mit dem Thema über einen kleinen Kreis engagierter Experten hinaus vergrößert werden.

Während der lokalen Leistungsschau des Berliner Open Data Days konnte gezeigt werden, dass der Zugang zu Daten der Berliner Verwaltung die Entwicklung innovativer Anwendungen möglich macht.

Über die Informationsaktivitäten zur Bereitstellung offener Daten ist es gelungen, die positiven Aspekte des Ansatzes offener Verwaltungsdaten vorzustellen und mögliche negative Aspekte mit Akteuren aus Verwaltung, Zivilgesellschaft, Wirtschaft und Wissenschaft zu diskutieren.

Gleichzeitig konnten so die verschiedenen relevanten Akteure rund um die Bereitstellung von Verwaltungsinformationen stärker vernetzt werden, als dies bislang der Fall war.

Über die öffentliche Thematisierung der Potenziale und Grenzen offener Daten mithilfe der verschiedenen Projekte und Veranstaltungen, der Befragung und weiterer Aktivitäten konnten zusätzliche Akteure innerhalb und außerhalb der Berliner Verwaltung identifiziert und eingebunden werden.

Die in Berlin seit dem Herbst 2010 offen geführte Diskussion zwischen Verwaltungsvertretern und vor allem zivilgesellschaftlichen Organisationen, aber auch der Wirtschaft und der Wissenschaft hat zu einem verfeinerten Verständnis offener Daten und der damit verbundenen Potenziale und Barrieren unter den Beteiligten geführt.

Sowohl untergesetzliche als auch gesetzliche Regelungen auf Landes- und Bezirksebene zur Bereitstellung von Daten für unterschiedliche Zielgruppen existieren in Berlin gerade in den Bereichen raum- und umweltbezogener Daten. Die Umsetzungen, bei denen bereits speziell Aspekte offener Daten aufgegriffen werden, sind jedoch noch ausbaufähig. Regelungen zur proaktiven Bereitstellung von Daten für Bürgerinnen und Bürger oder ein landesweites Lizenzregularium zur Verwendung durch Verwaltungseinheiten fehlen zum jetzigen Zeitpunkt noch.

Zur Umsetzung des Ansatzes offener Daten ist vielfach eine Anpassung existierender Informations- und Datenprozesse notwendig. Diese Geschäftsprozesse sind in den verschiedenen Ressorts in unterschiedlicher Qualität dokumentiert, teilweise fehlt eine Beschreibung vollständig. Dies erschwert die Identifikation und Änderung anzupassender Prozesse.

Bei Bereitstellung von Informationen spielt die Überprüfung der Datenqualität eine wichtige Rolle. Hierzu existiert in der Berliner Verwaltung bislang kein etablierter Ansatz der Qualitätssicherung, es finden sich nur ressortspezifische Herangehensweisen von unterschiedlicher Güte.

Zur mittel- bis langfristig erfolgreichen Entwicklung der Berliner Verwaltung hin zu einer verstärkten und routinemäßigen Bereitstellung von Verwaltungsdaten als offene Daten ist die Einbettung in eine Gesamtstrategie für offenes Regieren und Verwalten notwendig. Informationen der öffentlichen Verwaltung stellen in vielen gesellschaftlichen Entwicklungen eine wichtige Basis dar. Da sie daher vielfach Mittel zum Zweck sind, ist die Entfaltung einer Vorstellung von offenem Regierungshandeln als strategischer Option zur Lösung komplexer Herausforderungen wünschenswert.

Stärken	Schwächen
• intersektorelle Diskussion zu offenen Daten • breite lokale Akteurskonstellation • diverse lokale Beispiele für die Weiterverwendung offener Daten • Existierende Initiativen zur verwaltungsinternen Bereitstellung von Daten • Einbindung in nationale und internationale Open Data-Aktivitäten	• ausbaufähige gesetzliche Regelungen auf Landesebene zur proaktiven Datenbereitstellung • fehlende Berliner Gesamtstrategie für offenes Regieren und Verwalten • fehlende Dokumentation von Datenprozessen • keine etablierte Qualitätssicherung
Chancen	Risiken
• Entwurf einer Open Data-Strategie • erstes Bundesland mit einem Open Data-Portal • Notwendigkeit strukturierter Daten auch in verwandten Projekten erkannt • Open Data als Thema auf nationaler und internationaler Ebene • Landespolitik sieht Chancen von Open Data und Open Government	• Ausbreitung in die Fläche der Berliner Verwaltung • Veränderungen in der Landesregierung • tradierte Verwaltungskultur • schleppende Anpassung des rechtlichen Umfeldes • Vertrauensverlust durch verzögerte Umsetzung

Tabelle 2.2: SWOT-Analyse der Rahmenbedingungen im Land Berlin

Der Berliner Verwaltung liegt mit dieser Untersuchung der Entwurf einer Open Data-Strategie vor. Anhand der Analyseergebnisse und der abgeleiteten Handlungsempfehlungen kann eine Fortentwicklung der Bereitstellung offener Daten durch die Landesregierung unternommen werden.

Das Land Berlin ist das erste Bundesland mit einem offiziellen Portal für offene Verwaltungsdaten und kann von der damit einhergehenden Aufmerksamkeit der anderen föderalen Akteure profitieren. Gleichzeitig ist es gelungen, über internationale kommunale Netzwerke und EU-Projekte in einen Erfahrungsaustausch einzusteigen und von der »good practice« anderer Städte zu profitieren.

Die Vorteile strukturierter Daten wurden auch bereits in verwandten Initiativen in Berlin erkannt und werden dort teilweise schon seit mehreren Jahren detailliert verfolgt. Dies

bietet die Möglichkeit, Synergieeffekte im Hinblick auf Interoperabilität, thematische Aufmerksamkeit und Ressourcenbereitstellung der unterschiedlichen Projekte nutzbar zu machen.

Neben Akteuren der Berliner Verwaltung hat auch die Landespolitik die Chancen offener Daten als Bestandteil einer offenen Regierung und Verwaltung erkannt und sich zu eigen gemacht. Mehrere Landesparteien haben sich die Bereitstellung von Verwaltungsdaten für Dritte in ihre Parteiprogramme aufgenommen.

Im Herbst 2011 fand ein Wechsel der Berliner Landesregierung statt. Zum einen stellen Führungswechsel generell eine Herausforderung für den Verwaltungsapparat dar. Über einen gewissen Zeitraum werden manche Vorgänge und Entscheidungen aufgeschoben. Es kann daher in der praktischen Fortführung des Ansatzes offener Daten in Berlin zu Diskontinuitäten kommen. Zum anderen setzen neue Regierungskoalitionen mitunter neu Schwerpunkte in ihrem Handeln. Inwieweit dies die weitere Umsetzung von Open Data im Land Berlin betrifft, bleibt abzuwarten.

Die Anpassung des rechtlichen Umfelds der Bereitstellung offener Daten sollte in Relation stehen zu organisatorischen und technischen Veränderungen des öffentlichen Sektors. Rechtliche Regelungen können bei rechtzeitiger Umgestaltung als Wegbereiter für die Bereitstellung offener Daten wirken.

Der Wandel der existierenden Verwaltungskultur, nicht nur in Berlin, stellt sicher eine der größten Herausforderungen bei der Realisierung des freien Zugangs zu offenen Daten dar. Mit der bisherigen über 200-jährigen deutschen Verwaltungstradition sind einzelne Teilaspekte des offenen Regierungs- und Verwaltungshandelns, zu dem Open Data gezählt werden, nur mit Aufwand und Veränderungen vereinbar. Hier sind neben verschiedenen Projekten zur Öffnung von Verwaltung und Politik weitere Kommunikationsanstrengungen, die entsprechende Qualifizierung der Mitarbeiter sowie übergreifende Vorgaben der politisch-administrativen Spitze zielführend. Die Umsetzung benötigt ein stringentes Vorgehen und eine umfassende Vermittlung der Fortschritte und etwaiger Verzögerungen, um nicht an Vertrauen bei betroffenen Akteuren zu verlieren. Gleichzeitig bedarf es einer Entwicklung von unten her, die mit Veranstaltungen, Beispielanwendungen, unterschiedlichen Projekten und der Zusammenarbeit von Verwaltungsmitarbeitern, Zivilgesellschaft und weiteren Akteuren dazu beiträgt, eine offenerere, transparentere öffentliche Verwaltung zu erreichen.

3

Organisationsanalyse

Das Vorgehen bei der Organisationsanalyse umfasste zunächst eine Aufnahme des Status quo bei der Veröffentlichung von Daten der Berliner Behörden und eine Feststellung der notwendigen Änderungen im Hinblick auf die organisatorischen Gegebenheiten, die Qualitätssicherung, eventuelle zusätzliche Aufwände und die Ausbildung der Mitarbeiter. Aufbauend auf den jeweiligen Istzuständen wurden dann passende Sollprozesse entworfen.

Für die Erhebung von Istprozessen und möglicherweise existierenden Barrieren für Open Data in Berlin wurde zunächst ein Interviewleitfaden erstellt (siehe Anhang A), um Mitarbeiter der Behörden in qualitativer, strukturierter Weise zu befragen. Auf Grundlage dieses Leitfadens wurden circa 90-minütige Gespräche mit ausgewählten Vertretern von Berliner Institutionen mit öffentlichem Auftrag geführt: der Senatsverwaltung für Bildung, Wissenschaft und Forschung, dem Amt für Statistik Berlin-Brandenburg, der Zentral- und Landesbibliothek sowie der »Umwelt AG«, der Initiative für das Umweltportal des Landes und der Bezirke¹.

Dabei wurde zunächst je nach Bedarf eine Einführung in das Thema Open Data geliefert. Anschließend wurde erfasst, wie Veröffentlichungsprozesse typischerweise ablaufen. Abschließend wurde erörtert, wo aus Sicht des Interviewpartners Potenziale und Risiken für das betreffende Haus liegen.

In den folgenden Abschnitten sind für jeden Interviewpartner dargestellt:

- der Veröffentlichungsprozess in schematischer und tabellarischer Form

¹<http://www.berlin.de/umwelt>.

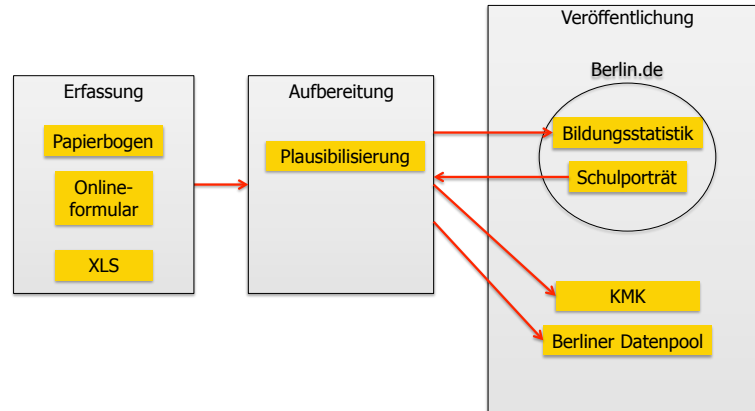


Abbildung 3.1: Schema des Ist-Veröffentlichungsprozesses bei der SenBWF

- der abgeleitete Sollprozess als Schema
- das Potenzial an Datensätzen, die unter Umständen veröffentlicht werden könnten
- Chancen und Risiken aus Sicht der Institution
- weitere Erkenntnisse

Abschließend werden die Vorschläge dahingehend analysiert, wie sich die Rahmenbedingungen für Open Data-Veröffentlichungen verbessern lassen (Abschnitt 3.5).

3.1 Senatsverwaltung für Bildung, Wissenschaft und Forschung

Bei der Senatsverwaltung für Bildung, Wissenschaft und Forschung (SenBWF) wurde die Abteilung Schule ausgewählt. Die erfassten Prozesse wurden tabellarisch festgehalten. Auf Basis der tabellarischen Prozesserhebung wurde eine schematische Grobdarstellung des Istprozesses abgeleitet.

Die Erfassung von Daten erfolgte auf drei Wegen: per Papierbogen, per Microsoft-Excel-Tabelle und per Onlineformular (vgl. Tabelle 3.1, Nummer 1). In der schematischen Darstellung in Abbildung 3.1 ist dieser Vorgang auf der linken Seite dargestellt.

Nach der Erfassung erfolgte der Schritt der Plausibilisierung, die vor allem in der Prüfung von Klassensummen besteht und in der Überführung in eine elektronische Datenbank bestand (vgl. Tabelle 3.1, Nummer 2, bzw. Abbildung 3.1, Mitte).

Die so konsolidierten Daten wurden auf verschiedenen Wegen veröffentlicht (vgl. Tabelle 3.1, Nummer 3, bzw. Abbildung 3.1, Mitte), die im Folgenden aufgelistet sind.

Nr.	Arbeitsaufgabe	Stelle	Speicher	Gebühr	Fristen	Ergebnisse	Medium	Speicher
1 a)	Datenerfassung per Papierbogen							
1.1 a)	Schulleiter füllen Fragebogen aus	Schulen			2 Wochen		Papier-bögen	
1.2 a)	Sichtung und manuelle Plausibilisierung	Statistik- abteilung der SenBWF	Papierbögen			konsolidierte Rohdaten	Papier-bögen	
1.3 a)	ggf. postalische Rückfrage	Statistik- abteilung der SenBWF	Papierbögen			konsolidierte Rohdaten	Papier-bögen	
1.4 a)	Digitalisierung der Rohdaten	Statistik- abteilung der SenBWF	konsolidierte Papierbögen					XLS
1 b)	Datenerfassung per XLS-Tabelle							
1.1 b)	Schulen füllen via WWW verteilte Tabelle aus	Schulen			2 Wochen			XLS
1.2 b)	Import der Tabellen	Statistik- abteilung der SenBWF	XLS					XLS
1 c)	Datenerfassung per Online-formular							
1.1 c)	Schulleiter füllen Online-formular	Schulen	WWW		2 Wochen			DB auf WWW- Server
1.2 c)	Onlineüberprüfung der Formulardaten	WWW-Server	WWW			konsolidierte Rohdaten		DB auf WWW- Server
2	Elektronische Plausibilisierung							
2.1	Automatisierte Plausibilitätsprüfung mit Visual-Basic-Programmen, vor allem Korrektheit von Summen	Statistik- abteilung der SenBWF	XLS, externe DB			Plausibili- sierte Daten		XLS, externe DB
2.2	Übertragung in interne Datenbank	Statistik- abteilung der SenBWF	XLS, externe DB					interne SQL-DB
3	Datenaufbereitung							
3.1 a)	Einzelanfragen von Presse, Verwaltung, Politik, Eltern beantworten	Statistik- abteilung der SenBWF	interne SQL- DB	unentgeltlich		Spezial- statistik	E-Mail	PDF, XLS etc.
3.1 b)	Jährliche Statistik erstellen	Statistik- abteilung der SenBWF	interne SQL- DB	unentgeltlich		Standard- produkt		PDF, XLS etc.
4	Veröffentlichung							
4.1	Der Berlin.de-Redakteur der SenBWF veröffentlicht eine Statistik mit dem Imperia CMS	Statistik- abteilung der SenBWF					WWW	PDF, XLS etc.

Tabelle 3.1: Groberfassung des Istprozesses bei der SenBWF

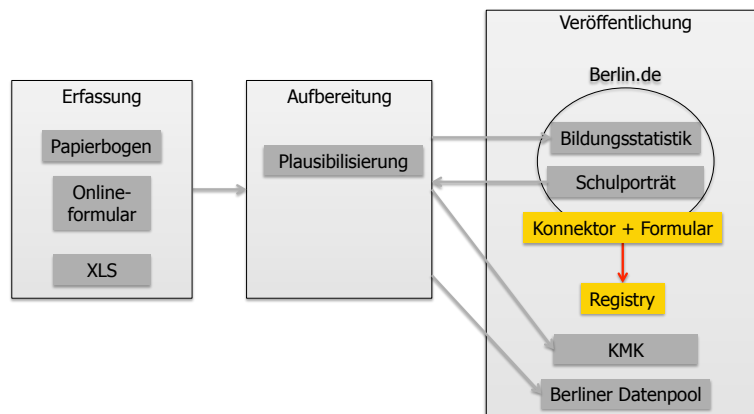


Abbildung 3.2: Schema des Soll-Veröffentlichungsprozesses bei der SenBWF

- Ein Berlin.de-Redakteur lädt Berichte durch das Imperia-CMS zu Berlin.de hoch, meist als PDF.
- Statistiken werden strukturiert und maschinenverarbeitbar an die entsprechende Stelle der Kultusminister-Konferenz weitergegeben, um dort bundeseinheitlich verarbeitet zu werden.
- Statistiken werden strukturiert und maschinenverarbeitbar an den Berliner abgestimmten Datenpool weitergegeben, also an ein verwaltungsinternes Medium.
- Eine weitere Möglichkeit für Datennutzer ist der Zugriff über eine Suchmaske auf Berlin.de². Hier können jedoch nur Einzelanfragen gestellt werden, eine programmatische Anbindung oder ein Herunterladen sind nicht möglich.

Abbildung 3.2 skizziert, wie der Veröffentlichungsprozess im Sinne von Open Data umorganisiert werden müsste: Die Redakteure, also Mitarbeiter der SenBWF, die für die Veröffentlichung von Berichten auf Berlin.de zuständig sind, besitzen eine Schlüsselposition. Über einen entsprechenden Konnektor bzw. ein erweitertes Formular könnten sie zusätzlich zum Hochladen die entsprechende Datei in einem zentralen Register anmelden.

Eine weitere notwendige Änderung betrifft das Format, in dem Berichte und darin implizit enthaltene Daten veröffentlicht werden. Da PDF-Dokumente praktisch keine Möglichkeit der automatischen Verarbeitung zulassen, ist es essenziell, dass die Daten möglichst unverarbeitet in einem strukturierten Format abgelegt werden.

Weitere, für ein Open Data-Portal interessante Datensätze wären Zahlen zu Abgängern, Lernmittelbefreiung, Bildungsgangempfehlung und Wiederholern.

Bei der umfassenden Veröffentlichung von Rohdaten sehen die Mitarbeiter der SenBWF folgende Risiken und Schwierigkeiten:

Auf mehreren Ebenen ist die Gefahr des Missbrauchs gegeben. Dazu gehören die Fehl- oder Umdeutung von Daten sowie die Zweckentfremdung z. B. für Werbung. Außerdem bereiten Datenschutzfragen Sorgen: Zwar werden nur Klassensummen erfasst, bei kleinen Klassen und speziellen Merkmalen lässt sich jedoch ein Personenbezug nicht ausschließen. Zusätzlich bleibt unklar, wie mit Fehlern umgegangen wird, die durch die Veröffentlichung entstehen bzw. bekannt werden.

In Anbetracht dieser Bedenken fehlt der Behörde ein politisches Mandat für Open Data, d. h., man kann die genannten Risiken nicht eingehen, ohne sich im Zweifelsfall auf eine politische Rückendeckung verlassen zu können.

Zusammenfassend lässt sich zur SenBWF festhalten:

- Viel wird schon veröffentlicht, aber nicht mit Open Data-Formaten und -Lizenzen.
- Viele weitere Datensätze sind für Open Data interessant.

²http://www.berlin.de/sen/bildung/schulverzeichnis_und_portraits/anwendung.

- Eine zentrale Beratungsstelle fehlt.
- Die politische Vorgabe für Open Data fehlt.

3.2 Amt für Statistik Berlin-Brandenburg

Das Amt für Statistik Berlin-Brandenburg (AfS) ist »der zentrale Dienstleister auf dem Gebiet der amtlichen Statistik für die Region Berlin-Brandenburg [...]. Es werden insgesamt über 250 Bundes- und EU-Statistiken sowie 25 Landesstatistiken aus allen Bereichen des gesellschaftlichen Lebens – auch für Wahlen und Volksabstimmungen – vom Amt für Statistik Berlin-Brandenburg erhoben, aufbereitet, zusammengefasst und ausgewertet.«³

Zur Erhebung und Veröffentlichung dieser Daten ist das Amt durch zahlreiche Gesetze und Verordnungen verpflichtet. Außerdem übernimmt es die zentrale Aufgabe, die entsprechenden Geheimhaltungsvorschriften umzusetzen. Das bedeutet, dass personenbezogene, also sensible Einzeldaten zu allgemein zugänglichen Berichten aggregiert werden. Der Zugang ist dabei je nach Anonymisierungsgrad gänzlich oder nur Personen verfügbar, die sich verpflichten, sich an bestimmte Auflagen zu halten, z. B. Wissenschaftlern.

Abbildung 3.3 skizziert den derzeitigen Veröffentlichungsprozess. Die Erfassung der Daten lässt sich im Rahmen dieser Studie nicht erschöpfend beleuchten. Hier kommen viele verschiedene Medien und Verfahren zum Einsatz, von ausgezählten Wahlzetteln bis zur Onlineerfassung von Eingaben. Zusätzlich zur Plausibilisierung werden die Daten in der Aufbereitungsphase entsprechend den den jeweiligen Gesetzen und Verordnungen aggregiert und entsprechend den jeweiligen Geheimhaltungsvorschriften anonymisiert.

Gegenwärtig unterstützt das AfS verschiedene Möglichkeiten zur Veröffentlichung von Datensätzen. Die verschiedenen Varianten sind in Tabelle 3.2, Nummer 2, aufgelistet. Im einfachsten Fall wird ein Standardprodukt auf den Internetseiten des AfS abgelegt. Für viele Daten wird hier schon das maschinenverarbeitbare XLS-Format neben PDF angeboten. Nutzungsrechte werden zwar eingeräumt, jedoch sind die Datensätze nicht mit einer freien Lizenz versehen, die für Rechtssicherheit bei Bürgern und Entwicklern sorgt. Auf demselben Weg werden Public Use Files bereitgestellt, also für die Öffentlichkeit nutzbare Datendateien, die mit verbreiteten Statistikprogrammen geöffnet werden können.

Wissenschaftler, die sich zur Wahrung des Datenschutzes verpflichten, haben zusätzlich Zugriff auf detailliertere Produkte. Außerdem können sie in Datenforschungszentren selbst recherchieren und per Datenfernabfrage Einzelanfragen stellen. Für diese Abfragen wird in jedem Einzelfall sichergestellt, dass alle Geheimhaltungsvorschriften berücksichtigt werden.

³<http://www.statistik-berlin-brandenburg.de/einzelseiten/inhalt-ueber-uns.asp>.

Nr.	Arbeitsaufgabe	Stelle	Input			Vorschriften	Gebühr	Output		
			Information	Medien	Speicher			Ergebnisse	Medien	Speicher
1	Statistik entsprechend Gesetz oder Verordnung erheben									
1.1	Erfassung der Einzeldaten	dezentral	personen-bezogene Einzeldaten	Papier		Gesetz/Verordnung		Einzelfalldaten liegen zur Weiterverarbeitung vor	Papier, Excel-Datei, Access-Datenbank	
2 a)	Gesetzlich vorgeschriebene Statistik erstellen und veröffentlichen									
2.1 a)	Einzelfalldaten entsprechend Geheimhaltungsvorschriften anonymisieren und zu Produkten aggregieren	innerhalb des AfS	personen-bezogene Einzeldaten	Papier, Excel-Datei, Access-Datenbank		Geheimhaltungsvorschrift		Produkt	PDF, XLS	
2.2 a)	Produkt veröffentlichen	innerhalb des AfS			PDF, XLS			öffentliches Dokument	PDF, XLS	Web-server
2 b)	Berechnung im Datenforschungszentrum									
2.1 b)	Gastwissenschaftler erstellt Datensatz aus Einzelfalldaten	Arbeitsplatz im Datenforschungszentrum	personen-bezogene Einzeldaten		lokaler PC (kontrollierte Software SPSS, SAS, Stata etc., kein Internetzugang)		250 Euro	exklusiv wissenschaftlich verwertbarer Datensatz		mobil, Papier
2 c)	Kontrollierte Datenabfrage									
2.1 c)	Wissenschaftler erstellt Abfrage-skript	dezentral	Dummydaten	XLS				Syntaxfile		
2.2 c)	AfS führt Abfrage durch und setzt Geheimhaltungsvorschriften um	AfS	Syntaxfile, Einzeldaten			Geheimhaltungsvorschrift	250 Euro	exklusiv wissenschaftlich verwertbarer Datensatz	SPSS, SAS, Stata	
2 d)	Bereitstellen von ScientificUseFiles									
2.1 d)	AfS stellt der Wissenschaft Einzeldaten bereit	AfS	personen-bezogene Einzeldaten			Geheimhaltungsvorschrift		faktisch anonymisierte Mikrodaten	SPSS, SAS, Stata	CD/DVD
2 e)	Bereitstellen von PublicUseFiles									
2.1 e)	AfS stellt der Öffentlichkeit Einzeldaten bereit	AfS	personen-bezogene Einzeldaten			Geheimhaltungsvorschrift	verschieden	absolut anonymisierte Mikrodaten	SPSS, SAS, Stata	CD/DVD
2 f)	Bereitstellen von PublicUseFiles									
2.1 f)	AfS stellt den Studenten Einzeldaten bereit	AfS	personen-bezogene Einzeldaten			Geheimhaltungsvorschrift	unentgeltlich	absolut anonymisierte Mikrodaten	SPSS, SAS, Stata	Web

Tabelle 3.2: Groberfassung des Istprozesses beim AfS

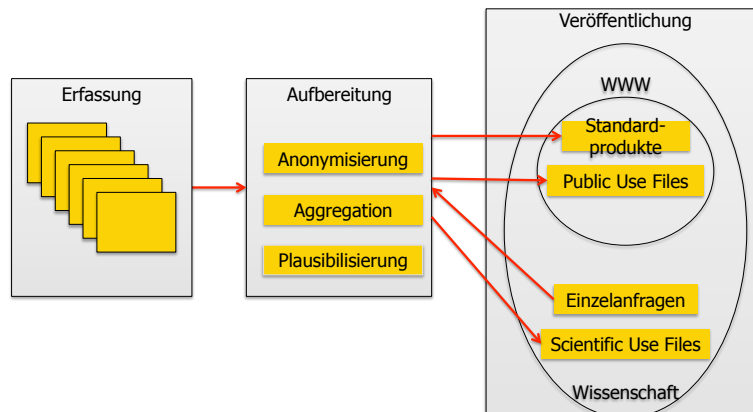


Abbildung 3.3: Schema des Ist-Veröffentlichungsprozesses beim AfS

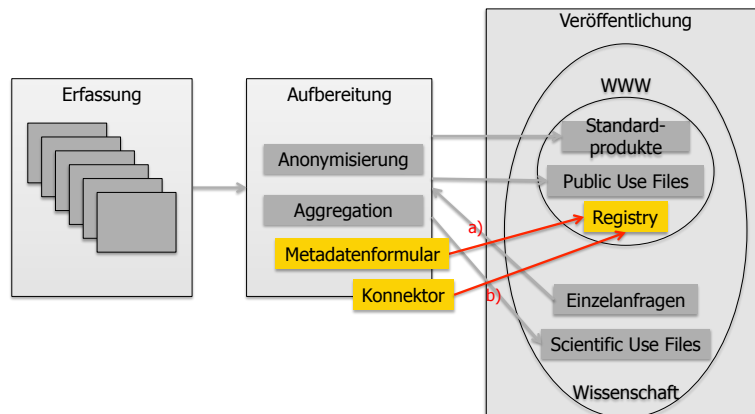


Abbildung 3.4: Schema des Soll-Veröffentlichungsprozesses beim AfS

Abbildung 3.4 skizziert, wie der Veröffentlichungsprozess im Sinne von Open Data umorganisiert werden müsste: Berichte, die als Excel-Datei bereitgestellt werden, müssten zusätzlich in einem zentralen Register gemeldet werden. Das kann durch einen vollautomatischen Konnektor oder durch ein Metadatenformular erfolgen, durch das Metadaten im Register eingetragen werden können.

An solchen Anpassungen ist das AfS durchaus interessiert: Von einem Open Data-Portal verspricht man sich vor allem eine bessere Sichtbarkeit der eigenen Produkte und einen Rückgang der aufwendigen unentgeltlichen Einzelanfragen. Generell scheint das Angebot des Hauses überraschend unbekannt zu sein. Durch ein für Berlin zentrales Portal würden die erzeugten Berichte womöglich besser genutzt.

Bei der umfassenden Veröffentlichung von Rohdaten hat das AfS aber auch Bedenken hinsichtlich der Vertrauenswürdigkeit und der Qualitätssicherung der Daten. Daher ist bei Produkten und Dienstleistungen, die von Daten des AfS abgeleitet sind, besonders wichtig, dass das AfS auch immer als Quelle genannt wird. Idealerweise gäbe es zusätzlich ein Verfahren, mit dem sich die Authentizität von Daten über Verarbeitungsschritte hinweg nachvollziehen lässt.

Des Weiteren äußert das AfS Bedenken hinsichtlich der Unbedenklichkeit von Daten mit Bezug zu Personen: Es ist sehr schwer auszuschließen, dass auch aus einem aggregierten, anonymisierten Datensatz doch noch eine Information über eine Einzelperson erschlossen werden kann, und zwar dann, wenn ein zweiter, bis dahin nicht berücksichtigter Datensatz hinzugezogen wird.

Als weitere Schwierigkeit nennt das Amt den Zielkonflikt zwischen Kostensenkungs- bzw. Einnahmedruck einerseits und der Idee der zweckungebundenen und kostenfreien Überlassung der Daten.

Eine bemerkenswerte Entwicklung beim AfS ist das interne Superstar-Projekt. Hier werden die Kernideen von Open Data schon in Angriff genommen – also Nutzbarmachung von Amtsdaten für Transparenz und wirtschaftliche Zwecke: Das Statistische Informationssystem StatIS BBB ermöglicht es, Auswertungen aus verschiedenen Statistiken in fachlich und regional tiefer Gliederung individuell zusammenzustellen. Es bietet folgende Vorteile:

- Starttabellen für den schnellen Einstieg
- fein untergliederte Daten für die Datenerkundung (Exploration)
- direkte Interaktion mit den Daten
- individuelle Erstellung von aussagekräftigen Tabellen
- jederzeitige Verfügbarkeit der aktuellen Daten
- Export der Datenzusammenstellung in andere Programme (z. B. Microsoft Excel)

Momentan befindet sich das Statistische Informationssystem StatIS BBB im Aufbau. Bereits verfügbar sind folgende Bevölkerungsstatistiken:

- Bevölkerungsfortschreibung
- Wanderungen
- Geburten
- Sterbefälle
- Mikrozensus
- Berliner Einwohnerregisterstatistik

Weitere Fachstatistiken sind in Vorbereitung. Einfache Tabellen können mit nur wenigen Klicks zusammengestellt und an die eigenen Wünsche angepasst werden. Zudem stehen vorgefertigte Tabellen zum Abruf bereit.

Zusammenfassend lässt sich zum AfS festhalten:

- Sehr viel wird schon übersichtlich in maschinenverarbeitbaren Formaten veröffentlicht, allerdings in vielen Fällen nicht unter freien Lizenzen; außerdem sind die Datensätze isoliert von denen anderer Behörden.
- Das AfS ist durchaus an der Veröffentlichung mit semantisch höherwertigen Technologien interessiert.
- Das AfS spielt bei Open Data eine zentrale Rolle durch seine Berufung und seine Kompetenz, sensible Basisdaten in allgemein nützliche Datensätze umzuwandeln.
- Die Herkunft und Authentizität von Daten sollte überprüfbar sein.
- Es besteht ein Zielkonflikt zwischen monetären Zwängen und Open Data.

3.3 Zentral- und Landesbibliothek Berlin

»Die Zentral- und Landesbibliothek (ZLB) ist die größte öffentliche Bibliothek in Deutschland. [...] Es werden über 3,4 Millionen gedruckte und elektronische Medien zur freien Nutzung angeboten. Täglich wird die ZLB von bis zu 5000 Gästen besucht – sie ist eine der meistfrequentierten Kultureinrichtungen Berlins.«⁴

Die Kernaufgabe der ZLB ist es, Information und Informationskompetenz an die Bürger zu vermitteln. Insoweit scheint die Verbindung zu einer Kultur der offenen Daten nahezu liegen. Trotzdem gibt es zurzeit kaum derartige Angebote, und auch für die Zukunft sind hier Schwierigkeiten sichtbar, insbesondere im Zusammenhang mit dem Urheberrecht.

Prinzipiell interessant wären für die zweckungebundene Veröffentlichung vier Datengruppen:

- die bibliografischen Daten
- digitale Medien
- Berliner Rechtsdokumente
- Nutzungsdaten

Letztere werden aus Datenschutzgründen nur da, wo unbedingt nötig, verarbeitet, also nicht im engeren Sinne erhoben.

Alle anderen genannten Datengruppen werden systematisch und in genau definierten Prozessen plausibilisiert, kategorisiert und verschlagwortet (vgl. Abbildung 3.5 und Tabelle 3.3).

Diese Katalogdaten sind jederzeit im Internet abrufbar und abfragbar,⁵ können jedoch aus ganz speziellen Gründen nicht zum Download angeboten werden: An den bibliografischen Daten besteht zum Teil nur ein Nutzungsrecht, da sie von der Nationalbibliothek lediglich lizenziert sind. Die digitalen Medien sind fast ausschließlich urheberrechtlich geschützt. Ähnliches gilt für die Berliner Rechtsliteratur, an der Verlage Urheberrechte halten. Die Protokolle der Sitzungen des Abgeordnetenhauses wären zwar prinzipiell geeignet, enthalten jedoch keine strukturierten Informationen, wie sie für Open Data nützlich wären.

Insoweit muss man zusammenfassend für die ZLB festhalten, dass hier durch eine bloße Umorganisation von Prozessen das durchaus große Potenzial der vorhandenen Informationen nicht mit einem Open Data-Ansatz erschlossen werden kann.

⁴http://www.zlb.de/ueber_uns.

⁵http://www.zlb.de/buecher_medien/onlinekatalog.

Nr.	Arbeitsaufgabe	Stelle	Input			Output	
			Information	Medien	Speicher	Ergebnisse	Medien Speicher
1 a)	Katalogdaten aufnehmen						
1.1 a)	Ggf. vorhandene bibliogra-fische Daten der Deutschen Nationalbibliothek beziehen	innerhalb der ZLB	MAB/MARC				interne DB
1.2 a)	Einträge bibliografisch vervollständigen	innerhalb der ZLB	MAB		interne DB		interne DB
1.3 a)	Einträge kategorisieren und verschlagworten	innerhalb der ZLB	MAB		interne DB		interne DB
1 b)	Medium digitalisieren						
1.1 b)	Scannen/fotografieren	innerhalb der ZLB	analoges Medium (Buch, Zeitschrift, Bild etc.)			digitales Abbild	Medien-server
1.2 b)	Mit bestehendem Eintrag verknüpfen	innerhalb der ZLB			interne DB		Medien-server
1 c)	Kommunale Rechtsliteratur, Plenumsprotokolle des Abgeordnetenhaus						
1.1 c)	Dokument einpflegen	innerhalb der ZLB		PDF			interne DB
1 d)	Nutzungsdaten verarbeiten						
1.1 d)	Nutzungsdaten erfassen	innerhalb der ZLB	personen-bezogene Nutzungsdaten				Nutzer-DB
1.2 d)	Nutzungsdaten löschen	innerhalb der ZLB			Nutzer-DB		
2	Katalogdaten anfragen						
2	Nutzer stellt Kataloganfrage	WWW		einfache oder komplexe Suche		HTML	

Tabelle 3.3: Groberfassung des Istprozesses bei der ZLB

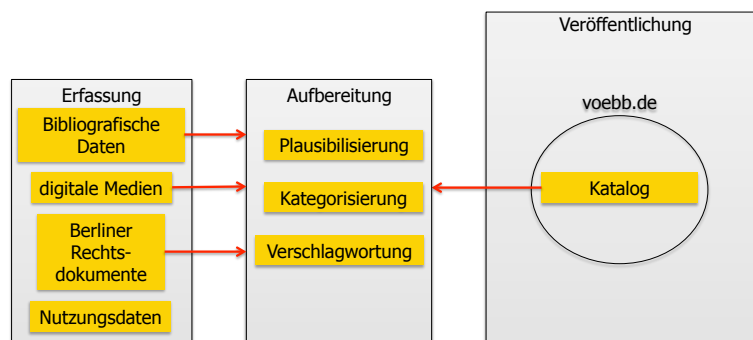


Abbildung 3.5: Schema des Ist-Veröffentlichungsprozesses bei der ZLB

3.4 Web-AG der Umweltämter von Land und Bezirken

Unter dem Namen »Web-AG« arbeitet eine Initiative von Verwaltungsmitarbeitern aus dem Bereich Umwelt, von denen einige beim Land Berlin und andere bei den Bezirksämtern beschäftigt sind. Ziel der Web-AG ist es, alle Berliner Umweltinformationen, insbesondere die gemäß UIG (vgl. Abschnitt 6.1.3), zentral unter berlin.de/umwelt auffindbar zu machen, unabhängig davon, durch welches Amt die Daten erfasst und aufbereitet wurden. Die Erstellung dieses zentralen Angebots ist vom Gesetzgeber nicht vorgeschrieben, dementsprechend ist es auch nicht mit eigenen Mitteln ausgestattet. Vielmehr beruht es auf dem Engagement einzelner Behördenmitarbeiter, die sich von einem koordinierten Webangebot positive Effekte für Bürger und Verwaltung versprechen.

Die Erfahrungen und Einschätzungen der Web-AG, die beim Interview mit ausgewählten Vertretern aufgenommen wurden, sind daher in zweierlei Hinsicht besonders relevant: Erstens werden hier schon seit einiger Zeit erste Schritte in Richtung Open Data im Umweltbereich umgesetzt, nämlich durch ein zentralisiertes Verzeichnis für dezentral verarbeitete Daten, das von Mitarbeitern und nicht von Verordnungen in Gang gesetzt wurde (»bottom-up« versus »top-down«). Zweitens kann hier beobachtet werden, inwieweit eine Koordination über zwei föderale Ebenen möglich ist.

Da an dem Gespräch Vertreter mehrerer Häuser beteiligt waren und die Erfassung und Verarbeitung von Daten in jedem Haus anders abläuft, wäre es unzweckmäßig gewesen, die Veröffentlichungsprozesse so detailliert wie bei den anderen durchgeführten Interviews zu ergründen. Für den eigentlichen Veröffentlichungsschritt ist jedoch festzuhalten, dass es in den teilnehmenden Ämtern jeweils einen designierten Redakteur der Web-AG gibt. Dieser Umstand legt einen Sollprozess nahe, der im Wesentlichen dem entspricht, der schon für die SenBWF vorgeschlagen wurde: Durch eine möglichst kleine Erweiterung des Imperia-Formulars für den Upload von Dateien könnte ein Eintrag des Datensatzes im Datenregister erfolgen. Dieser Prozess ist in Abbildung 3.6 schematisch dargestellt. Die Anpassung würde dazu führen, dass Umweltdaten separat gesucht werden müssen, sondern dort gefunden werden können, wo sich auch allen anderen offenen Datensätzen Berlins finden lassen.

Zusätzlich zu den bereits bereitgestellten Informationen sieht die Web-AG vor allem in zwei Bereichen Potenzial für weitere Veröffentlichungen. Dazu gehören die Daten, die in Berlin tätige Unternehmen den Umweltbehörden ohnehin melden müssen. Insbesondere im Hinblick auf Transparenz bei der Umweltbelastung durch die lokale Wirtschaft gibt es hier starke Argumente für eine Freigabe als Open Data. Des Weiteren existieren viele Informationen, die von Umweltschützern und -aktivisten bzw. von entsprechenden Nicht-regierungsorganisationen gesammelt wurden. Im Gegensatz zu Behördendaten ist hier die Neutralität bisweilen fraglich, aber auch hier kann die Veröffentlichung in einem Open Data-Portal mit entsprechender Kennzeichnung der Bereitsteller der Transparenz nur zuträglich sein.

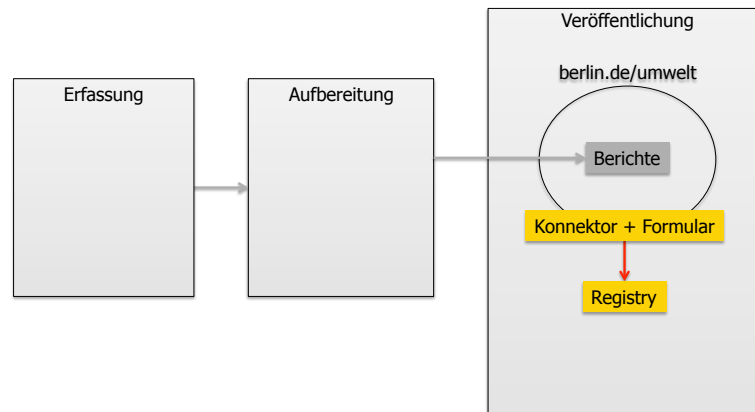


Abbildung 3.6: Schema des Soll-Veröffentlichungsprozesses bei der Web-AG

Auch für die Web-AG stellt sich die Frage der Finanzierung von Open Data-Aktivitäten. Insbesondere wird auf ein zentrales Problem verwiesen: Prinzipiell stellt die Pflege eines weiteren Portals einen Mehraufwand gegenüber der gesetzlichen Pflicht dar. Dieses Mehr wird zwar unter Umständen im Landesdurchschnitt kompensiert, etwa durch steigende Umsatzsteuereinnahmen. Allerdings treten diese Einnahmen vollständig an anderer Stelle auf. Zudem wird befürchtet, dass nach dem Bedienen erster Open Data-Forderungen weitere Ansprüche gestellt werden, denen die Ämter bei den zur Verfügung stehenden knappen Ressourcen nicht mehr gerecht werden können. Ebenfalls im Hinblick auf die Finanzlage soll vermieden werden, dass mehrere bzw. häufig wechselnde Systeme gepflegt werden müssen; insoweit ist ein berlin- und ressortzentrales Register zwar erwünscht, aber nur wenn es Aussicht auf Langlebigkeit hat.

Weiterhin ist zu bedenken, dass die Erfassung und Verarbeitung von Umweltdaten in den einzelnen Bezirken in unterschiedlicher Art und Weise und auch in unterschiedlichem Umfang erfolgt. Zum einen gilt es, diese Heterogenität zu berücksichtigen. Zweitens muss man sich bewusst sein, dass die einzelnen Ämter in eine Konkurrenzsituation gebracht werden, sobald transparent und nachvollziehbar wird, welcher Bezirk welche Umweltdaten erfasst und veröffentlicht.

Gerade die Tatsache, dass die Web-AG von sich aus erste Schritte in Richtung Open Data gegangen ist, zeigt, dass die teilnehmenden Mitarbeiter bei allen Bedenken große Chancen in dieser Bewegung sehen:

Der Wechsel weg von der Bedienung von Einzelanfragen hin zu einer »Selbstbedienung« durch Bürger, Datenjournalisten und Unternehmen geht einher mit einer großen Entlastung für die Ämter und kann damit bei der Kostensenkung helfen.

Gleichzeitig könnte es den Umweltbehörden zu mehr Geltung verhelfen bzw. ihre Leistungen stärker ins Blickfeld der öffentlichen Beachtung rücken. Viele hochwertige und

bereits heute veröffentlichte Informationen werden von der Öffentlichkeit unter Umständen stark unterbewertet, da die Verwaltung nur begrenzt für eine ansprechende Umsetzung in Visualisierungen und Anwendungen sorgen kann. Das Engagement Dritter könnte hier die Verdienste der Verwaltung besser zur Geltung bringen.

Des Weiteren verspricht sich die Web-AG von einem Open Data-Portal einen Kommunikations-Rückkanal vom Bürger zur Verwaltung. Heute sind die Hürden für Bürger verhältnismäßig hoch, Wünsche, Fragen oder auch Lob so an die Behörden zu tragen, dass einerseits der richtige Mitarbeiter erreicht wird, andererseits transparent verfolgt werden kann, ob und wie die Eingaben der Bürger berücksichtigt worden sind.

Abschließend ist zu erwähnen, dass der von Open Data intendierte Transparenzeffekt zu einem großen Teil mit dem Anliegen der Umweltämter übereinstimmt, nämlich dafür zu sorgen, dass Umweltinformationen, insbesondere aus dem wirtschaftsnahen Umfeld, frei zur Recherche zur Verfügung stehen.

3.5 Fazit

In den durchgeführten Interviews hat sich gezeigt, dass in den Berliner Verwaltungseinrichtungen ein großes Potenzial für Open Data liegt: Ein großer und weiter wachsender Bestand an Daten liegt bereits elektronisch vor und ließe sich verhältnismäßig leicht in einem maschinenverarbeitbaren Format exportieren. Bereits kleine Anpassungen bestehender Veröffentlichungsprozesse ermöglichen hier einen schnellen Einstieg in Open Data. Die Institutionen haben zudem häufig ein Eigeninteresse an der Bereitstellung offener Daten bekundet.

Diese Veränderungen können durch verschiedene organisatorische Maßnahmen begünstigt werden. Dazu gehört die weitergehende Analyse und Beratung jeder individuellen Institution. Die im Rahmen der Studie durchgeführten Interviews haben gezeigt, dass die Abläufe sich in den Häusern unterscheiden, aber häufig durch eine genaue Beobachtung ein aufwandsarmer Einstieg gefunden werden kann. Idealerweise werden dabei jeweils so viel Wissen und so viele Fertigkeiten transferiert, dass im Anschluss jede Einrichtung über Personal mit ausreichend Fachwissen im Hinblick auf offene Daten verfügt. Zusätzlich wird jedoch eine zentrale Anlaufstelle benötigt, um gewonnene Erkenntnisse weiterzugeben und, wo nötig, institutionsübergreifende Lösungen zu finden.

Die Maßnahmen können jedoch nur Erfolg bringen, wenn es ein klares politisches Bekenntnis zu Open Data gibt. Häufig besteht ein Zielkonflikt zwischen der Freigabe von Daten und dem Datenschutz einerseits und Einnahmezwängen sowie dem tradierten Verwaltungsverständnis andererseits. Diesen Konflikt können Verwaltungseinrichtungen nur mit entsprechender Rückendeckung lösen.

4

Architektur

In diesem Abschnitt werden Architekturempfehlungen und eine technische Konzeption für eine Berliner Open Data-Plattform beschrieben. Die vorgestellte Grobspezifikation kann als Basis für Umsetzungsprojekte genutzt werden.

Die Herleitung der technischen Konzeption der Open Data-Plattform für Berlin wurde auf Basis der Vorstudie zur City Data Cloud Berlin (Barnickel et al., 2010) und unter Berücksichtigung von spezifischen Anforderungen der identifizierten Bereitsteller und Nutzer der Daten durchgeführt. Zudem berücksichtigen die Architekturempfehlungen internationale Standards für die Bereitstellung und Verlinkung öffentlicher Daten und zur Sicherstellung der Interoperabilität zu anderen Open Data-Plattformen.

Im Folgenden werden die Ergebnisse beschrieben, die im Rahmen von internen Architekturworkshops, Expertengesprächen und Literaturrecherchen unter Berücksichtigung von internationalen Good Practice-Ansätzen sowie eigener Expertise und eigener Erfahrungen, u. a. aus den Projekten Open Cities¹, Datenpool Berlin und Engage², erarbeitet wurden. Nach dem Ansatz einer prozessorientierten Anforderungsanalyse wird zunächst ein generischer Sollprozess für die Bereitstellung von öffentlichen Daten beschrieben. Dieser beinhaltet die grundlegenden Rollen und beschreibt das Aufgabenspektrum, das diesen zugeordnet werden kann. Auf Grundlage des Sollprozesses werden anschließend konkrete Anwendungsfälle und damit verbundene notwendige Funktionalitäten der Berliner Open Data-Plattform abgeleitet. Davon ausgehend wird eine Grobarchitektur entworfen, die eine mögliche Realisierung der Berliner Open Data-Plattform nach Domänen, Akteuren und Komponenten sowie deren Schnittstellen strukturiert.

¹<http://opencities.net>.

²<http://www.engage.project.eu>.

4.1 Generischer Soll-Prozess mit Abbildung auf Systemkomponenten

Im Rahmen der Organisationanalyse wurden bereits Sollprozesse zur Bereitstellung von öffentlichen Daten in Berlin entworfen. Die prozessorientierte Anforderungsanalyse für die Berliner Open Data-Plattform abstrahiert von diesen spezifischen Sollprozessen und leitet davon ausgehend einen generischen Sollprozess her. Abbildung 4.1 zeigt die wesentlichen Schritte bei der Bereitstellung von öffentlichen Daten und identifiziert die darin beteiligten Akteure, Domänen und notwendigen Kernkomponenten der Open Data-Plattform.

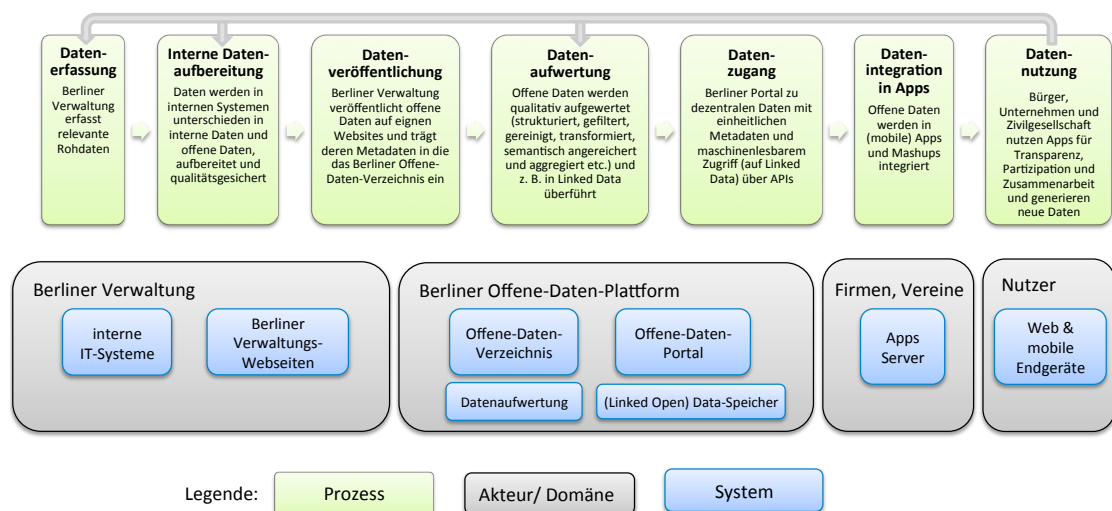


Abbildung 4.1: Generischer Sollprozess mit Abbildung auf die Plattformkomponenten

Der erste Schritt beschreibt die initiale Datenerfassung. Die Berliner Verwaltung erfasst relevante Rohdaten in verschiedenen Formaten auf Basis von aufgabenspezifischen Verfahrensweisen, die in diesem Rahmen nicht weiter betrachtet werden.

Der zweite Schritt befasst sich mit der internen Datenaufbereitung, um aufzuzeigen, dass die Quellen von offenen Daten, wie sie im Rahmen dieser Studie betrachtet werden, zunächst interne IT-Systeme der Berliner Verwaltung umfassen. Verwaltungsmitarbeiter führen auf deren Basis eine Qualitätssicherung der Daten durch und nutzen diese für die spezifischen Verwaltungsaufgaben. Rohdaten, die beispielsweise personenbezogene Daten enthalten, werden zu statistischen Daten anonymisiert und aggregiert. Schließlich kann unterschieden werden, welche Daten weiterhin nur intern weiterverwendet werden und welche als offene Daten veröffentlicht werden können.

Der folgende Schritt beschreibt die Veröffentlichung der Daten. Die Berliner Verwaltung veröffentlicht ausgewählte offene Daten auf eigenen, dezentral verwalteten Websites, wie

z. B. auf den entsprechenden Bereichen von www.berlin.de. Um einen einheitlichen Zugang mit übergreifender Suche und Navigation zu den verschiedenen Datenquellen zu ermöglichen, werden zusätzliche Beschreibungen der Datensätze, sogenannte Metadaten, zentral in einem stadtweiten Register für offene Daten registriert.

Darauf aufbauend kann optional eine Datenaufwertung durchgeführt werden. Dieser Schritt dient der Erhöhung der Nachnutzbarkeit und somit der Steigerung des Potenzials von offenen Daten, mehrwertbringend in webbasierten Mashups und mobilen Applikationen weiterverwendet zu werden. Um die dafür notwendige Maschinenlesbarkeit und Interpretierbarkeit sicherzustellen, können Datensätze in semantisch ausdrucksstärkere Formate wie z. B. Linked Data überführt werden, die es ermöglichen, den Kontext und die Beziehungen zwischen Daten explizit und interoperabel auszudrücken. Die Verwendung von Linked Open Data und zugehörigen Technologien vereinfacht es, aufwendige Filterungs- und Aggregationsmechanismen durchzuführen und Datensätze, die in verschiedenen Vokabularen beschrieben sind, zusammenzuführen, um daraus neue Schlussfolgerungen abzuleiten. Datensätze im Linked Data-Format können zentral auf der Open Data-Plattform gespeichert werden, die notwendige Werkzeuge und Dienste für die Überführung und persistente Bereitstellung zur Verfügung stellt.

Der nächste Schritt beschreibt den Datenzugang über ein Webportal, das einen einheitlichen Zugang zu den offenen Daten liefert (basierend auf standardisierten Metadaten). Darüber hinaus wird der maschinenlesbare Zugang über ein Application Programming Interface (API) für die Nachnutzung verdeutlicht.

Darauf aufbauend kann die Datenintegration in Apps erfolgen. Entwickler aus dem privaten Sektor und der Zivilgesellschaft entwickeln (mobile) Anwendungen und vertreiben sie über externe Infrastrukturen wie Applikationsserver und App Stores, die nicht Teil der Berliner Open Data-Plattform sind.

Abschließend findet die eigentliche Nutzung der Daten statt. Zum einen nutzen Bürger, Wirtschaft und Zivilgesellschaft die webbasierten und mobilen Anwendungen für mehr Transparenz, Partizipation und Zusammenarbeit in Berlin. Zum anderen können so perspektivisch auch neue Daten und Verknüpfungen zwischen Daten generiert werden, die dann wiederum bei der Datenerfassung, der internen Datenaufbereitung oder der Datenaufwertung berücksichtigt werden und somit zur weiteren Bereitstellung von offenen Daten führen.

Die hier dargestellte prozessorientierte Perspektive mit der ersten Identifikation von Akteuren, Domänen und Kernkomponenten ermöglicht ein generelles Verständnis der zu konzipierenden Open Data-Plattform und einen Überblick über ihre Bestandteile. Im folgenden Abschnitt kann die Anforderungsanalyse darauf aufbauend anhand von konkreten Anwendungsfällen und damit verbundenen Funktionalitäten vertieft werden.

4.2 Anwendungsfälle

Mithilfe der im Folgenden dargestellten Anwendungsfälle werden die Funktionalitäten der Open Data-Plattform beschrieben. Sie helfen in der Planungs- und Umsetzungsphase, ein einheitliches Verständnis zwischen den beteiligten Akteuren auf Basis einer strukturierten Darstellung in Form von Anwendungsfalldiagrammen herzustellen.

Anwendungsfälle beschreiben mögliche Szenarien, die eintreten können, wenn ein Akteur versucht, mithilfe des betrachteten Systems ein bestimmtes fachliches Ziel zu erreichen. Sie beschreiben, was inhaltlich beim Versuch der Zielerreichung geschehen kann, und abstrahieren von konkreten technischen Lösungen.³.

Ausgehend von der obigen Prozessanalyse können zunächst folgende Akteure identifiziert werden:

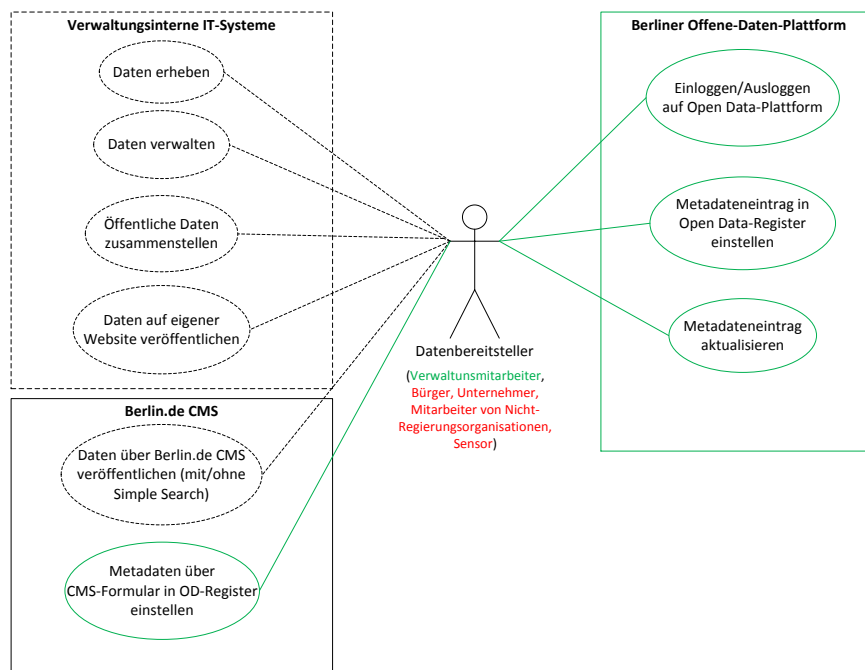
- **Datenbereitsteller:** Datenbereitsteller sind in erster Linie Mitarbeiter der Berliner Verwaltung. Sie bilden den zentralen Akteur in den Prozessschritten von der Datenerfassung bis zur Datenveröffentlichung. In einer weiteren Ausbaustufe ist ebenfalls denkbar, dass Bürger, Unternehmen oder Organisationen aus der Zivilgesellschaft auch als Datenbereitsteller auftreten. Darüber hinaus können auch Sensordatenquellen als Datenbereitsteller aufgefasst werden, die automatisiert Daten an die Open Data-Plattform übermitteln.
- **Plattformbetreiber:** Der Plattformbetreiber ist für die kontinuierliche Bereitstellung und Wartung der Open Data-Plattform verantwortlich.
- **Direkter Datennutzer:** Der direkte Datennutzer agiert unmittelbar mit den über die Open Data-Plattform bereitgestellten Daten. Diese Akteursgruppe umfasst damit z. B. Applikationsentwickler, Datenjournalisten, Wissenschaftler, Bürger, Unternehmen und Organisationen aus der Zivilgesellschaft sowie Mitarbeiter aus der Verwaltung im Rahmen von verwaltungsübergreifender Zusammenarbeit.
- **Indirekter Datennutzer:** Der indirekte Datennutzer agiert mittelbar mit den bereitgestellten offenen Daten durch die Nutzung von (mobilen) Applikation oder Web-Mashups, die auf den über die Open Data-Plattform veröffentlichten Daten aufbauen. Diese Akteursgruppe umfasst die Endanwender wie Bürger, Unternehmen oder Organisationen aus der Zivilgesellschaft sowie Mitarbeiter aus der Verwaltung im Rahmen von verwaltungsübergreifender Zusammenarbeit.
- **Datenaufbereiter:** Der Datenaufbereiter bezieht sich auf den optionalen Prozessschritt Datenaufwertung. Datenaufbereiter können sowohl Mitarbeiter für Qualitätssicherung aus der Berliner Verwaltung, als auch sogenannte Datenpaten aus der Zivilgesellschaft oder externe Dienstleister sein.

³<http://de.wikipedia.org/wiki/Anwendungsfall>.

Im Folgenden werden die funktionalen Anforderungen an die Open Data-Plattform aus Sicht der aufgeführten Akteure beschrieben und in Form von Anwendungsfällen dargestellt. Anschließend werden die einzelnen Anwendungsfälle näher erläutert und die daraus resultierenden Funktionen beschrieben. Die grün markierten Anwendungsfälle beziehen sich auf den Funktionsumfang, der für eine erste Ausbaustufe geplant ist, die rot markierten entsprechend auf eine zweite Ausbaustufe.

4.2.1 Akteur: Datenbereitsteller

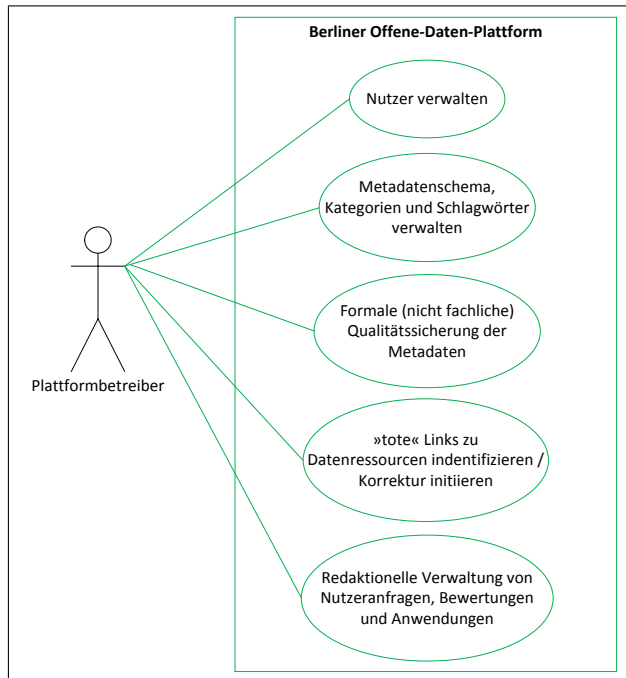
Die folgende Abbildung gibt eine Übersicht zu den Anwendungsfällen des Akteurs Datenbereitsteller. Es werden drei Systemgrenzen identifiziert: verwaltungsinterne IT-Systeme, das Berlin.de Content Management System (CMS), zuständig für die Webseiten der Berliner Verwaltung und die Berliner Offene-Daten-Plattform. Die Anwendungsfälle, die den verwaltungsinternen IT-Systemen zugeordnet sind, sind zur Vollständigkeit dargestellt. Sie werden in diesem Rahmen jedoch nicht weiter betrachtet, da keine direkten Schnittstellen zur Offenen-Daten-Plattform geplant sind.



Nr.	Anwendungsfall	Kurzbeschreibung
1	Daten über das Berlin.de CMS veröffentlichen (mit/ohne Simple Search)	Bereits verfügbare Funktionalität des CMS Imperia, das für die Redaktion der Berliner Verwaltungswebseiten auf berlin.de verwendet wird. Simple Search beschreibt eine dynamische Datenbereitstellungsfunktion, die es ermöglicht, in Tabellen strukturierte Datensätze auf Berlin.de über eine Schnittstelle dynamisch (d. h. spezifisch parametrisierbar über RESTful-Webservices mit JSON) abzufragen.
2	Metadaten über CMS-Formular in Offene-Daten-Register einstellen	Durch ein in das CMS Imperia integriertes Formular kann der Onlineredakteur auf Berlin.de Metadaten zu einem von ihm neu bereitgestellten Datensatz direkt über eine Schnittstelle in das Open Data-Register einstellen.
3	Ein- und Ausloggen auf der Open Data-Plattform	Um sicherzustellen, dass nur berechtigte Nutzer Metadaten und Datensätze in die Open Data-Plattform einstellen, müssen sie sich über ein Berechtigungssystem ein- und ausloggen und zuvor dafür registriert werden.
4	Metadateneintrag in Offene-Daten-Register einstellen	Über ein Webformular können Datenbereitsteller Metadaten zu von ihnen veröffentlichten Datensätzen in das Offene-Daten-Register einstellen. Siehe hierzu auch die Anlage F: Handreichung zum Datenregister.
5	Metadateneintrag aktualisieren	Der Datenbereitsteller kann von ihm erstellte Metadateneinträge im Open Data-Register aktualisieren.

Tabelle 4.1: Anwendungsfälle für den Akteur Datenbereitsteller

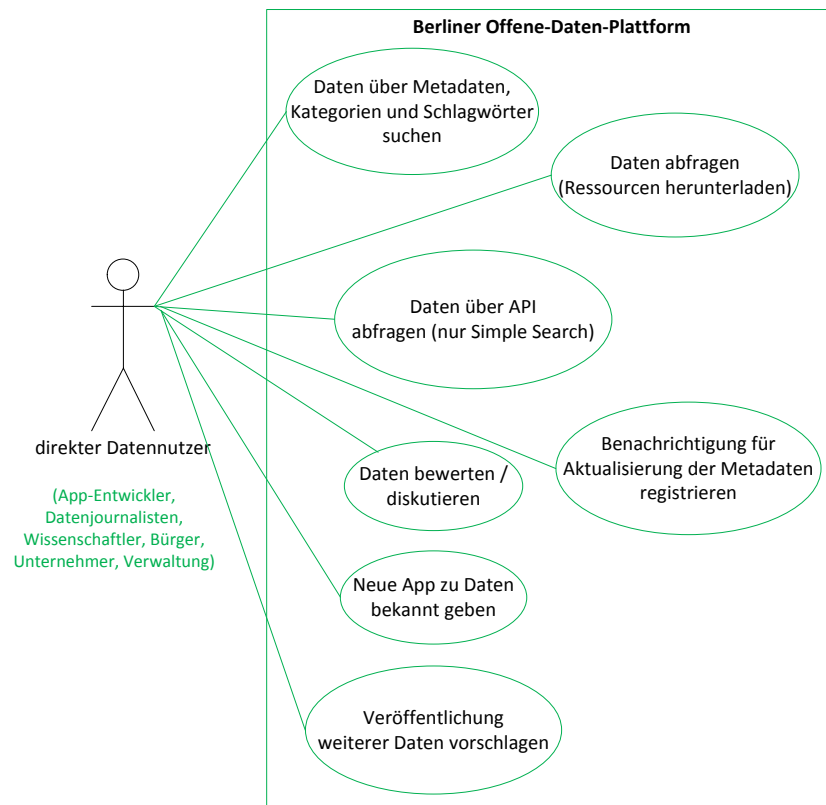
4.2.2 Akteur: Plattformbetreiber



Nr.	Anwendungsfall	Kurzbeschreibung
6	Nutzer verwalten	Der Plattformbetreiber kann Nutzerkonten mit CRUD-Funktionen (Create, Read, Update, Delete) verwalten.
7	Metadatenschema, Kategorien und Schlagwörter verwalten	Der Plattformbetreiber kann Metadatenschema, Kategorien und Schlagwörter mit CRUD-Funktionen verwalten.
8	Formale (nicht fachliche) Qualitätssicherung der Metadaten	Die Metadaten werden automatisch auf Vollständigkeit bezüglich der Pflichtfelder und der korrekten Datentypen überprüft.
9	»Tote« Links zu Datenressourcen indentifizieren / Korrektur initiieren	»Tote« Links zu Datenressourcen werden automatisch identifiziert und es wird eine (gesammelte) Benachrichtigungs-E-Mail versandt.
10	Redaktionelle Verwaltung von Nutzeranfragen, Bewertungen und Anwendungen	Eingehende Nutzeranfragen, Bewertungen und verlinkte und beschriebene Anwendungen müssen überprüft und ihre Beantwortung bzw. ihre Übernahme muss initiiert werden.

Tabelle 4.2: Anwendungsfälle für den Akteur Plattformbetreiber

4.2.3 Akteur: direkter Datennutzer

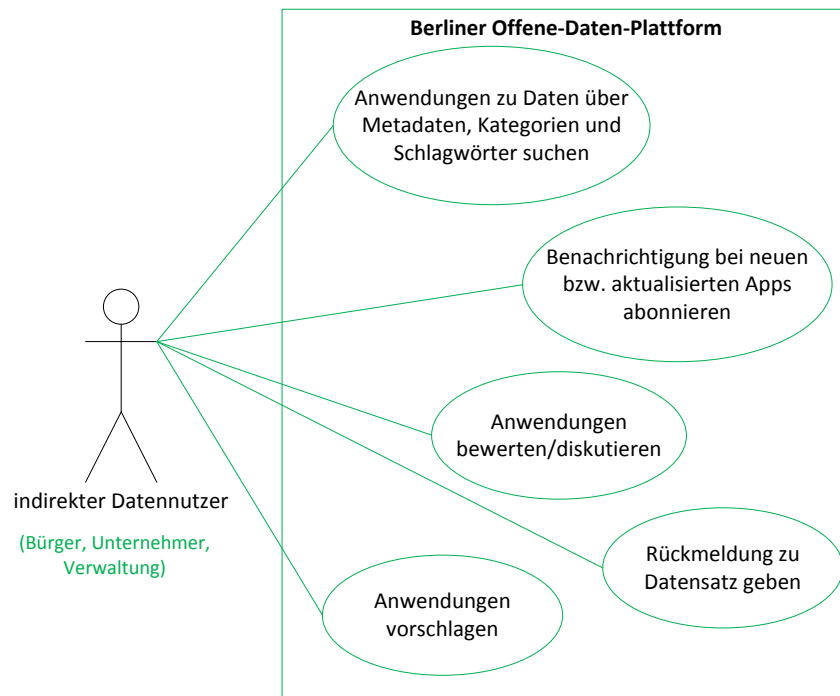


Nr.	Anwendungsfall	Kurzbeschreibung
11	Daten über Metadaten, Kategorien und Schlagwörter suchen	Direkte Datennutzer können Datensätze anhand der Metadaten, Kategorien und Schlagwörter und einer Volltextsuche suchen. Die Volltextsuche bezieht auch Portalinhalte mit ein, jedoch keine Inhalte der extern bereitgestellten Datensätze.
12	Daten abfragen (Ressourcen herunterladen)	Direkte Datennutzer können die Datensätze über den Uniform Resource Locator (URL) im zugehörigen Metadatensatz herunterladen

Nr.	Anwendungsfall	Kurzbeschreibung
13	Daten über API abfragen (nur Simple Search)	Ausgewählte Datensätze, die von Berlin.de-Onlineredakteuren mit den SimpleSearch-Baukasten bereitgestellt wurden (http://support.berlin.de/wiki/index.php/SimpleSearch-Baukasten), können über eine RESTful-Webservice-API (http://support.berlin.de/wiki/index.php/SimpleSearch_API) abgefragt werden. Die URL zur API wird analog der URL zur Datensatzressource im zugehörigen Metadatensatz gepflegt.
14	Benachrichtigung für Aktualisierung der Metadaten registrieren	Um automatisch über neu eingestellte Datensätze benachrichtigt zu werden, können RSS-Feeds für einzelne Datensätze, Kategorien, Schlagwörter und das gesamte Portal abonniert werden.
15	Daten bewerten/diskutieren	Zu jedem Datensatz soll eine 5-Sterne-Bewertungsauswahl angeboten werden. Die Bewertung kann sich z. B. auf Qualität und Relevanz des Datensatzes beziehen. Zu jedem Datensatz soll die Möglichkeit einer Onlinediskussion gegeben werden. Beiträge können nur nach Log-in geschrieben werden, lesbar sollen sie jedoch auch für nicht angemeldete Nutzer sein.
16	Neue App zu Daten bekannt geben	Es soll ein Onlineformular bereitgestellt werden, mit dem App-Entwickler Informationen (Kurzbeschreibung, Screenshots, URL und Datenbezug zu welchem etc.) bereitstellen können. Nach Freigabe durch den Plattformbetreiber (gegebenenfalls delegiert an einen Redakteur) sollen die Informationen im Anwendungenbereich (Apps) des Open Data-Portals sichtbar sein. Anwendungen sollten auf die Datensätze im Portal verweisen, die sie verwenden, und Datensätze ebenso auf Apps, in denen sie verwendet werden.
17	Veröffentlichung weiterer Daten vorschlagen	Über ein separates Forum, das z. B. nach den Datenkategorien strukturiert ist, sollen Vorschläge für die Veröffentlichung weiterer Datensätze eingereicht werden können.

Tabelle 4.3: Anwendungsfälle für den Akteur direkter Datennutzer

4.2.4 Akteur: indirekter Datennutzer

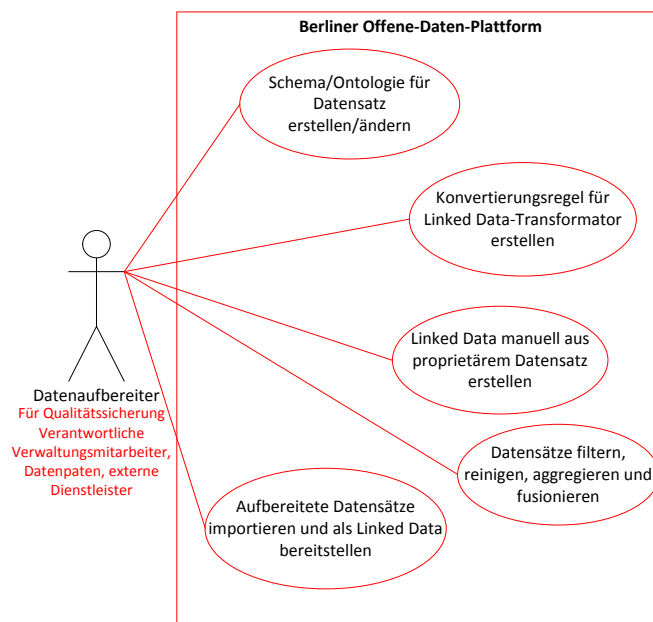


Nr.	Anwendungsfall	Kurzbeschreibung
18	Anwendungen zu Daten über Metadaten, Kategorien und Schlagwörter suchen	Extern bereit gestellte Anwendungen, die nach Apps-Bereichen strukturiert und mit einer Kurzbeschreibung vorgestellt werden, können über Metadatenfelder, Kategorien oder Schlagwörter gesucht werden.
19	Benachrichtigung bei neuen bzw. aktualisierten Apps abonnieren	Über RSS-Feeds können Benachrichtigungen abonniert werden, die über neue und aktualisierte Apps informieren.
20	Anwendungen bewerten/ diskutieren	Analog zu den Datensätzen können auch Anwendungen z. B. über ein 5-Sterne-System bewertet werden. Ebenfalls wird ein Diskussionsbereich zur Verfügung gestellt.

Nr.	Anwendungsfall	Kurzbeschreibung
21	Rückmeldung zu Datensatz geben	Im Rahmen des Diskussionsbereichs zu den Anwendungen soll auch die Möglichkeit gegeben werden, Rückmeldungen zu den verwendeten Datensätzen zu geben. Die Meldungen sollen dann an den Ansprechpartner des entsprechenden Datensatzes weitergeleitet werden.
22	Anwendungen vorschlagen	Über einen interaktiven Forumsbereich soll die Möglichkeit gegeben werden, Anwendungen vorzuschlagen.

Tabelle 4.4: Anwendungsfälle für den Akteur indirekter Datennutzer

4.2.5 Akteur: Datenaufbereiter



Nr.	Anwendungsfall	Kurzbeschreibung
23	Schema/Ontologie für Datensatz erstellen/ändern	Der Datenaufbereiter erstellt ein Schema oder eine Ontologie (vgl. Glossar) für einen oder mehrere Datensätze. Sie werden über die Plattform bereitgestellt und können geändert und gelöscht werden.

Nr.	Anwendungsfall	Kurzbeschreibung
24	Konvertierungsregel für Linked Data-Transformator erstellen	Der Datenaufbereiter erstellt Konvertierungsregeln für ein externes Transformatorwerkzeug, um vorhandene Datensätze z. B. im CSV- oder Excel-Format in Linked Data umzuwandeln.
25	Linked Data manuell aus proprietärem Datensatz erstellen	Auf Basis der Konvertierungsregeln kann mithilfe eines externen Transformatorwerkzeugs aus einem proprietären Datensatz (z. B. im Excel-Format) ein Linked Data-Datensatz erstellt werden.
26	Datensätze filtern, reinigen, aggregieren und fusionieren	Mithilfe externer Werkzeuge können Datensätze gefiltert, gereinigt, aggregiert und fusioniert werden.
27	Aufbereitete Datensätze importieren und als Linked Data bereitstellen	Über externe Werkzeuge aufbereitete Datensätze können in die Plattform importiert werden und als Linked Data über API und entsprechende Abfragesprachen (z. B. SPARQL) bereitgestellt werden.

Tabelle 4.5: Anwendungsfälle für den Akteur Datenaufbereiter

4.3 Ausbaustufen der Architektur

In diesem Abschnitt wird die Architektur der Berliner Open Data-Plattform in zwei Ausbaustufen beschrieben. Die Architektur gliedert sich in Komponenten, mit denen die im vorigen Abschnitt beschriebenen Anwendungsfälle umgesetzt werden können. In der Architektur werden auch Domänen an den Systemgrenzen der Berliner Open Data-Plattform beschrieben, um deren Interaktion mit der Berliner Open Data-Plattform zu verdeutlichen. Die Domänen an den Systemgrenzen umfassen Behörden, direkte Datennutzer, indirekte Datennutzer und externe Betreiber von App Stores.

Die Architektur wird analog zur Aufteilung der Anwendungsfälle in zwei Ausbaustufen dargestellt. In der ersten Ausbaustufe wird ein zentraler Zugang über einheitliche Meta-daten zu dezentral gepflegten und gehaltenen Daten beschrieben (grüne Komponenten und Schnittstellen). In der zweiten Ausbaustufe wird zusätzlich die Möglichkeit gegeben, ausgewählte Daten semantisch aufzubereiten und als Linked Data in der Berliner Open Data-Plattform zu speichern und über diese maschinenlesbar bereitzustellen.

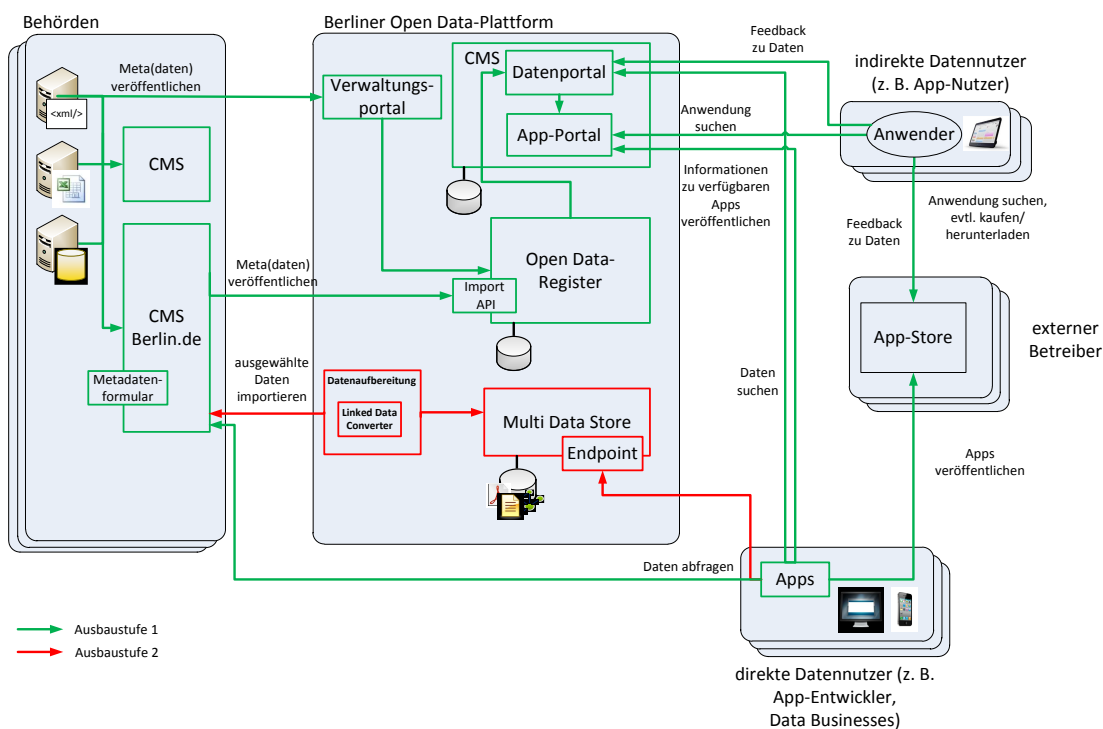


Abbildung 4.2: Grobarchitektur und Schnittstellen

Die folgenden Abschnitte beschreiben die fünf Domänen mit ihren wesentlichen Komponenten und den Schnittstellen zwischen ihnen.

4.3.1 Domäne Behörden

Datensätze der Behörden in verschiedenen Formaten – z. B. Hypertext Markup Language (HTML), PDF, Comma Separated Value (CSV), Extensible Markup Language (XML), Excel etc. – werden dezentral im CMS gepflegt und gespeichert. Viele Berliner Behörden nutzen das auf Imperia basierende CMS zur Veröffentlichung von Webinhalten auf Berlin.de. Andere Behörden, wie z. B. das Statistikamt, nutzen ein eigenes CMS und veröffentlichen ihre Inhalte auch auf eigenen Webseiten. Um die Auffindbarkeit der Datensätze über das zentrale Berliner Open Data-Portal zu erhöhen, werden einheitliche Metadaten dieser Datensätze über das Verwaltungsportal der Berliner Open Data-Plattform veröffentlicht.

Zusätzlich besteht die Möglichkeit, im Redaktionsprozess des CMS von Berlin.de bereits direkt über ein Metadatenformular die Metadaten zu neuen Datensätzen zu erfassen. Diese werden dann automatisch über eine Webservice-Schnittstelle (Import-API) an das Open Data-Register der Berliner Open Data-Plattform übermittelt.

4.3.2 Berliner Open Data-Plattform

Die Berliner Open Data-Plattform besteht aus fünf wesentlichen konzeptionellen Komponenten, die im Folgenden beschrieben werden.

4.3.2.1 Content Management-System (CMS)

Auf Basis des CMS der Berliner Open Data-Plattform werden das Berliner Datenportal und das zugehörige Apps-Portal (Apps-Bereich innerhalb des Datenportals) bereitgestellt. Das Datenportal bietet eine öffentliche Weboberfläche zum Navigieren, Suchen, Diskutieren und zum Zugriff auf offene Daten der Berliner Verwaltung bzw. der Stadt. Im Apps-Portal werden Informationen und Verweise zu Anwendungen gegeben, die Dritte auf Basis der Daten entwickelt haben. Ein CMS wird als Grundlage verwendet, um ein ansprechendes Webportal zu gestalten und Funktionen wie Blogs, Foren, News und Feeds einfach einbinden zu können.

4.3.2.2 Verwaltungsportal

Das Verwaltungsportal bietet den Berliner Verwaltungsmitarbeitern oder autorisierten Dritten eine nichtöffentliche Weboberfläche zur Veröffentlichung von Metadaten auf der Berliner Open Data-Plattform. Über ein Formular eingegebene Metadaten zu dezentral bereitgestellten Datensätzen werden direkt im Open Data-Register gespeichert und dann anschließend im Open Data-Portal der Öffentlichkeit zugänglich gemacht.

4.3.2.3 Open Data-Register

Das Open Data-Register speichert Metadaten zu den Datensätzen, die über das Datenportal auffindbar und abrufbar sind. Es werden zwei Möglichkeiten zum Eintragen von Metadaten in das Open Data-Register angeboten:

1. Verwendung des Verwaltungsportals und manuelle Eingabe der Metadaten über ein Webformular
2. Verwendung der Import-API des Open Data-Registers mit direktem Aufruf aus den CMS der Verwaltungen und somit die automatische Publikation der Metadaten

4.3.2.4 Multi/Linked Data Store

Neben der Bereitstellung von Metadaten zu Datensätzen, die dezentral gespeichert und bereitgestellt werden, soll die Berliner Open Data-Plattform in einer zweiten Ausbaustufe auch die Möglichkeit bieten, angereicherte und überführte ausgewählte Datensätze als Linked Data zentral zu speichern und bereitzustellen. Der Linked Data-Ansatz ermöglicht die Verknüpfung und Aggregation von Datensätzen aus unterschiedlichen und heterogenen Datenquellen und bietet umfangreiche Filter- und Abfragemechanismen. Die Überführung zu Linked Data wird durch entsprechende Werkzeuge zur Datenaufbereitung unterstützt, die Datenbereitstellern bzw. Datenpaten zur Verfügung gestellt werden. Linked Data werden im Resource Description Framework (RDF)-Format überführt und im Linked Data Store der Berliner Open Data-Plattform gespeichert. Maschinenlesbarer und maschineninterpretierbarer Zugang für extern entwickelte (mobile) Applikationen und Mashups wird durch Webschnittstellen, die auf einer Abfragesprache für RDF basieren, bereitgestellt – Simple Protocol and RDF Query Language (SPARQL). Neben dem Angebot von Linked Data soll darüber hinaus auch die Möglichkeit gegeben werden, Datensätze in anderen Datenformaten, z. B. CSV, Excel Spreadsheet (XLS) und XML, in einem zentralen Speicher der Open Data-Plattform zu speichern, sodass der Linked Data Store als Multi Data Store umgesetzt werden sollte.

4.3.3 Datennutzer und externe App Stores

- Direkte Datennutzer: Der direkte Datennutzer agiert unmittelbar mit dem über die Open Data-Plattform bereitgestellten Daten. Diese Akteursgruppe umfasst damit z. B. Applikationsentwickler, Datenjournalisten, Wissenschaftler, Bürger, Unternehmen und Organisationen aus der Zivilgesellschaft sowie Mitarbeiter aus der Verwaltung im Rahmen von verwaltungsübergreifender Zusammenarbeit. Sie suchen Datensätze auf dem Datenportal, fragen diese über Webschnittstellen ab und geben Rückmeldungen über die Qualität der Daten wie z. B. deren Weiterverarbeitbarkeit.

- Indirekte Datennutzer: Der indirekte Datennutzer agiert auch mit den bereitgestellten offenen Daten, aber mittelbar, d. h. durch die Nutzung von (mobilen) Applikationen oder Web-Mashups, die auf den über die Open Data-Plattform veröffentlichten Daten aufbauen. Diese Akteursgruppe umfasst die Endanwender wie Bürger, Unternehmen oder Organisationen aus der Zivilgesellschaft sowie Mitarbeiter aus der Verwaltung im Rahmen von verwaltungsübergreifender Zusammenarbeit. Die Apps werden zum einen im Appsbereich des Datenportals gesucht und gefunden oder auch über herkömmliche (mobile) App Stores wie z. B. von Apple oder Google bezogen.

4.3.4 Umsetzung der Architektur

Die vorgestellte Grobspezifikation der Open Data-Plattform kann als Basis für Umsetzungsprojekte genutzt werden. Die in diesem Kapitel ausgeführten Prozess- und Anwendungsfallbeschreibungen bilden eine gute Grundlage für die Lastenhefterstellung z. B. im Rahmen von Ausschreibungs- oder Vergabeverfahren. Hervorzuheben ist, dass die hier beschriebene erste Ausbaustufe bereits weitgehend im Berliner Open Data-Pilotprojekt umgesetzt wurde. Eine mögliche Weiterentwicklung für einen nachhaltigen Betrieb sollte daher auf den vorhandenen Komponenten aufbauen und diese weiterführen.

5

Geschäftsmodelle

Neben der Förderung von Transparenz und Partizipation beruht die Mehrwertgenerierung der Berliner Open Data-Plattform auch darauf, dass öffentliche Daten für die Entwicklung innovativer kommerzieller Anwendungen genutzt und öffentliche mit kommerziellen Datenbeständen verknüpft werden. Damit ergibt sich neben dem Ansatz, Daten kostenlos bzw. grenzkostenorientiert zur Verfügung zu stellen, auch die Perspektive, kommerzielle Geschäftsmodelle zu etablieren. Nach Schätzungen werden im EU-Markt mit (kostenlos oder gegen Gebühren) weiterverwendeten öffentlichen Daten mindestens 27 Milliarden Euro im Jahr umgesetzt.¹

Grundsätzlich ist die Realisierung eines Geschäftsmodells ein wichtiger Baustein für die Verstetigung der Berliner Open Data-Plattform. Nachfolgend werden daher zunächst potenzielle Produkte und Kunden betrachtet. Anschließend werden Monetarisierungspotenziale eingeschätzt und Modelle für geschäftliche Gestaltungsoptionen skizziert.

5.1 Produkte und Kunden

Das zentrale und unmittelbare Produkt der Plattform sind die bereitgestellten Daten, für die sowohl Bürger, Unternehmen – insbesondere auch aus der Medienbranche – als auch Verwaltungen als Kunden in Frage kommen. Mittelbare Produkte sind Dienste und Dienstleistungen rund um die Daten. Dazu zählen elektronische Dienste für die Aggregation, Auswertung und Visualisierung von Daten, die von Anwendungsentwicklern, aber

¹Pressemeldung der Europäischen Kommission vom 16. Juni 2011: »Digitale Agenda: Preise für die kreative Weiterverwendung offener Daten«

auch von Unternehmen und Verwaltungen, z. B. im Rahmen der Analyse bestimmter Problemfelder genutzt werden. Darüber hinaus sind nichtelektronisch erbrachte Dienstleistungen, wie z. B. die manuelle Auftragsrecherche im Datenbestand und die Veredelung von Daten, denkbar. Bei Onlineplattformen ist zusätzlich die Platzierung von Werbung eine mögliche Einnahmequelle. Kunden sind hier Unternehmen, die ihre Produkte und Dienstleistungen kontextsensitiv bewerben.

Produkte	Kunden
Daten (z. B. Geodaten)	Bürger, Unternehmen, Verwaltung
Dienste und Dienstleistungen (z. B. Datenaggregation)	(Bürger), Unternehmen, Verwaltung
Werbung (z. B. stadtbezogene Werbung analog berlin.de)	Unternehmen

Tabelle 5.1: Produkte und Kunden offener Daten

5.2 Monetarisierungspotentiale

Bei der Betrachtung der Monetarisierungspotenziale wird nachfolgend in erster Linie das Produkt Daten berücksichtigt: Zum einen handelt es sich dabei um das Kernprodukt des Angebots, zum anderen konzentriert sich der technische Realisierungsumfang der pilotierten Plattform auf die Bereitstellung von Daten und nicht auf Zusatzdienste. Bei den im Anschluss skizzierten übergeordneten Optionen für Geschäftsmodelle werden Dienste, Dienstleistungen und Werbung jedoch als mögliche Zusatzprodukte mitbetrachtet.

Zur Bewertung der Monetarisierungspotenziale lässt sich eine Matrix aufstellen, deren Dimensionen die Herkunft der Daten (öffentlich, privat) und die Art der Datennutzung (privat, kommerziell) sind (vgl. Tabelle 5.2):

- Die Bereitstellung öffentlicher Daten für die private Nutzung birgt kein bzw. ein zu vernachlässigendes Monetarisierungspotenzial. Hier hat sich der Grundsatz, dass Open Government Data für Bürgerinnen und Bürger kostenfrei angeboten werden sollen, weitestgehend durchgesetzt. Demzufolge ist davon auszugehen, dass die Zahlungsbereitschaft in der Zielgruppe außerordentlich gering ist.
- Bei der Bereitstellung öffentlicher Daten für die kommerzielle Nutzung hingegen ist die Erhebung von Kostenbeiträgen als realistisch anzusehen. Mit der kommerziellen Nutzung der Daten verbessern Unternehmen ihre Angebote mit neuen bzw. verbesserten Produkten und können dadurch mehr Umsatz und Gewinn erzielen. Damit kann die grundsätzliche Zahlungsbereitschaft als hoch eingeschätzt werden. Nach dem geltenden Informationsweiterverwendungsgesetz können für Verwaltungsdaten Entgelte erhoben werden, die sich im Rahmen der Kosten der Erfassung, Re-

produktion und Verbreitung zuzüglich einer angemessenen Gewinnspanne bewegen. Somit sind Einnahmen im beschränkten Umfang erzielbar.²

- Eine weitere Option ist die Bereitstellung nichtöffentlicher Daten zur privaten Nutzung. Dabei ist zu unterscheiden zwischen Daten, die von Bürgerinnen und Bürgern erhoben und für die Allgemeinheit bereitgestellt werden (z. B. private Restaurantbewertungen), und Daten die von Unternehmen für die nichtkommerzielle Nutzung bereitgestellt werden (z. B. Testberichte zu Produkten). Für erstgenannte Daten ist der Community-Gedanke bestimmend: Die privaten Datenberesteller profitieren auf der anderen Seite als Nutzer von anderen privat bereitgestellten Daten. Der Anspruch, Entgelte für diese Daten zu erheben, aber auch die Zahlungsbereitschaft ist gering. Für Daten, die von Unternehmen für die nichtkommerzielle Nutzung angeboten werden, besteht hingegen bereits heute ein Markt, der jedoch aufgrund der viel zitierten »Kostenlos-Mentalität« im Internet und einhergehend mit der schleppenden Akzeptanz von Premium-Angeboten im Vergleich zum Business-to-Business-Bereich gering ausfällt. Die kostenpflichtige Bereitstellung von Daten für kommerzielle Zwecke ist jedoch gegenüber der kostenfreien Bereitstellung unter Betrachtung der gesamtwirtschaftlichen Effekte (höhere Steuereinnahmen durch höhere Unternehmensgewinne und neu entstehende Unternehmen) kritisch abzuwägen.
- Die kostenpflichtige Bereitstellung nichtöffentlicher Daten für die kommerzielle Nutzung (z. B. Daten von Marktforschungsunternehmen) ist heute bereits geläufig und bildet einen Markt mit hohen Wachstumsraten. Das Monetarisierungspotenzial ist dementsprechend als hoch zu bewerten. Für die Umsetzung der beiden letztgenannten Optionen ist die Plattform in Richtung eines Marktplatzes zu erweitern, der es auch externen Akteuren ermöglicht, Daten bereitzustellen, einzukaufen und abzurechnen.

	private Nutzung	kommerzielle Nutzung
öffentliche Daten	kein	beschränkt
private bzw. privatwirtschaftliche Daten	beschränkt	hoch

Tabelle 5.2: Monetarisierungspotenziale offener Daten

Auf Basis der obigen Betrachtungen können drei übergeordnete Geschäftsmodelle für die Berliner Open Data-Plattform entworfen werden.

²Die kostenpflichtige Bereitstellung von Daten für kommerzielle Zwecke ist jedoch gegenüber der kostenfreien Bereitstellung unter Betrachtung der gesamtwirtschaftlichen Effekte (höhere Steuereinnahmen durch höhere Unternehmensgewinne und neu entstehende Unternehmen) kritisch abzuwägen.

5.3 Mögliche Geschäftsmodelle für Berlin

Die anhand der Vorüberlegungen identifizierbaren drei Geschäftsmodelle sind (I) der Betrieb der Plattform als Costcenter, finanziert durch das Land Berlin, (II) ein differenzierendes Modell, in dem die private Nutzung kostenfrei ist, die kommerzielle Nutzung und Verwertung der Daten aber kostenpflichtig ist, und (III) ein Marktplatzmodell, in dem die Plattform auch externen, privatwirtschaftlichen Anbietern offensteht, die eigene Daten anbieten. Die Möglichkeit, die jeweiligen Modelle zu verfolgen, hängt dabei neben den wirtschaftlichen Fragestellungen auch vom Ausbaustadium der Plattform ab. Die verschiedenen Modelle werden im Folgenden kurz anhand ihres Nutzenversprechens, ihrer Ausgestaltung und ihres Ertragsmodells beschrieben. Zusätzlich werden noch kurz Ergänzungsprodukte beschrieben, die in jedes der beschriebenen Geschäftsmodelle Eingang finden könnten.

5.3.1 Option I: Betrieb als Costcenter der Stadt Berlin

a) Ausgestaltung des Modells:

Dieses Modell geht davon aus, dass die Berliner Open Data-Plattform zentral vom Land Berlin finanziert wird und von Bürgern, Unternehmen und Verwaltungen kostenfrei genutzt werden kann. Ein wesentlicher Vorteil des Modells sind die relativ geringen Anforderungen an zusätzliche technische Komponenten und organisatorische Prozesse, da auf Abrechnungsfunktionalitäten verzichtet werden kann. Es ist zudem flexibel mit verschiedenen bestehenden Geschäftsmodellen der Berliner Verwaltung kombinierbar und auch technisch und organisatorisch flexibel ausgestaltbar. So kann die Berliner Open Data-Plattform einerseits als strukturierter, durchsuchbarer Katalog der Berliner Verwaltungsdaten aufgebaut werden, der — entsprechend der vorgeschlagenen ersten Ausbaustufe — Verwaltungsdaten um Metadaten anreichert und verlinkt, während die Daten dezentral gepflegt und bereitgestellt werden. Interessant wäre auch ein weiterer Ausbau der Plattform, der die Zusammenarbeit und Informationsteilung innerhalb der Verwaltung weiter vereinfachen würde. Im Land Berlin könnte in diesem Modell die BerlinOnline Stadtportal GmbH & Co. KG — wie auch bereits in der Pilotphase geschehen — die Gestaltung und den Betrieb des Datenportals (daten.berlin.de) übernehmen. Der Betrieb der technischen Komponente des Datenregisters, das in der Pilotphase von Fraunhofer FO-KUS gehostet und verwaltet wird, könnte hingegen durch das IT-Dienstleistungszentrum Berlin (ITDZ) erfolgen. Das Datenregister hält lediglich die Metadaten und Verweise zu den offenen Verwaltungsdaten vor. Als Zusatzleistung, die ebenfalls durch das ITDZ angeboten werden könnte, bietet sich daher auch das Hosting von Datensätzen an.

b) Nutzenversprechen:

Das wesentliche Ziel des Betriebs der Berliner Open Data-Plattform als Costcenter ist es, Bürgern, Unternehmen und Verwaltungen einen besseren Zugang zu öffentlichen Daten zu ermöglichen und die Nutzung dieser Daten zu vereinfachen. Bürger, Unternehmen

und Verwaltungen profitieren dabei im Wesentlichen von der Zeitersparnis beim Suchen und gegebenenfalls Aufbereiten der relevanten Daten. Gerade Unternehmen ermöglicht eine offene, strukturierte Bereitstellung der Daten, zudem weitere Geschäftsfelder. Das Ziel einer als Costcenter betriebenen Berliner Open Data-Plattform ist es, diesen Nutzen mit seinen verbundenen Wohlfahrtsgewinnen zu maximieren. Die Kosten werden daher vom Land Berlin getragen. Dieses Modell kombiniert eine zentrale Plattform, die (»top-down«) von der Senatsverwaltung für Wirtschaft, Technologie und Forschung bereitgestellt wird, mit der dezentralen (»bottom-up«) Bereitstellung von Daten durch einzelne Senatsverwaltungen und schränkt daher einzelne Berliner Akteure in ihrer Handlungs- und Entscheidungsfreiheit ein.

c) Erlösmodell:

Entsprechend der Ausrichtung des Geschäftsmodells als Costcenter wird im Wesentlichen auf zusätzliche externe Erlöse verzichtet. Als neue Erlösquellen kommen letztlich lediglich solche aus den Ergänzungsprodukten wie Werbung und Serviceleistungen in Frage.

Der Ansatz der Ausbaustufe 1, der wie ein Verzeichnis der öffentlichen Berliner Daten funktioniert, trägt dazu bei, bereits bestehende Umsätze zu verbessern, da die bestehenden Erlösmodelle mit öffentlichen Daten im Land Berlin nicht angegriffen werden, während der Zugang zu diesen Angeboten verbessert wird. Im Land Berlin könnte die BerlinOnline Stadtportal GmbH & Co. KG einen Teil ihres Betriebsaufwandes — sofern politisch gewünscht — über Werbung refinanzieren. Es ist jedoch davon auszugehen, dass für den Betrieb sowie die Pflege und Weiterentwicklung der Plattform zusätzliche Finanzmittel benötigt werden. Diese müssten durch das Land Berlin zur Verfügung gestellt werden. Gleiches gilt für den Betrieb des Datenregisters, der vom ITDZ übernommen werden könnte. Aufwände für Leistungen des Datenhostings sind hingegen in diesem Modell jeweils von den Dateneigentümern zu übernehmen.

5.3.2 Option II: Private Nutzung kostenfrei, kommerzielle Nutzung kostenpflichtig

a) Ausgestaltung des Modells:

Dieses Modell geht davon aus, dass private Nutzer die Berliner Open Data-Plattform kostenlos nutzen können, während kommerzielle Nutzer für den Service bzw. für die Lizenz zur Datennutzung zahlen müssen. Damit dieses Modell für die Berliner Open Data-Plattform angewendet werden kann, ist eine Einigung mit den datenbereitstellenden Verwaltungen erforderlich. Die Anforderungen sowohl an den Ausbaugrad der Plattform (z. B. Zahlungsprozesse und Zahlungskomponenten) als auch an die organisatorischen Abstimmungen sind damit höher als in der ersten beschriebenen Option. Auch in diesem Modell könnte die BerlinOnline Stadtportal GmbH & Co. KG als Betreiber des Portals auftreten und das ITDZ Leistungen zum Betrieb des Datenregisters und eventueller Abrechnungskomponenten sowie zusätzliche Dienstleistungen im Bereich des Datenhostings anbieten.

b) Nutzenversprechen:

Auch bei diesem Modell ist das wesentliche Ziel des Betriebs der Berliner Open Data-Plattform, Bürgern, Unternehmen und Verwaltungen einen besseren Zugang zu öffentlichen Daten zu ermöglichen und die Nutzung dieser Daten zu vereinfachen. Ein wesentlicher Unterschied ist jedoch, dass die Verwaltung an Profiten, die mit ihren Daten erzielt werden, zu partizipieren sucht und somit zusätzlich Extraeinnahmen erzielen sollte. Neben dem Ziel, Wohlfahrtsgewinne durch eine Offenlegung von öffentlichen Daten und eine Verbesserung des Zugangs zu erzielen, wird also auch das Ziel verfolgt, einen Teil der Kosten für die Erhebung, Aufbereitung und Bereitstellung zumindest von den Akteuren zurückzubekommen, die ihr Geschäft auf diesen Daten aufbauen. Das Modell erfordert höhere Abstimmungsaufwände, da zu beachten ist, dass die Bezirke – in ihrer Rolle als Datenbereitsteller – im Land Berlin ein erhebliches Mitspracherecht haben. Die damit verbundene hohe Anzahl an einzubindenden Akteuren und die Berücksichtigung individueller Anforderungen können die Umsetzung des Modells erheblich erschweren.

c) Erlösmodell:

Da Nutzer der Berliner Open Data-Plattform nicht ohne Weiteres in ihrer Rolle als Bürger oder Unternehmen identifiziert werden können, ist es kaum möglich, jegliche unternehmerische Nutzung als kommerzielle Nutzung abzurechnen. Es erscheint daher sinnvoll, kommerzielle Nutzung so zu definieren, dass diese einer Nutzung der Berliner Open Data-Plattform entspricht, bei der ein Produkt auf Basis oder mithilfe der über die Plattform bereitgestellten Daten angeboten wird. Dies führt dazu, dass private Nutzer nicht unbedingt deckungsgleich mit Bürgern und kommerzielle nicht unbedingt deckungsgleich mit Unternehmen sind, da auch Unternehmen öffentliche Daten (z. B. für ihre eigene Marktanalyse) nutzen können, obwohl sie ihr Geschäft mit völlig anderen Produkten und Dienstleistungen machen. Kommerzielle Nutzer wären als solche zu erkennen, wenn sie kommerzielle Produkte anbieten, die offensichtlich auf Daten der Berliner Open Data-Plattform basieren. Zusätzlich kommen Erlöse aus den Ergänzungsprodukten wie Werbung und Serviceleistungen in Frage, wobei gerade bei den Serviceleistungen auch eine Möglichkeit bestünde, die Angebote für private Nutzer und für kommerzielle Nutzer zu differenzieren. Hierbei würde das kommerzielle Angebot z. B. Datenstrukturen und erweiterte Schnittstellen bereitstellen, die eine maschinelle Auswertung und Verarbeitung der Daten weiter vereinfachen. Ein Nachteil wäre allerdings, dass bei einem solchen exklusiven Zugang zu besseren Schnittstellen die private Nutzung leiden könnte und erwünschte nichtkommerzielle Anwendungen eventuell nicht realisiert würden.

Erlöse könnten in diesem Modell über prozentuale Beteiligungen am Umsatz der Produkte, die Daten der Berliner Open Data-Plattform verwenden, als feste Beträge pro Zugang oder als feste Beträge pro Abruf abgerechnet werden. Gerade das Modell der prozentualen Beteiligungen am Umsatz ist aber nur mit hohem Aufwand durchsetzbar. In jedem Fall ist der Aufwand beim Kontrollieren verwendeter Inhalte und beim Eintreiben von Entgelten zu berücksichtigen. Das Erlösmodell mit festen Beträgen für Zugang oder Abruf ist einfacher zu realisieren, führt aber zu einer geringeren Abschöpfung der Zahlungsbereitschaft, da die Preise für einige Produkte zu hoch wären, sodass diese nicht

realisiert würden, für andere Produkte hingegen niedriger als die tatsächliche Zahlungsbereitschaft.

Ein Erlösmodell, das mit festen Beträgen pro Abruf arbeitet, ist mit dem Modell vergleichbar, das Google für seine Maps-Dienste verwendet. Im Land Berlin würde dieses Modell insgesamt etwas höhere Kosten bei den Betreibern verursachen als das Modell eines Costcenters, da zusätzliche Zahlungskomponenten und Prozesse sowie ein Abrechnungssystem gegenüber den Datenbereitstellern benötigt würden. Für eine Refinanzierung der Betriebsleistungen, der Pflege und Weiterentwicklung der Plattform und der Kontrolle und Abrechnung der Datennutzung käme zunächst ebenfalls Werbung als Erlösquelle in Frage. Durch die höheren Kosten kann aber mit einer Kostendeckung durch Werbung gerechnet werden. Die zusätzlichen Kosten müssten also entweder vom Land Berlin getragen werden, sodass die vollen Umsätze mit den Daten an die Datenbereitsteller weitergereicht werden könnten, oder es müssten Transaktionsgebühren erhoben werden, die eine Kostendeckung sicherstellen müssten und von den Erlösen der Datenbereitsteller abzuziehen wären.³ Auch die Kosten für den Betrieb und die Wartung des Datenregisters, das vom ITDZ Berlin bereitgestellt werden könnte, müssten entweder vom Land Berlin übernommen werden oder in die vom Plattformbetreiber erhobenen Transaktionsgebühren eingepreist werden (hier wären Abstimmungen zwischen dem Plattformbetreiber und dem Betreiber des Datenregisters nötig). Leistungen des Datenhostings (ebenso gegebenenfalls ITDZ) sollten, wenn sie optional sind, von den Dateneigentümern übernommen werden, und wenn sie verpflichtender Bestandteil der Plattform sind, vom Land Berlin übernommen werden bzw. in die Transaktionsgebühren eingepreist werden.

5.3.3 Option III: Marktplatzmodell – Plattform mit Transaktionsgebühren

a) Ausgestaltung des Modells:

Dieses Modell geht davon aus, dass die Berliner Open Data-Plattform als Marktplatz aufgebaut wird, der nicht nur der Berliner Verwaltung zur Verfügung steht, sondern auch von privatwirtschaftlichen Datenanbietern genutzt wird. Dabei soll sich die Plattform über Gebühren der Datenanbieter finanzieren. Um dieses Modell zu etablieren, muss die Plattform ebenfalls mindestens in Ausbaustufe 2 vorliegen und zudem über Zahlungskomponenten verfügen. Darüber hinaus sollte zusätzlich die Möglichkeit bestehen, die privatwirtschaftlich eingebrachten Daten in ihrer Qualität zu sichern und Umsatzanteile der Datenanbieter automatisch durchzureichen. Beide Aspekte führen zu einem weiter erhöhten organisatorischen und technischen Aufwand. Im Land Berlin könnte in diesem Modell die BerlinOnline Stadtportal GmbH & Co. KG sinnvollerweise den Betrieb des Datenportals (daten.berlin.de) übernehmen. Das ITDZ könnte hingegen wiederum Leistungen zur Pflege des Datenregisters und zusätzlich die in dieser Ausbaustufe zwingend erforderliche Dienstleistung im Bereich des Datenhostings anbieten.

³Es wäre auch ein Mischmodell denkbar, das beide Ansätze vereint.

b) Nutzenversprechen:

Dieses Modell bietet Datenanbietern einen deutlichen Nutzenvorteil, da es ihnen einen weiteren Vertriebsweg eröffnet. Für Datennutzer liegt der Vorteil hingegen auch bei diesem Modell im höheren Komfort des Zugangs, wobei von der Zeitersparnis beim Suchen und gegebenenfalls von den besser aufbereiteten Daten profitiert wird; ein Vorteil, der durch die Bündelung von Daten der öffentlichen Verwaltung und Daten, die von privaten Anbietern bereitgestellt werden, sogar noch steigt. Die eventuell höheren Preise dürften allerdings zulasten des Vorteils der Datennutzer gehen. Ein wesentlicher Vorteil dieses Modells ist jedoch, dass hier für die Betreiber und Finanziere der Berliner Open Data-Plattform die Möglichkeit besteht, einen größeren Teil der Kosten zurückzuerhalten.

c) Erlösmodell:

Das Erlösmodell des Marktplatzes bietet insbesondere zwei verschiedene Optionen, die jedoch auch kombiniert werden können. Zum einen könnten fixe Beträge für die Platzierung eines Datenangebots genommen werden und zum anderen könnten Transaktionsgebühren von jedem verkauften Datenpaket erhoben werden. Beim Modell der Transaktionsgebühren ergeben sich wiederum zwei Varianten: Bei der ersten Variante wird ein fixer Preis pro Transaktion erhoben, bei der zweiten Variante wäre ein prozentualer Betrag vom Umsatz eines bestimmten Datensatzes entscheidend, wobei die Datenbereitsteller den Preis ihrer Daten selbst festlegen würden und von diesem Preis bei jeder Transaktion einen prozentualen Anteil abgeben würden. Die Datenanbieter können sich dabei dazu entschließen, ihre Daten unentgeltlich bereitzustellen, was wiederum stärker den Nutzern der Plattform zugute käme.

Bei hoher Sichtbarkeit der Plattform käme gegebenenfalls auch der Vertrieb von Apps, die über den Marktplatz bezogen werden und auf den Daten aufsetzen, in Frage. Hierbei würden ebenfalls Transaktionsgebühren erhoben. In Bezug auf den Vertrieb von Apps ist aber mit einer harten Konkurrenz durch etablierte Plattformen (iTunes, Android Market, Windows Marketplace etc.) zu rechnen. Im Land Berlin würde dieses Modell von allen Modellen wohl insgesamt die höchsten Kosten beim Betreiber des Portals (z. B. BerlinOnline Stadtportal GmbH & Co. KG) verursachen. Auch beim Betreiber des Datenregisters und des Datenhostings würde das Modell höhere Kosten verursachen, da neben öffentlichen Daten nun auch private Daten über die Plattform angeboten werden könnten und damit sowohl das Volumen als auch die Kundeninteraktionen steigen dürften. Es würden zudem auch hier zusätzliche Zahlungskomponenten und Prozesse sowie ein Abrechnungssystem gegenüber den Datenbereitstellern benötigt. Für eine Refinanzierung der Betriebsleistungen, der Pflege und Weiterentwicklung der Plattform und der Kontrolle und Abrechnung der Datennutzung kämen insbesondere Transaktionsgebühren in Frage (auch hier wären Abstimmungen zwischen dem Portalbetreiber und dem Betreiber des Datenregisters nötig). Eine unterschiedliche Behandlung von kommerziellen Anbietern und Verwaltungsanbietern wäre dabei aber möglich.

5.3.4 Zusatzgeschäft

Potenzial für ein Zusatzgeschäft besteht im Wesentlichen in zwei Bereichen:

5.3.4.1 Werbung

Werbung ist eine der klassischen Einnahmequellen im Internet. Der Vorteil bei Internetwerbung ist im Allgemeinen, dass eine sehr gute Zielgruppenselektion und somit eine sehr gezielte Ansprache dieser Zielgruppe möglich ist. Die hohe Konkurrenz führt dennoch zu relativ geringen Preisen für Werbeplätze. Gängig sind Preise im Bereich von –10 Euro pro 1000 Seitenaufrufe, wobei die Preise abhängig von der Attraktivität der Zielgruppe schwanken. Da die Vermarktung der Werbeplätze eigenständig relativ schwierig ist, sollte zusätzlich ein Anteil von circa 40 % für die Vermarkter einkalkuliert werden. Für die Berliner Open Data-Plattform kann Werbung Zusatzeinnahmen bringen, wobei nicht davon auszugehen ist, dass die Einnahmen auch nur annähernd den Betriebskosten entsprechen würden. Auch kann aggressiv platzierte Werbung Nutzer abschrecken und damit den Gesamtnutzen der Plattform für Berlin signifikant reduzieren. Hier hat eine Abwägung zu erfolgen, inwiefern die Zusatzeinnahmen dies rechtfertigen.

5.3.4.2 Zusatzleistungen

Unter Zusatzleistungen fallen insbesondere Dienste und Dienstleistungen, die sich im Bereich der Datenaufbereitung, -analyse und -recherche bewegen. So ist es z. B. denkbar, die Maschinenlesbarkeit von Daten zu verbessern, um deren Nutzbarkeit für automatisierte Auswertungen oder für komplexe Anwendungen zu erhöhen. Zusatzleistungen sind insbesondere für Nutzer mit kommerziellen Interessen attraktiv und bieten der Berliner Open Data-Plattform die Möglichkeit, mit diesem Angebot die höhere Zahlungsbereitschaft der Unternehmen, die direkt an den Daten verdienen, abzuschöpfen, wobei diesen gleichzeitig auch ein höherer Service geboten wird. Im Unterscheid zum Werbegeschäft erfordert das Zusatzgeschäft jedoch signifikante Mehraufwände, sodass eine Abwägung der Sinnhaftigkeit vorgenommen werden sollte.

5.4 Fazit

Mit der Fertigstellung der Berliner Open Data-Plattform in der ersten Ausbaustufe erscheint zunächst nur die erste Option des Betriebs als Costcenter in Frage zu kommen, wobei mit dem Zusatzprodukt Werbung ein kleinerer Beitrag zur Kostendeckung des Betriebs erzielt werden könnte. Die Datenbereitsteller aufseiten der Verwaltung könnten jedoch von einem besseren Zugang der Datennutzer profitieren, da dieser idealerweise ihre eigenen Umsätze mit kostenpflichtigen Daten steigen lassen sollte.

Mit zunehmendem Ausbau der Plattform werden die Optionen II und III interessant, bei denen auch zusätzliche Produkte aus dem Bereich der Datenaufbereitung eine größere Rolle spielen dürften. Eine Umstellung des Geschäftsmodells von Option I auf Option II oder III sollte jedoch im Vorfeld mit den relevanten Akteuren aufseiten sowohl der Datenbereitsteller als auch der Plattformbetreiber diskutiert werden.

6

Rechtliche Rahmenbedingungen

Als Teil dieser Konzeptstudie wird eine Internetplattform umgesetzt, die in erster Linie den Berliner Bürgern die Möglichkeit geben soll, aufbereitete Basisdaten von Berliner Behörden zu nutzen. Somit wird u. a. auch der in der EU-Richtlinie 2003/98/EG formulierten Forderung der EU-Kommission nachgekommen, den Rückstand zu den USA auf dem Informationsmarkt aufzuholen und generell die Schaffung höherwertiger Mehrwertinformationsprodukte zu ermöglichen (Püschel, DuD 2006, S. 481). Um dieses Ziel in Deutschland zu verwirklichen, wurde durch die Umsetzung der EU-Direktive 2003/98/EG in deutsches Recht der Grundsatz der beschränkten Aktenöffentlichkeit in der Verwaltung gelockert (Sydow, NVwZ 2008, S. 481). Für die Verwirklichung von Open Data muss es also möglich sein, problemlos an die Daten der Behörden zu kommen und diese uneingeschränkt weiterzuverwenden. Trotzdem ist die kommerzielle Weiterverwendung von Daten öffentlicher Stellen nicht vorbehaltlos garantiert. Zahlreiche gesetzliche Ausnahmen, aber auch rechtliche Widersprüche erschweren eine Weiterverwendung öffentlicher Daten. In diesem Bericht soll erklärt werden, wie einerseits derzeit Daten von öffentlichen Stellen im Bundesland Berlin aus Sicht einer natürlichen oder juristischen Person beschafft werden können und wie andererseits Behörden rechtswirksam Datensätze bereitstellen können. Dabei wird eine Übersicht darüber gegeben, welche Gesetze als Rechtsgrundlage in Betracht kommen, welche Gesetze vorrangig angewendet werden müssen und inwiefern die Bestimmungen des Datenschutzes die Akteneinsicht begrenzen können. Außerdem wird die Art und Weise der Lizenzierung von Daten öffentlicher Stellen untersucht. Zum Schluss wird auf den Schutz von personenbezogenen Daten von möglichen Nutzern der Plattform eingegangen.

6.1 Rohdatenbeschaffung aus Sicht einer natürlichen oder juristischen Person

6.1.1 Informationszugang und Akteneinsicht in Berlin gemäß BerlIFG

Zuerst muss die Information, die weiterverwendet werden soll, beschafft werden. Das IWG regelt die Weiterverwendung, gemäß § 3 I 2 IWG aber nicht den Zugang zur Information, es baut nur auf bestehende Zugangsregelungen auf (Polenz, NVwZ 2010, S. 488). Es gibt mehrere Wege, Informationen öffentlicher Behörden zu erhalten. Welcher Weg am sinnvollsten ist, hängt von der gewünschten Information ab. Handelt es sich um eigene personenbezogene Daten, können Datenschutzvorschriften wie z. B. §§ 6, 19 und 34 BDSG einen Anspruch begründen. Als Verfahrensbeteiligter kann man sich auf Akteneinsicht gemäß § 29 VwVfG berufen (Schmitz, NVwZ 2005, S. 985). Im Rahmen einer ODP wird es sich regelmäßig nicht um eine Verfahrensbeteiligung oder eigene Daten handeln, sondern um wirtschaftlich verwertbare Basisdaten der Behörden. Für diese Informationen ist das BerlIFG einschlägig.

6.1.1.1 Antragsberechtigung gemäß § 3 I 1 BerlIFG

Das BerlIFG gewährt gemäß § 3 I BerlIFG jedem Menschen und allen privaten juristischen Personen ein Recht auf Akteneinsicht in Akten, die von öffentlichen Stellen in Berlin geführt werden. Die rechtliche Gesellschaftsform, die die Betreiberin der Plattform sein wird, wird also selbst antragsberechtigt sein, wenn sie eine Kapitalgesellschaft (z. B. GmbH) sein wird. Nicht anspruchsberechtigt sind juristische Personen des öffentlichen Rechts (Husein, LKV 2010, S. 338). Das Oberverwaltungsgericht Berlin-Brandenburg entschied am 2. Oktober 2007, dass ein Anspruch auf Akteneinsicht nur für Akten besteht, »wenn grundsätzlich lediglich auf solche amtlichen Informationen, die tatsächlich bei der Behörde vorhanden sind« (LSK 2008, 330203), Zugriff erfolgen soll. So könnte die Behörde durch vorsätzliche Weggabe der gewünschten Akten gezielt einen Anspruch aus dem BerlIFG vereiteln (Hartleb, NVwZ 2009, S. 826), da eine Wiederbeschaffungspflicht der Behörde nur besteht, wenn zum Zeitpunkt des Antrages die Akte der Behörde vorlag und sie Verfügungsberechtigt war (Husein, LKV 2010, S. 339), und selbst dann müsste der Dritte die Akte freiwillig zurückgeben (Hartleb, NVwZ 2009, S. 825).

6.1.1.2 Anwendungsbereich gemäß § 2 I BerlIFG

Verpflichtet zur Auskunft sind gemäß § 2 I BerlIFG alle öffentlichen Stellen des Landes Berlin. Dazu gehören Behörden, nicht rechtsfähige Anstalten, Krankenhausbetriebe und Eigenbetriebe des Landes Berlin, landesunmittelbare Körperschaften des öffentlichen Rechts, Stiftungen und Anstalten, aber auch Private im Rahmen der Ausübung hoheitlicher Befugnisse. Ausnahmen gelten bei Gerichten und Staatsanwaltschaften. Hier wird

Auskunft oder Einsicht nur erteilt, soweit es sich um Akten handelt, die Verwaltungsaufgaben betreffen (Partsch, LKV 2001, S. 98). Für den Zugang zu Umweltinformationen im Land Berlin gilt gemäß § 18 a I BerlIFG das Umweltinformationsgesetz des Bundes.

6.1.1.3 Akten i. S. d. § 3 II BerlIFG

»Akten im Sinne dieses Gesetzes sind alle schriftlich, elektronisch, optisch, akustisch oder auf andere Weise festgehaltenen Gedankenverkörperungen und sonstige Aufzeichnungen, insbesondere Schriftstücke, Magnetbänder, Disketten, Filme, Fotos, Tonbänder, Pläne, Diagramme, Bilder und Karten, soweit sie amtlichen Zwecken dienen« (§ 3 II BerlIFG). Entscheidend für ein Zugangsrecht ist, ob die Akte Teil eines Verwaltungsvorgangs geworden ist. Dazu gehören in Berlin auch der Akte zugehörige Entwürfe und Notizen, auch wenn diese nicht Bestandteil der Akte werden sollen (Hussein, LKV 2010, S. 339).

Amtliche Zwecke Im IFG des Bundes werden »amtliche Informationen« genannt, die man mit den Akten im BerlIFG, die einem amtlichen Zweck dienen, vergleichen könnte. Zu den »amtlichen Informationen« hat der Verwaltungsgerichtshof München entschieden, dass bei einem Begehren nach Informationszugang aus kommerziellen Gründen keine amtlichen Informationen vorliegen würden (MMR, 2009, S. 292). Dies wird auf der Bundesebene von Schoch als gesetzeswidrig bezeichnet, da der Begriff »amtliche Information« »nach der Legaldefinition (§ 2 Nr. 1 IFG) nicht aus der Perspektive des Antragstellers, sondern nach den gesetzlich festgelegten objektiven Merkmalen zu bestimmen« (Schoch, NJW 2009, S. 2990) sei. Diese Argumentation lässt sich auch auf das Land Berlin übertragen. Weiterhin sei es nicht zulässig, die (kommerziellen) Motive des Anspruchstellers zu hinterfragen, da ja gerade ein voraussetzungsloser Anspruch gewährt werden soll (ebd.). Diese Argumentation könnte auf das Land Berlin übertragen werden, wenn man der Ansicht ist, dass der Zweck des BerlIFG, staatliches Handeln zu kontrollieren und die demokratische Meinungs- und Willensbildung zu fördern, einen voraussetzungslosen Anspruch beinhaltet.

Aktenbegriff Im BerlIFG heißen die Daten im Gegensatz zu den meisten anderen Landesgesetzen nicht Informationen, sondern Akten. Dies heißt nach dem Oberverwaltungsgericht Berlin-Brandenburg, dass das BerlIFG nur ein aktenbezogenes Informationsrecht gewährt, was einen rechtlichen Unterschied zum reinen Informationsrecht darstellt (Hussein, LKV 2010, S. 339). Ein Journalist darf demnach keine Einsicht in oder Auskunft über den Terminkalender des Regierenden Bürgermeisters gewährt bekommen, da es sich bei diesem um keine Akte im verwaltungsrechtlichem Sinne handele (Schoch, NJW 2009, S. 2989). Bei den Tischkalendern von Mitarbeitern des Berliner Landesamtes für Mess- und Eichwesen lässt sich aber ein Verwaltungsbezug herstellen, »da sie der Übersicht bezüglich der kontrollierten Unternehmen und als Grundlage der Jahresstatistik dienen. (Husein, LKV 2010, S. 339)«

6.1.1.4 Umfang des Informationsrechts gemäß § 4 I BerlIFG

Gemäß § 4 I BerlIFG besteht grundsätzlich ein Recht auf Akteneinsicht oder Aktenauskunft im vollen Umfang des Antrags. Jedoch muss die Akteneinsicht verweigert werden, wenn personenbezogene Daten oder Betriebs- und Geschäftsgeheimnisse betroffen sind. Das Informationsrecht kann verweigert werden, wenn öffentliche Belange oder behördliche Entscheidungsprozesse geschützt werden sollen.

6.1.1.5 Personenbezogene Daten

Gemäß § 4 I 1 BerlDSG sind personenbezogene Daten Einzelangaben über persönliche oder sachliche Verhältnisse einer bestimmten oder bestimmbaren natürlichen Person (Betroffener). Trotz personenbezogener Daten besteht ein Informationsrecht, wenn gemäß § 6 II BerlIFG eine Einwilligung des Betroffenen vorliegt oder gemäß § 6 I BerlIFG das Informationsinteresse die schutzwürdigen Belange des Betroffenen überwiegen. Dies ist in der Regel der Fall, wenn der Anspruchsteller keine Privatinteressen (Rache, Neugier) verfolgt und der Betroffene

- an einem Verwaltungs- oder sonstigen Verfahren beteiligt ist,
- vorgeschriebene Erklärungen, Anzeigen oder Auskünfte gegenüber einer Behörde abgegeben hat,
- von einer Behörde überwacht wurde,
- gemäß § 6 II 1 Nr. 1 lit. e BerlIFG Eigentümer, Pächter, Mieter, Gutachter oder Sachverständiger ist
- oder als Amtsträger an Verwaltungsvorgängen mitwirkt (Husein, LKV 2010, S. 340 ff.).

6.1.1.6 Betriebs- und Geschäftsgeheimnisse

Als Betriebs- und Geschäftsgeheimnisse werden allgemein alle auf ein Unternehmen bezogenen Tatsachen, Umstände und Vorgänge verstanden, die nicht offenkundig, sondern nur einem begrenzten Personenkreis zugänglich sind und an deren Nichtverbreitung der Rechtsträger ein berechtigtes Interesse hat (Husein, LKV 2010, S. 342). Auch ein durch die Akteneinsicht eintretender nicht nur unwesentlicher wirtschaftlicher Nachteil des Betroffenen kann, ohne dass ein Betriebsgeheimnis vorliegt, zur Verweigerung des Antrags führen (ebd.). Trotzdem kann aber ein Anspruch bestehen, wenn gemäß § 7 1 BerlIFG das Informationsinteresse schwerer wiegt als schutzwürdige Belange des Betroffenen.

6.1.1.7 Schutz des Urheberrechts

Wenn der Überlassung von Ablichtungen gemäß § 13 V 2 BerlIFG Urheberrechte entgegenstehen, muss die Behörde eine Einwilligung des Urhebers einfordern. Wenn dieser die Einwilligung verweigert, besteht kein Anspruch auf Ablichtungen von Akten. Der Anspruch auf eine Einsicht in diese, gegebenenfalls mit Anfertigung von Notizen gemäß § 13 V 4 BerlIFG, bleibt davon jedoch unberührt. Es wird davon ausgegangen, dass in der Überlassung der Akten an die Verwaltung durch den Urheber zumindest eine Einwilligung auf Einsicht vorliegt (Parsch, LKV 2001, S. 98).

6.1.1.8 Schutz besonderer öffentlicher Belange und des behördlichen Entscheidungsprozesses

Gemäß §§ 9, 10 und 11 BerlIFG besteht ein Recht auf Akteneinsicht oder -auskunft nicht, wenn

- durch die Einsicht oder Auskunft bevorstehende Maßnahmen zur Rechtsdurchsetzung oder Strafverfolgung vereitelt würden,
- ein laufendes Verwaltungsverfahren hinsichtlich der Entscheidungsgründe beeinflusst werden könnte
- oder die Einsicht oder Auskunft zur schwerwiegenden Gefährdung des Gemeinwohls führen könnte.

6.1.1.9 Verfahren

Der Antrag ist gemäß § 13 I 1 BerlIFG mündlich oder schriftlich zu stellen. Eine Begründung ist nicht erforderlich, außer es sind personenbezogene Daten oder Betriebsgeheimnisse betroffen (Husein, LKV 2010, S. 342). Wenn Rechte von Dritten oder Betroffenen von der Akteneinsicht oder Auskunft betroffen sind, müssen diese an der Entscheidung über einen Informationszugang durch eine Anhörung gemäß § 14 II 1 BerlIFG beteiligt werden. Gemäß § 15 BerlIFG muss bei Zugangsverweigerung eine Begründung der Behörde mit Nennung der Inhalte der vorenthaltenen Information erfolgen. Ebenso muss die Behörde begründen, warum keine beschränkte Einsicht oder Auskunft gemäß § 12 BerlIFG erfolgen kann. Bei einer eventuellen Zurückweisung des Antrags muss der Bescheid innerhalb von zwei Wochen erfolgen.

6.1.1.10 Verbot der gewerblichen Nutzung gemäß § 13 VII BerlIFG

In § 13 VII BerlIFG ist ein generelles Verbot der kommerziellen Weiterverwendung von Daten öffentlicher Stellen formuliert, weil der Berliner Landesgesetzgeber, anders als die

Bundesregierung, der Auffassung ist, dass ohne das Verbot auch eine kommerzielle Weiterverwendung der erlangten Informationen schon im BerlIFG mit geregelt sei (Schoch, NVwZ 2006, S. 875): »Die Veröffentlichung, Speicherung oder Sammlung von durch Akteneinsichten oder Aktenauskünfte erhaltenen Informationen zu gewerblichen Zwecken ist nicht zulässig.« Da öffentliche Stellen nicht antragsberechtigt sind, gilt das Verbot nur für private Antragsteller. Wer dem zuwiderhandelt, begeht gemäß § 22 I BerlIFG eine Ordnungswidrigkeit und muss gem. § 22 II BerlIFG zudem mit einer Geldbuße von bis zu 5000 Euro rechnen. Danach könnten z. B. App-Entwickler kein Entgelt mehr für ihre Produkte verlangen. Diese Regelung könnte jedoch gegen höherrangiges Recht verstoßen.

Verhältnis des § 13 VII BerlIFG zur EU-Richtlinie 2003/98/EG In Art. 2 Nr. 4 Richtlinie 2003/98/EG definiert die Richtlinie die Weiterverwendung als »Nutzung von Dokumenten, die im Besitz öffentlicher Stellen sind, durch natürliche oder juristische Personen für kommerzielle oder nichtkommerzielle Zwecke, die sich von dem ursprünglichen Zweck im Rahmen des öffentlichen Auftrags, für den die Dokumente erstellt wurden, unterscheiden«. Die Richtlinie bestimmt den kommerziellen Zweck als Teil der Definition der Weiterverwendung. Somit könnte der § 13 VII BerlIFG gegen höherrangiges EU-Recht verstoßen (Schoch, NVwZ 2006, S. 875). EU-Richtlinien entfalten jedoch erst unmittelbare Geltung, wenn sie vom betreffenden Mitgliedsstaat nicht fristgemäß umgesetzt wurden (Karpenstein in: Büchting/Heussen, Rn. 20). Mit dem Inkrafttreten des IWG 2006 kam Deutschland jedoch der Aufforderung zur Umsetzung nach. Fraglich ist daher, ob § 13 VII BerlIFG gegen die Bestimmungen des IWG verstößt.

Verhältnis des § 13 VII BerlIFG zum IWG Gemäß § 2 Nr. 3 IWG ist eine Weiterverwendung »jede Nutzung von Informationen, die über die Erfüllung einer öffentlichen Aufgabe hinausgeht und in der Regel auf die Erzielung von Entgelt gerichtet ist«. Das IWG sieht somit ebenfalls eine gewerbliche Nutzung der erlangten Informationen vor. Somit verstößt der § 13 VII BerlIFG auch gegen nationales Recht, denn im Bereich der konkurrierenden Gesetzgebung haben gemäß Art. 72 I GG die Länder nur so lange die Befugnis zu Gesetzgebung, wie der Bund nicht von seiner Gesetzgebungszuständigkeit Gebrauch macht.

Schlussfolgerung Der § 13 VII BerlIFG hat wegen Verstoß gegen höherrangiges nationales Recht und die EU-Richtlinie 2003/98/EG seine Rechtswirksamkeit verloren (Schoch, NVwZ 2006, S. 875) und könnte durch eine Verweisung auf die Bestimmungen des IWG ersetzt werden (Püschel, DuD 2006, S. 488). Nutzer von durch die ODP bereitgestellten Daten können diese also auch für gewerbliche Zwecke anbieten.

6.1.2 Die Informationsweiterverwendung gemäß IWG

Das IWG knüpft an bestehende Informationszugangsrechte an, um dann die kommerzielle Weiterverwendung der durch das Informationszugangsrecht erlangten Daten zu regeln (Sydow, NVwZ 2008, S. 483). In diesem Fall knüpft das IWG an das BerlIFG (der Widerspruch, der dieser Verknüpfung im Fall der kommerziellen Weiterverwendung durch § 13 VII BerlIFG innewohnt, wurde in 6.1.1.10 aufgelöst), das UIG und alle anderen fachgesetzlichen Zugangsregelungen an. Die Weiterverwendung muss bereits von der Behörde gestattet worden sein, dann ist auch Dritten in nicht diskriminierender Weise die Weiterverwendung zu gestatten (Altmeppen/Kahlen, MMR 2006, S. 498).

6.1.2.1 Anwendungsbereich

Gemäß § 1 I IWG gilt das IWG für alle bei öffentlichen Stellen vorhandenen Informationen. Öffentliche Stellen erfüllen im Allgemeininteresse liegende Aufgaben nicht-gewerblicher Art und werden überwiegend von anderen öffentlichen Stellen finanziert oder kontrolliert (Altmeppen/Kahlen, MMR 2006, S. 500). Sie sind nicht verpflichtet, »Anpassungen vorzunehmen oder die Erstellung von bestimmten Informationen für ihre Weiterverwendung fortzusetzen« (ebd.). Informationen sind gemäß § 2 Nr. 2 IWG jede Aufzeichnung, unabhängig von der Art ihrer Speicherung, »z. B. Informationen über Geografie, Wetter, Tourismus, Soziales, Wirtschaft, Geschäftsleben und Patentwesen. (ebd.)« Ausgeschlossen sind gemäß § 1 II Nr. 4 IWG Informationen, die durch Urheberrechte geschützt sind. Das heißt, dass, selbst wenn eine Einwilligung zur Nutzung des Rechteinhabers vorliegt (z. B. gemäß § 13 V 2 BerlIFG), das IWG nicht als Rechtsgrundlage zur Weiterverwendung dienen kann. Eine Weiterverwendung wäre in diesem Fall aber trotzdem möglich, wenn keine weiteren Rechte entgegenstehen, nur eben nicht nach dem IWG. Außerdem sind Informationen vom Anwendungsbereich ausgeschlossen, die sich gemäß § 1 II Nr. 5-7 IWG im Besitz von öffentlich-rechtlichen Rundfunkanstalten oder Bildungs-, Kultur- und Forschungseinrichtungen befinden oder gemäß § 1 II Nr. 1-2 IWG nicht zugänglich oder nicht voraussetzungslos zugänglich sind (Schoch, NVwZ 2006, S. 875). Gemäß § 1 II Nr. 3 IWG gilt das IWG ebenfalls nicht für Informationen, deren Erstellung nicht unter die öffentlichen Aufgaben fällt. Eine Behörde fertigte z. B. für ein Unternehmen ein Gutachten über Windverhältnisse an. Diese Informationen sollen von dem IWG ausgeschlossen sein (ebd.). Dies sei gemäß Schoch jedoch in keiner Weise zulässig, da eine öffentliche Stelle immer im Rahmen ihrer öffentlichen Aufgaben zu agieren habe, daher auch das Gutachten in Erfüllung der Aufgabenzuständigkeit angefertigt worden sei und somit vom IWG erfasst werde (Schoch, NVwZ 2006, S. 876).

6.1.2.2 Weiterverwendung

Eine Weiterverwendung ist gemäß § 2 Nr. 3 IWG jede Nutzung der Information, die über die Erfüllung öffentlicher Aufgaben hinausgeht und auf die Erzielung eines Entgelts ge-

richtet ist. Gemäß § 2 Nr. 3 2 IWG stellt die reine intellektuelle Wahrnehmung einer Information und Verwertung des dadurch erlangten Wissens regelmäßig keine Verwertung dar. Der Schwerpunkt der Weiterverwendung liegt bei der Auswirkung des Informationsangebotes auf den Binnenmarkt. Dabei handelt es sich einerseits um Dienste gegen Entgelt, aber auch nichtkommerzielle Informationsdienste können eine Auswirkung auf den Markt haben und fallen somit in den Anwendungsbereich des IWG (Altmeppen/-Kahlen, NVwZ 2006, S. 501).

6.1.2.3 Gleichbehandlung und Transparenz

Öffentliche Stellen müssen die Informationen gemäß §§ 3, 4 IWG »nicht diskriminierend, transparent, zeitnah, ohne überhöhte Entgelte und nicht exklusiv« (Sydow, NVwZ 2008, S. 484) zur Verfügung stellen. Die Gleichbehandlung bezieht sich auf die grundsätzliche Weiterverwendung wie auch auf die Bedingungen, unter denen die Weiterverwendung gestattet wurde (ebd.). Das heißt, dass kein Weiterverwender schlechter gestellt werden darf, sobald die Information einmal freigegeben wurde. Gemäß § 3 IV IWG ist die Vergabe von Exklusivrechten an der Weiterverwendung von Informationen nicht gestattet, außer der Dienst oder das Angebot sind von großem öffentlichem Interesse und ein kommerzieller Verleger wäre ohne Exklusivrechte nicht interessiert (Altmeppen/Kahlen, NVwZ 2006, S. 501). Dies trifft z. B. auf die Vereinbarung des Bundes mit dem Bundesanzeigerverlag über die Veröffentlichung von Gesetzen im Bundesanzeiger zu (Schoch, NVwZ 2006, S. 876). Ebenso ist die Gleichbehandlung auf die Gebührenerhebung anzuwenden. Gem. § 4 III IWG stellt die Obergrenze von Gebühren im Falle einer Gebührenvereinbarung für die Bereitstellung von Informationen und Gestattung ihrer Weiterverwendung die Summe der Kosten ihrer Erfassung, Erstellung, Reproduktion und Verbreitung zuzüglich einer angemessenen Gewinnspanne dar. Die Behörde kann geringere Gebühren verlangen, wenn der Nutzer für das erstellte Mehrwertprodukt kein Entgelt fordert (Altmeppen/Kahlen, NVwZ 2006, S. 502). Gemäß § 4 IV 1 IWG sind Nutzungsbestimmungen und eventuelle Entgelte von der Behörde im Voraus festzulegen und wenn möglich elektronisch zu veröffentlichen.

6.1.2.4 Verfahren

Privatrechtliche Vereinbarungen Die Weiterverwendung kann durch die öffentliche Stelle gemäß § 4 II 2 und 3 IWG auch durch einen privatrechtlichen Vertrag mit enthaltenen Nutzungsbedingungen gewährt werden, die verhältnismäßig sein müssen, nicht zu Wettbewerbsverzerrungen führen und die Möglichkeit der Weiterverwendung nicht unnötig einschränken dürfen (Altmeppen/Kahlen, NVwZ 2006, S. 502).

Verweigerung der Weiterverwendung Gemäß § 4 V IWG muss die Behörde bei einer Verweigerung der Weiterverwendung eine Begründung liefern und auf Rechtsschutz-

möglichkeiten hinweisen. Ist der Grund ein entgegenstehendes Urheberrecht, muss der Rechteinhaber genannt werden, soweit dies zulässig ist.

6.1.3 Zugang nach dem Umweltinformationsgesetz

6.1.3.1 Anwendungsbereich

Auskunftspflichtig sind alle informationspflichtigen Stellen. Das sind alle Stellen, die in der Trägerschaft des Bundes stehen (Reidt/Schiller in: Landmann/Rohmer, UIG § 1, Rn. 19). Davon ausgenommen sind oberste Bundesbehörden und Gerichte des Bundes gemäß § 2 I Nr. 1 a und b UIG. Inhaltlich kann an den Behördenbegriff angeknüpft werden (Reidt/Schiller in: Landmann/Rohmer, UIG § 2, Rn. 4). Durch den Verweis des § 18 a BerlIFG gilt das UIG auch für das Land Berlin, mit Ausnahme der §§ 11 bis 14 UIG, die hauptsächlich kosten- und bundesspezifische Normen enthalten.

6.1.3.2 Umweltinformationen

Die Art der Speicherung der Umweltinformationen ist unwichtig, sie müssen nur bei der Behörde abgelegt worden sein (Rn. 32). Der Begriff Umweltinformation ist sehr weit auszulegen (Rn. 31), sodass alles, was sich unter die Kategorien Zustand von Umweltbestandteilen, Umweltfaktoren, Maßnahmen oder Tätigkeiten mit Umweltbezug, Berichte zur Umsetzung des Umweltrechts, wirtschaftliche Analysen und Zustand der menschlichen Gesundheit und Sicherheit, Lebensbedingungen der Menschen sowie Kulturstätten und Bauwerke subsumieren lässt, als Umweltinformation gilt. Es besteht, wie auch bei den anderen Gesetzen, keine Pflicht zur Informationsbeschaffung.

6.1.3.3 Zugang und Weiterverwendung

Das UIG ist ein Informationszugangsgesetz speziell für Umweltdaten und soll den Umweltschutz verbessern. Es ist aber von der Art und Funktionsweise dem BerlIFG sehr ähnlich. Genau wie im BerlIFG ist jede natürliche und juristische Person des Privatrechts anspruchsberechtigt, ein besonderes Interesse muss nicht geltend gemacht werden und der Begriff des Anspruchstellers ist weit auszulegen (Reidt/Schiller in: Landmann/Rohmer, UIG § 3, Rn. 4-5). Der Zugang erfolgt durch Akteneinsicht oder Aktenauskunft, Kopien, z. B. Fotokopien, dürfen angefertigt werden (Rn. 13). Der Anspruchsteller hat ein gewisses Wahlrecht, wie der Zugang erfolgen soll, dieser kann aber bei Erfüllungsschwierigkeiten aufseiten der Behörde abgelehnt werden (Rn. 15-21). Wenn andere Zugangsrechte für die Umweltinformationen bestehen, so bleiben diese gemäß § 3 I 2 UIG parallel neben denen des UIG bestehen (Rn. 29). Die Weiterverwendung erfolgt wie bei dem BerlIFG nach dem IWG (Altmeppen/Kahlen, NVwZ 2006, S. 500)

6.1.3.4 Ablehnungsgründe

Abgelehnt werden kann ein Antrag gemäß §§ 8 und 9 UIG, wenn z. B. personenbezogene Daten, Rechte Dritter oder Betriebs- und Geschäftsgeheimnisse betroffen sind und wenn öffentliche Belange geschützt werden müssen oder Urheberrechte betroffen sind. Wenn das öffentliche Interesse an den Umweltdaten aber das Geheimhaltungsinteresse Dritter überwiegt, kann der Zugang zu den Umweltinformationen auch ohne die Zustimmung betroffener Dritter erfolgen (Schoch, NJW 2009, S. 2992).

6.1.4 Zugang nach dem Verbraucherinformationsgesetz

Wie das BerlIFG und das UIG ist auch das VIG ein Informationszugangsgesetz, speziell für gesundheitsbezogene Verbraucherinformationen aus dem Lebensmittelbereich (Domeier/Werner in: Leupold/Glossner, Teil 8, Rn. 219). Es funktioniert nach genau dem gleichen Muster. Das Gesetz gewährt dem Antragsteller Zugang zu Verbraucherinformationen, es muss dabei kein besonderes Interesse geltend gemacht werden. Wie auch in den anderen Informationszugangsgesetzen findet eine Abwägung zwischen öffentlichem Interesse und den Datenschutzrechten Dritter statt. Das öffentliche Interesse kann das Datenschutzinteresse Dritter überwiegen, wenn ein Verstoß gegen das Futter- und Lebensmittelrecht vorliegt (Schoch, NJW 2009, S. 2992). Teilweise wird das VIG als »symbolische Gesetzgebung« angesehen, die keinerlei Neuerungen gegenüber dem IFG und somit auch dem BerlIFG zu bieten habe (Schoch, NVwZ 2006, S. 877). Alle Verbraucherinformationen sind für den Antragsteller auch über das BerlIFG zu bekommen.

6.1.5 Geodatenzugangsgesetz Berlin

Mit dem GeoZG Bln wurde die EU-Richtlinie 2007/2/EG in ein Landesgesetz im Land Berlin umgesetzt. Es soll dazu dienen, im Land Berlin eine Geodateninfrastruktur aufzubauen. Das GeoZG Bln versteht den Begriff »Zugang« im Sinne eines Zugangs zu der Geodateninfrastruktur und dem Geoportal, von wo aus man online interaktiv die Dienste durchsuchen und nutzen kann. Nicht gemeint ist ein Zugang zu den Rohdaten im Sinne des IWG und des BerlIFG. Daher wurde den Behörden gemäß § 13 II GeoZG Bln auch eine Möglichkeit gegeben, die Weiterverwendung von Darstellungsdiensten zu unterbinden (Polenz, NVwZ 2010, S. 488) bzw. Geldleistungen für diese und sogar auch für die Nutzung zu fordern. Auch Nutzungs- und Lizenzbedingungen dürfen gemäß § 13 I GeoZG Bln gestellt werden. In diesem Zusammenhang sind auch die Geolizenzen¹ als Verträge zwischen einer begrenzten Anzahl von Nutzern und Arbeitsplätzen und der Behörde zu sehen, in denen Preis und Leistung festgelegt werden. Das GeoZG Bln ist damit keine geeignete Anspruchsgrundlage, um Basisdaten für einen Open Data-Dienst zu erlangen. Ebenso wenig sind Geolizenzen für OGD zu gebrauchen.

¹<https://www.geolizenz.org/index/lizenzen.php>.

Es wäre aber möglich, die Rohdaten über das BerlIFG zu erlangen und gemäß dem IWG weiterzuverwenden.

6.1.6 Ausblick auf das eGovernment- und Organisationsgesetz in Berlin

Im Rahmen der Schaffung eines eGovernment-Gesetzes in Berlin bietet sich eine weitere Möglichkeit, die informelle Verwaltung zu regeln. Nach dem BerlIFG hat der Antragsteller bisher das Recht, ohne Begründung die Herausgabe von bestimmten Daten von den Behörden zu verlangen. Diese prüfen aufgrund der Anfrage, ob eine Herausgabe möglich ist. Der Impuls geht vom Antragsteller aus. Nun sollen die Behörden proaktiv Datensätze veröffentlichen. Im aktuellen § 8 I Ref-E EGovG Bln werden die Behörden verpflichtet, die Datenbestände im Land Berlin bereitzustellen. Dies soll gemäß § 8 II Ref-E EGovG Bln über ein zentrales Datenportal geschehen. Durch diesen Paragraphen könnten die Open Data-Grundsätze verpflichtend für die Behörden werden. Ein wichtiger Bestandteil von eGovernment ist das informative Verwaltungshandeln. Damit ist die Kommunikation zwischen dem Staat und dem Bürger und der Wirtschaft gemeint (Ernst in: Hoeren/Sieber, Teil, 28.1, Rn. 47). Durch die elektronische Verwaltung der großen Datenmengen der Behörden steigt die Effizienz und Effektivität der Datendistribution (Ernst in: Hoeren/Sieber, Teil 28.1, Rn. 48). Unabhängig von den Einschränkungen der dem Veröffentlichungswunsch entgegenstehenden Datenschutzbestimmungen ist dies eine Chance für eine Förderung einer Informationsgesellschaft in Berlin. Allerdings ist zu bedenken, dass Ansprüche gegen die Behörden entstehen können, wenn diese der Bereitstellungspflicht nicht nachkommen. Weiterhin könnten sich Behörden mit Mehrausgaben konfrontiert sehen. Mehr Personal würde benötigt werden, um technische Abläufe zu regeln und eine Abwägung zwischen öffentlichem Interesse und Datenschutz vorzunehmen. Letztendlich ist fraglich, ob der § 8 Ref-E EGovG Bln nicht eher im BerlIFG statuiert werden sollte, wie es im § 11 Bremer Informationsfreiheitsgesetz der Fall ist.

6.2 Behördliche Veröffentlichung und Lizenzierung von PSI

6.2.1 Datenschutzrechtliche Vorgaben für die Veröffentlichung von PSI

Die Bereitstellung von PSI erfolgt zumeist in zwei Schritten. Zuerst muss die Behörde die zu veröffentlichenden Datensätze auf Urheberrecht und Datenschutz prüfen. Danach kann die Datei lizenziert werden. Daher muss zuerst auf die Vorstufe zur Veröffentlichung eingegangen werden. Danach wird ein geeignetes Lizenzmodell für die PSI erörtert.

Im ersten Teil des Berichts wurde erörtert, wie eine private Person Daten von Behörden erlangen kann. Grenzen sind dem Begehren hauptsächlich durch den Schutz personenbezogener Daten sowie von Betriebs- und Geschäftsgeheimnissen gesetzt. Wenn die Behörde die Daten nun selbst veröffentlichen möchte, kann nichts anderes gelten. Die

Behörde muss genauso, als ob eine private Person über die Informationsfreiheitsgesetze Auskunft verlangen würde, prüfen, ob die begehrte Information freigegeben werden kann oder gegebenenfalls anonymisiert werden muss oder nicht veröffentlicht werden darf. Die Behörden brauchen flexible Instrumente, anhand deren schnell und einfach festgestellt werden kann, ob ein Datensatz zur Freigabe geeignet ist. Der folgende zweistufige Test orientiert sich an den Datenschutzparagrafen des BerlIFG, des BerlDSG sowie des UIG.

6.2.1.1 Stufe 1: Identifizierung des Rechteinhabers / kein Verstoß gegen Urheberrechte Dritter

Geklärt muss werden, wer der Urheberrechtsinhaber ist. Ist dies die Behörde, dann kann veröffentlicht werden. Ist dies eine dritte Person (Dritter im Sinne von nichtinvolviert in die Beziehung Behörde-Nutzer), muss diese um Erlaubnis gefragt und sich gegebenenfalls geeinigt werden. Wenn die dritte Person die Zustimmung verweigert, darf die Datei nicht veröffentlicht werden.

6.2.1.2 Stufe 2: Prüfung der Personenbezogenheit

Schwerer festzustellen ist dagegen ein eventueller Personenbezug von Daten, selbst wenn die Behörde jeden Datensatz eingehend prüft. In Zukunft könnte es Möglichkeiten geben, einen heute noch unkritischen Datensatz personenbezogen zu machen. Einen vollständig datenschutzrechtlich unbedenklichen Datensatz wird es daher so nicht geben. Momentan Unkritisches sollte man trotzdem aufgrund des Informationsinteresses des Bürgers veröffentlichen. Wenn die Behörde bei einem Datensatz keinen Personenbezug herstellen kann, muss das für eine Veröffentlichung ausreichen. Wenn jedoch ein Personenbezug festgestellt werden kann oder durch die Veröffentlichung ein Betriebsgeheimnis oder die öffentliche Sicherheit gefährdet würde, muss die Datei unter Verschluss gehalten werden. Rechtlich sollte eine Absicherung erfolgen, indem wie in der englischen Open Government License² darauf verwiesen wird, dass die Nutzung von eventuell personenbezogenen Daten in der Information untersagt ist. Weiterhin sollte man sich das Recht vorbehalten, veröffentlichte Daten bei Bekanntwerden eines Personenbezugs sofort zurückzurufen.

6.2.1.3 Schadensersatzansprüche gegen die Behörde

Sollte die Behörde trotzdem einen Datensatz veröffentlichen, der einen Personenbezug möglich macht, ohne dass dies von den Ausnahmetatbeständen gerechtfertigt ist, kann der Betroffene gemäß § 18 I BerlDSG klagen. Es sind somit nur natürliche Personen anspruchsberechtigt. Der Betroffene müsste allerdings geltend machen können, dass ihm ein Vermögensschaden entstanden ist. Ansonsten bleibt die Möglichkeit, gemäß § 18 I 3

²<http://www.nationalarchives.gov.uk/doc/open-government-licence>.

BerlDSG auch Schäden, die keine Vermögensschäden sind (z. B. Ehrverletzungen), geltend zu machen. Allerdings kommt hier das Erfordernis der Schwere hinzu. Es muss durch die Veröffentlichung der personenbezogenen Daten ein Bild in der Öffentlichkeit geprägt werden, dass die Handlungs- und Entscheidungsfähigkeit des Betroffenen erheblich einschränkt (NJW-RR, 2005, S. 1565). Dieses wird restriktiv gehandhabt.

6.2.2 Daten- und Datenbanklizenzierung

Bei der ODP wird, sobald sich die Regierungen auf eine Freigabe ihrer PSI geeinigt haben, die Lizenzierung eine tragende Rolle spielen, wie sich anhand der jetzt schon bestehenden europäischen ODP z. B. in Paris, Wien, Barcelona und England aufzeigen lässt. Auch der Lizenzierung von Datenbanken kommt eine größere Bedeutung zu, da die Informationen immer öfter in Form von Datenbanken veröffentlicht werden. Der Bürger soll in Zukunft nicht aktiv um sein Recht auf Informationsfreiheit vor der Behörde kämpfen müssen, sondern ein breites Angebot an Datensätzen zur Verfügung gestellt bekommen. Damit wird ein wichtiger Schritt in Richtung Informationsgesellschaft gemacht, die zu fördern eines der wichtigsten Politikziele der EU ist (Schoch, NVwZ 2006, S. 872). Bei der Lizenzierung gehen die Lizenzgeber aber durchaus unterschiedliche Wege. Es gilt daher, ein Modell zu finden, das dem Berliner Szenario am besten entspricht.

6.2.2.1 Separater Schutz der Datenbank

Generell wird, wie auch in der europäischen Datenbankrichtlinie 96/9/EG, zwischen Inhalt und Datenbank unterschieden. Unter Datenbank wird jede Sammlung von Daten verstanden, unter Daten jegliche Information (Harte-Bavendamm in: Kilian/Heussen, Teil 5, Rechtsschutz von Datenbanken, Rn. 2). Mit dem Inkrafttreten der Datenbankrichtlinie werden auch Investitionen in reine alphabetische Aufzählungen geschützt, die nicht als geistige Schöpfung aufgrund von individueller Auswahl, Wertung, Rangfolge oder Aktualität gelten (v. Lewinski in: Loewenheim, § 54, Rn. 36). Fraglich ist, wie nun die bloße Art und Weise, wie die Daten zusammengestellt sind, geschützt werden soll. Paris und England haben die Datenbank getrennt lizenziert bzw. freigegeben^{3 4}, während Barcelona und Wien auf solch eine getrennte Lizenzierung verzichtet und nur die Datensätze freigegeben haben^{5 6}.

Open Database License (ODbL) Paris hat die Datenbank unter eine modifizierte Version der Open Database License gestellt. Diese funktioniert nach dem Prinzip der Copyleft-Lizenzen. Das heißt, der Lizenzgeber verzichtet weitestgehend auf seine Urheberrechte und berechtigt den Lizenznehmer, die Datenbank beliebig weiterzuverwenden,

³http://opendata.paris.fr/opendata/jsp/site/Portal.jsp?page_id=10.

⁴<http://www.nationalarchives.gov.uk/doc/open-government-licence>.

⁵<http://w20.bcn.cat/opendata/Usos.aspx?lang=CAT>.

⁶<http://data.wien.gv.at/nutzungsbedingungen/>.

solange er das Ergebnis unter derselben oder einer kompatiblen Lizenz veröffentlicht. Wer sein Produkt, das auf den freien Inhalten aufbaut, nicht unter eine Copyleft-Lizenz stellen möchte, kann mit dem Lizenzgeber, in diesem Fall der Stadt Paris, einen individuellen Vertrag aushandeln, sogenanntes Dual Licensing.

Attributionslizenz In England wird die Datenbank im Rahmen der »Open Government License for Public Sector Information« unter eine »Attribution License« gestellt. Der Nutzer muss nur die ursprüngliche Herkunft der Datenbank nennen und darf sie sonst beliebig weiterverwenden, auch kommerziell und unter eigener Lizenz. Ein Beispiel für eine solche Datenbanklizenz ist die ODC Attribution License.⁷

Keine Freigabe der bloßen Datenbank Die ODP in Barcelona und Wien erwähnen in ihren Nutzungsbedingungen die reine Datenbankstruktur nicht. Mit der Lizenzierung der Datensätze unter CC-Lizenzen ist den Datenbanken insoweit Rechnung getragen, als der Entnahme und Weiterverwendung der Datensätze das speziell für Datenbanken geschaffene Schutzrecht (sogenanntes Sui-generis-Schutzrecht) nicht entgegensteht. Da aber keine eigene Lizenz für die Datenbank verwendet wird, kommt für diese das Sui-generis-Schutzrecht zum Tragen. Somit ist die reine Datenbankstruktur in diesen Fällen nicht freigegeben und darf von den Nutzern der ODP nicht verwendet werden.

Fazit Gerade bei PSI ist die Relevanz der Datenbank gering, es sollen eben gerade »raw data« zur Verfügung gestellt werden, um die Kosten für die Behörden nicht unnötig in die Höhe zu treiben und die Schöpfung des wirtschaftlichen oder wissenschaftlichen Potenzials dem Bürger bzw. der Wirtschaft zu überlassen. Dieses Prinzip gilt nicht nur für einzelne Informationen oder Datensätze, sondern ebenfalls für die Datenbanken. Auch die Datenbank wird »raw« zur Verfügung gestellt. Damit sind alphabetische oder thematische Zusammenstellungen zu erwarten, eventuell sortiert nach Aufrufen. Die Wartungsarbeiten halten sich in Grenzen und könnten eventuell von einer Software übernommen werden. Somit fällt der Datenbank ein relativ geringer Schutz zu, da in ihr kaum eine geistige Schöpfung gesehen werden kann, und auch die intensive Wartung, die als Hauptargument für den Schutz reiner Aufzählungen von Informationen genannt wird (v. Lewinski in: Loewenheim, § 54, Rn. 36), liegt hier nicht vor. Maßgeblich sind bei Rohdaten die Dateninhalte. Es geht gerade nicht um eine Aufbereitung. Bei dem Datenbankschutz ist aber die Aufbereitung einer Datenbank das einzige Merkmal, das einen Schutz rechtfertigen kann. Wenn der Nutzer also mehrere höchstens thematisch geordnete Datensätze weiterverwenden will, die jeweils eine Weiterverwendung erlauben, muss der Nutzer darüber hinaus trotzdem den Schutz der primitiven, vielleicht sogar willkürlichen Anordnung beachten. Daher leuchtet es nicht ein, wieso die Datenbank eine Lizenz wie die ODbL erhalten soll, die fordert, alle aus der Datenbank entstehenden

⁷<http://www.opendatacommons.org/licenses/by/1-0>.

Produkte unter dieselbe Lizenz zu stellen. Dies mag den Vorteil haben, dass die Zusammenstellung der Datenbank in allen weiterverwendeten Variationen frei zugänglich sein mag. Der Nachteil ist, dass die großzügig gewährten Nutzungsrechte der ODbL praktisch wertlos sind, wenn die in der Datenbank enthaltenen Daten urheberrechtlich geschützt sind. Die Datenbanklizenz suggeriert dann dem Durchschnittsbenutzer Freiheit, die in Wirklichkeit nicht gegeben ist. Dies kann beim Nutzer zu Irrtümern und so zur Rechtsunsicherheit im Dschungel der Lizenzen führen. Darüber hinaus fördert die ODbL das Ziel der wirtschaftlichen Potenzialnutzung der Richtlinie 2003/98/EG nur schwach, da eine wirtschaftliche Nutzung mit der ODbL praktisch sinnlos ist, wenn das kommerziell angebotene Produkt ebenso kostenlos nutzbar sein muss. Andererseits könnten auf diese Weise Anreize für Unternehmen geschaffen werden, die Datenbankstruktur unter einer anderen, möglicherweise kostenpflichtigen Lizenz zu erwerben, wenn die Behörde von der Freigabe der Datenbankstruktur bei einer wirtschaftlichen Nutzung profitieren will.

Die Attributionslizenz mit Namensnennung als Bedingung hat den Vorteil, dass die natürlichen oder juristischen Personen hier freier in der u. a. auch wirtschaftlichen Verwertungsauswahl sind und mögliche Verwirrungen für den Laien durch die Unterscheidung von Datensatz und Datenbank ausgeschlossen sind. Es ist aber fraglich, ob überhaupt ein wirtschaftlicher Wert in der reinen Datenbankstruktur von OGD gesehen werden kann. Es ist auch nicht der Sinn der ODP, bloße Datenbankstrukturen zur Verfügung zu stellen. Vielmehr sollen behördliche Inhalte und Informationen freigegeben werden. Weiterhin ist auch nicht ersichtlich, dass die Datenbankstruktur ohne angeschlossene Inhalte von Interesse für die Nutzer wäre. Somit spricht nichts gegen eine Außerklassung der Datenbanklizenz. Die Berücksichtigung, die die Datenbank im Legal Code der CC-Lizenzen erfährt, sollte in der Praxis ausreichen. Wenn allerdings dem Bürger die maximale Rechtssicherheit und Verwertungsfreiheit verschafft werden soll, sollte die Datenbank unter die ODC-Attributionslizenz gestellt werden.

6.2.2.2 Lizenzierung der Daten

Bei der Lizenzierung der Daten sollte beachtet werden, dass den Nutzern auf verständliche Art und Weise nahegebracht wird, wie sie die Daten nutzen können. Barcelona und Wien nutzen eine Lizenz aus dem Angebot der Creative Commons-Lizenzen, England nutzt eine selbst entwickelte Lizenz, die aber kompatibel mit den CC-Lizenzen ist⁸ und praktisch die gleichen Anforderungen stellt und Freiheiten gewährt wie die CC BY 3.0.⁹

CC BY 3.0 Diese Lizenz gewährt dem Lizenznehmer, die Informationen beliebig weiterzuverwenden, auch kommerziell und unter eigener Lizenz, solange als Quelle der originäre Urheber genannt wird.¹⁰

⁸<http://data.gov.uk/terms-conditions>.

⁹http://www.epsiplus.net/news/news/uk_opendata_licences_based_on_cc.

¹⁰<http://creativecommons.org/licenses/by/3.0/legalcode>.

CC BY ND 3.0 Diese Lizenz soll in Barcelona auf die urheberrechtsgeschützten Daten angewendet werden. Man darf diese Daten weiterverteilen, auch kommerziell, allerdings dürfen sie nicht verändert werden. Auch die Quelle muss genannt werden.¹¹ Diese Lizenz eignet sich für die Veröffentlichung von Material, das durch Rechte Dritter belastet ist.

CC0 Eine weitere Lizenz ist die CC0-Lizenz, die dem Nutzer ähnliche Nutzungsmöglichkeiten gewährt wie die CC BY 3.0. Das Erfordernis der Namensnennung fällt hier jedoch weg. Praktisch an dieser Lizenz ist, dass jede Art von Inhalt oder Werk eingeschlossen ist, also auch die Datenbank.¹² Dazu muss der Rechteinhaber auf alle Rechte verzichten, soweit das Gesetz dies zulässt. In den USA werden die meisten PSI unter eine identische Public Domain-Lizenz gestellt.

Das IWG (siehe 6.1.2) gestattet das Recht, aus Daten, die durch Informationsfreiheitsgesetze erlangt wurden, Informationsmehrwertprodukte zu gewinnen. Diese dürfen dann auch kommerziell genutzt werden. Diese Rechte entsprechen den Rechten, die auch die CC BY 3.0 gewährt, nur wird im IWG nicht die Namensnennung gefordert. Somit gewährt das IWG Rechte, die denen der CC0 identisch sind. Für die von der Behörde veröffentlichten Daten gelten somit, soweit keine Ausnahmetatbestände des IFG greifen, die Bestimmungen des IWG, die denen der CC0 entsprechen, jedoch nur in Bezug auf die Daten. Für solche Daten bliebe eigentlich keine andere Wahl als eine CC0-Lizenz, denn der Antragsteller darf diese durch Zugangsgesetze erlangten Daten unabhängig von der vergebenen Lizenz verwenden. Dies spricht für eine Anwendung der CC0-Lizenz.

Open Government License for Public Sector Information England hat eine Lizenz entwickelt, die die der CC BY 3.0 sehr ähnlich und mit dieser auch in der neuesten Version kompatibel ist.¹³ Der Lizenzgeber gewährt dem Nutzer der Information, der mit der Nutzung automatisch in die Lizenzbedingungen einwilligt, eine weltweite, gebührenfreie, unbefristete und nicht exklusive Lizenz für die Informationen wie auch die Datenbanken. Solange der Nutzer die Quelle der Information und der Datenbank bekannt macht und einen Hinweis auf den Anbieter hinzufügt, soll es dem Nutzer erlaubt sein, die Information und die Datenbank

- zu kopieren, zu veröffentlichen, zu verteilen und zu übermitteln,
- zu bearbeiten und
- kommerziell zu nutzen, indem sie mit anderen Informationen oder Datenbanken verbunden wird oder in ein schon vorhandenes eigenes Produkt oder in eine eigene Anwendung eingebaut wird.

Der Nutzer soll die Information oder die Datenbank so weiterverwenden, dass

¹¹<http://creativecommons.org/licenses/by-nd/3.0/legalcode>.

¹²<http://creativecommons.org/publicdomain/zero/1.0/legalcode>.

¹³http://www.epsipius.net/news/news/uk_opendata_licences_based_on_cc.

- keine Verwechslungsgefahr mit einem staatlichen Angebot entstehen kann,
- sie nicht falsch dargestellt wird und irreführend ist,
- sie nicht gegen die Bestimmungen der EU-Richtlinie 2003/98/EG verstößt.

Weiterhin werden noch Einschränkungen bezüglich des Datenschutzes und des Schutzes von Rechten Dritter geltend gemacht.¹⁴

Geolizenzen Die sogenannten Geolizenzen wurden entwickelt, um den Behörden mit Geodatenbestand den Vertrieb ihrer Geodatenmehrwertprodukte (z. B. Geodatendarstellungsdienste), einfacher zu machen. Diese Lizenzen sind allgemeine Vorlagen speziell für Geodaten, die sich an den CC-Lizenzen orientieren. Es gibt Lizenzen, die die kommerzielle Nutzung erlauben oder verbieten sowie die Weiterverarbeitung und die Nutzung in öffentlichen Netzwerken regeln.¹⁵ Der große Unterschied der Geolizenzen zu den CC-Lizenzen besteht in der Möglichkeit, verschiedene Felder auszufüllen wie Produktname, Lizenzgeber, Lizenznehmer, Preis und Anzahl der Arbeitsplätze (für die die Lizenz gelten soll). Es ist nicht zwingend, dass die Behörde ein Entgelt für die Nutzung der Daten durch die Lizenz in Rechnung stellt, aber gemäß § 13 I GeoZG Bln darf sie Geldleistungen dafür verlangen. Der Lizenzgegenstand wird auch nicht wie bei OGD die Rohdatei sein, sondern schon aufbereitete Daten, inklusive Serviceleistung für einen bestimmten Zeitraum. Die Behörde kann den Vertrag auch kündigen. Die Lizenzen sind als ein Mittel anzusehen, das den Behörden länderübergreifend das eventuell auch kostenpflichtige Lizenzieren ihrer Daten vereinfachen soll. Sie sind somit nicht geeignet, Open Data-Inhalte zu lizenzieren.

Fazit Das katalanische Modell (Barcelona) scheint zwischen Urhebern und Nutzern zu vermitteln, indem die urheberrechtsgeschützten Werke zwar freigegeben werden, aber nicht bearbeitet werden dürfen. Auch Behörden können Urheberrechte an den erhobenen und aufbereiteten Daten haben. Jedoch kann die Behörde kein Eigenleben entfalten und muss stets im öffentlichen Auftrag handeln (Schoch, NvWZ 2006, S. 875). Daraus könnte man folgern, dass eine Public Domain-Lizenz die beste Lösung zumindest für jene Daten wäre, an denen Dritte keine Urheberrechte geltend machen können. Dies entspräche auch der Transformation des IWG in eine Lizenz. Auf diesem Wege würden die PSI zu Allgemeingut. Dies wäre wünschenswert und ist vor allem auch geltendes Recht, jedoch scheinen sich viele Datenanbieter eine Namensnennung zu wünschen. Daher sollten diese ihre Daten zumindest unter der Bedingung der Namensnennung freigeben, wie es z. B. die CC BY 3.0-Lizenz statuiert. Die CC BY 3.0-Lizenz ist eine sogenannte Attributionslizenz und ist damit auch mit der ODC-Attributionslizenz für Datenbanken kompatibel. Diese zwei Lizenzen entsprechen praktisch der »Open Government License for Public Sector Information« des Vereinigten Königreiches Großbritannien und Nordirland. Berlin könnte sich ebenfalls eine eigene Lizenz nach englischem Vorbild geben, jedoch geht

¹⁴<http://www.nationalarchives.gov.uk/doc/open-government-licence>.

¹⁵<https://www.geolizenz.org/index/lizenzen.php>.

der europäische Trend bei OGD in Richtung CC-Attributionslizenz. Auch wenn die OGL die gleichen Nutzungsbedingungen aufstellt, vereinfacht die Nutzung einer identischen Lizenz zumindest für die Daten die Schaffung eines übersichtlichen, homogenen und nutzerfreundlichen europäischen und auch weltweiten Informationsmarktes für OGD. Daher ist der CC BY 3.0-Lizenz für die Daten der Vorzug zu geben. In Kombination mit der ODC-Lizenz für die Datenbank werden die PSI so zukunftsicher freigegeben.

6.2.3 Zusätzliche Nutzungsbedingungen

Darüber hinaus können noch zusätzliche Nutzungsbedingungen hinzugefügt werden, wenn der Lizenzgeber spezielle Sachverhalte abdecken möchte, die in der abstrakten Lizenz nicht behandelt wurden. Zum Beispiel stellt Barcelona bei beiden Lizenzarten die Bedingung, dass die Daten nicht verfälscht dargestellt werden oder ihr Informationsgehalt verzerrt wird. Außerdem soll die Information mit einem Hinweis auf ihre letzte Aktualisierung versehen werden.¹⁶ In Wien dürfen die Informationen nur genutzt werden, wenn sie nicht z. B. rassistisch oder sexistisch verwendet werden.¹⁷ In England verbietet die Lizenz jegliche Nutzung von personenbezogenen Informationen. Allerdings könnten ausschweifende angehängte Nutzungsbedingungen Inhalten der Lizenz entgegenstehen. Dadurch würde wieder der Nutzer verwirrt. Daher sollte man die Anhänge an die Lizenzen so knapp wie möglich halten. Die Nutzungsbedingungen der Berliner ODP sind eine Art Mischung aus allen vorhergenannten Nutzungsbedingungen. Sie sind so weit akzeptabel, einige Änderungen wären aber sinnvoll. Zuerst wäre es von Vorteil, wenn der Nutzer wüsste, was ein Vorschlag oder aber ein verbindlicher Vertragsgegenstand ist. Eine Unterteilung in generelle Hinweise und Bitten und in konkrete Erweiterungen der Lizenz würde sich daher anbieten. Man sollte zudem vermeiden, undeutliche und alles umfassende Begriffe zu verwenden wie »geltendes Recht«, »Respekt vor Privatheit« und »alle Zwecke«. Diese Sachverhalte sind klar und deutlich schon in der eigentlichen Lizenz geregelt. Ebenso ist die Untersagung der Verbreitung von beleidigendem und unrechtmäßigem Material genau wie allgemein »andere Handlungen« schon in den Gesetzen wie StGB und BGB geregelt. Der Ausschluss einer Aktualisierungspflicht könnte gegen den § 8 I Ref-E EGovG Bln verstoßen (siehe 6.1.6). Zuletzt könnte man noch der Bedingung, dass auf Berliner Daten basierende Apps von der Stadt Berlin beworben werden dürfen, hinzufügen, dass dies nur bei einer bereits vom Autor vorgenommenen Veröffentlichung greifen soll.

¹⁶http://w20.bcn.cat/opendata/Usos.aspx?lang=CAT#cond_us.

¹⁷<http://data.wien.gv.at/nutzungsbedingungen>.

6.3 Datenschutz der Nutzer der Open Data-Plattform

6.3.1 Anwendbarkeit des TMG auf die Open Data Internet-Plattform

Fraglich ist, welches Datenschutzgesetz für eine ODP anzuwenden ist. Nach § 1 I 1 TMG sind Telemediendienste alle elektronischen Informations- und Kommunikationsdienste, soweit sie nicht Telekommunikationsdienste nach § 3 Nr. 24 TKG, telekommunikationsgestützte Dienste nach § 3 Nr. 25 TKG oder Rundfunk nach § 2 RStV sind. Das heißt, die PSI-Plattform müsste Daten nicht nur übermitteln, sondern aufbereiten und Inhalte regeln (Schmitz in: Hoeren/Sieber, Teil 16.2, Rn. 57-59). Die Plattform stellt die Daten der öffentlichen Stellen nicht nur gebündelt zur Verfügung, sondern verknüpft sie mit anderen Informationen zu neuen Mehrwertinformationsprodukten. Ebenfalls hat der Gesetzgeber in der Gesetzesbegründung für das TMG als Beispiel für ein Telemedium einen Online-Informationssdienst von Verkehrs-, Wetter-, Umwelt- oder Börsendaten genannt (vgl. BT-Drs. 16/3078). Somit finden die Datenschutzvorschriften des TMG Anwendung auf die ODP.

6.3.2 Vertragsverhältnis zwischen Nutzer und der ODP

Der Nutzer hat die Möglichkeit, auf alle Daten ohne vorherige Registrierung zuzugreifen. Um die Kommentarfunktion zu nutzen, muss sich der Nutzer unter einem Pseudonym registrieren. Der Nutzer wird sich also in jedem Fall anonym auf der Plattform bewegen. Ein Nutzungsvertrag kommt zwar schon allein durch die Lizenzen, die vorgeben, unter welchen Bedingungen Daten und Datenbank zu nutzen sind, zustande. Dadurch werden aber in keiner Form personenbezogene Daten verarbeitet. Der durch die Lizenzen vereinbarte Nutzungsvertrag ist in Bezug auf den Datenschutz der Nutzer also unbeachtlich.

6.3.3 Personenbezogene Daten

Fraglich ist, ob überhaupt personenbezogene Daten entstehen. Die bei der Nutzung der ODP anfallenden Daten gelten dann als personenbezogen, wenn sie bei der Nutzung des Dienstes entstehen und den Bezug auf eine bestimmbare (natürliche) Person ermöglichen (Schmitz in: Hoeren/Sieber, Teil 16.2, Rn. 76-77). Der Informationsdienst wird zunächst die IP des Nutzers kennen. Hier fehlt es jedoch regelmäßig an einem Personenbezug, da die ODP kein Access Provider ist und so die IP zuordnen könnte (Schmitz in: Hoeren/Sieber, Teil 16.2, Rn. 84). Die gesetzten Cookies weisen auch keinen Personenbezug auf, weil sie nicht mit anderen Informationen über die bürgerliche Existenz des Nutzers in Verbindung gebracht werden können. Eventuell wird die Plattform ein anonymes Nutzungsprofil erstellen, wozu aber keine Einwilligung nötig ist (Rn. 224). Somit wird der Datenschutz auf Plattformnutzerseite ausreichend berücksichtigt.

6.3.4 Personenbezogene Daten der Datenanbieter

In mehreren Gerichtsentscheidungen war zu klären, ob der Name eines Mitarbeiters einer öffentlichen Verwaltung bei Veröffentlichungen der Behörde dem Datenschutz unterliegt. Dies umfasste auch die Kontaktdaten des Bearbeiters. Das Oberverwaltungsgericht Rheinland-Pfalz, Az. 2 A 10413/07, hat am 10. September 2007 entschieden, dass »im Interesse einer transparenten, bürgernahen öffentlichen Verwaltung der Dienstherr von Rechts wegen nicht gehindert ist, Namen, Funktionen und dienstliche Erreichbarkeit jedenfalls solcher Beamter, die mit Außenkontakten betraut sind, auch ohne deren Einverständnis im Internet bekannt zu geben«. Insbesondere sind Telefonnummer oder E-Mail-Adresse keine schützenswerten Bestandteile der Personalakte. Vielmehr dienen diese Informationen der sachgerechten Kommunikation mit den Bürgern. Da der Kläger Beschwerde einlegte, wurde die Sache nun vor dem Bundesverwaltungsgericht verhandelt. Es bestätigte das erstinstanzliche Urteil (Az. BVerwG 2 B 131.07 vom 12. März 2008) und bekräftigte die Position des Dienstherrn: Sofern eine juristische Person des öffentlichen Rechts befugt sei, ihre behördliche und organisatorische Struktur zu regeln, könne der Dienstherr ohne ausdrückliche Ermächtigungsgrundlage bestimmen, wie die Behörde nach außen auftrete. Wie er das im Einzelnen ausgestalte, liege in seinem Ermessen. Kein Bediensteter habe jedoch einen Anspruch darauf, anonym zu bleiben, es sei denn, legitime Interessen wie z. B. Sicherheitsbedenken stünden dem entgegen.

6.4 Zusammenfassung und Fazit

Als geeignete Anspruchsgrundlagen zur erstmaligen Freigabe von behördlichen Informationen kommen in Berlin gegenwärtig aus Perspektive der privaten natürlichen und juristischen Person vor allem das BerlIFG und durch seine Verweise in § 18 a BerlIFG für Umweltinformationen auch das UIG in Betracht. Der Anspruch auf die kommerzielle Weiterverwendung der einmal durch IFG oder UIG freigegebenen Information entsteht durch das IWG. Das Verbot der kommerziellen Weiterverwendung gemäß § 13 VII BerlIFG ist unbeachtlich. Haupthindernis für einen Zugang oder eine Weiterverwendung wird wohl die mögliche Verletzung von Betriebs- und Geschäftsgeheimnissen sein. Hier kann aber z. B. gerade bei Umweltangelegenheiten gut mit dem öffentlichen Informationsinteresse argumentiert werden. Das GeoZG Bln kommt als Anspruchsgrundlage nicht in Frage, da es den Aufbau einer eher staatlichen Geodateninfrastruktur mit eigenen Darstellungsdiensten fördern möchte, die im Endeffekt selbst Mehrwertinformationsprodukte sind. In gewisser Weise steht zumindest im Geodatenbereich das europäische Geodateninfrastrukturvorhaben in Konkurrenz zu der ODP. Es verpflichtet die Behörden, gerade speziell für dieses Geodatenportalvorhaben ihre Daten bereitzustellen, und begründet so für die Behörden eine Pflicht zum Aufbau einer Geodaten-ODP. Dass die Behörden ihre Darstellungsdienste nicht kostenfrei weiterverwendet sehen wollen, ist verständlich. Nur steht die behördliche Verpflichtung zum Aufbau eines Geodatenportals der wirtschaftlichen Schaffung von Mehrwertinformationsprodukten entgegen, es entsteht eine direkte

Konkurrenzsituation. Diese kann durchaus von der EU gewollt sein. Die Wirtschaft ist so angehalten, über einen reinen Darstellungsdienst hinaus Informationsprodukte über mehrere Ebenen zu verknüpfen.

Als Beispiel für ein Lizenzmodell für die behördlich bereitgestellten Informationen kann das Attributionsmodell empfohlen werden. Da Behörden im öffentlichen Auftrag handeln, sollte maximal eine Namensnennung als Bedingung für die Nutzung der Daten reichen. Die CC BY 3.0-Lizenz bietet sich an und wird für OGD z.B. schon in Wien und Barcelona genutzt, die vom Inhalt fast identische OGL in England. Wegen Unklarheiten des Verhältnisses von Datenbank und Daten sollte die eventuell veröffentlichte Datenbank eine eigene Lizenz bekommen. Eine Attributionslizenz wäre hier ebenfalls empfehlenswert, denn auf diesem Weg könnten eventuelle Kompatibilitätsprobleme bei einer Doppellizenzierung umgangen werden.

Der Nutzer wird sich auf der Plattform nur anonym bewegen. Dies ermöglicht maximalen Nutzerdatenschutz. Um dem Datenschutz bei der OGD-Veröffentlichung Genüge zu tun, sollte den Nutzungsbedingungen ein generelles Nutzungsverbot von personenbezogenen oder sicherheitsgefährdenden Daten hinzugefügt werden. Zusätzlich sollten die Behörden vor der Veröffentlichung die PSI auf Personenbezogenheit prüfen.

7

Pilotsystem

Aufbauend auf den Architekturempfehlungen aus Phase 1, wurde ein Prototyp für eine Berliner Open Data-Plattform mit ersten Pilotdaten entwickelt und mit der BerlinOnline GmbH auf <http://daten.berlin.de> pilotiert. Die Test- und Entwicklungsplattform wurde am 14. September 2011 freigeschaltet. Damit konnten frühzeitig die Machbarkeit und das Potenzial offener Daten in Berlin aufgezeigt werden. Bei der Bereitstellung und Weiterverarbeitung von öffentlichen Daten konnten wichtige praktische Erfahrungen gewonnen werden, die in den weiteren nutzer- und bedarfsorientierten Ausbau der Plattform einfließen können. Anschließend sollen das Berliner Open Data-Portal und das Berliner Datenregister in einer Übergangsphase weiterbetrieben werden, bis sie in einen Wirkbetrieb überführt werden können.

Gemäß den Empfehlungen aus dem Abschnitt »Architektur« wurde von FOKUS ein Datenregister und von BerlinOnline ein dazugehöriges Webportal realisiert.

Das Portal dient direkten und indirekten Datennutzern als einheitlicher Zugang zu offenen Daten in Berlin mit Funktionalitäten zur Suche (auf Basis von Kategorien und Schlagwörtern) und zur Navigation im Datenbestand. Das Datenportal wird näher in Abschnitt 7.1 beschrieben.

Das Datenregister (Abschnitt 7.2) dient Datenbereitstellern und Plattformbetreibern als Zugangspunkt zum Registrieren von neuen Datensätzen bzw. deren Metadaten, die über das Portal veröffentlicht werden sollen. Zur Realisierung des Datenregisters wurde die Software CKAN der Open Knowledge Foundation¹ verwendet, die an die Anforderungen des Berliner Pilotsystems angepasst wurde. Sie bietet eine Administrationsoberfläche

¹<http://okfn.org>.

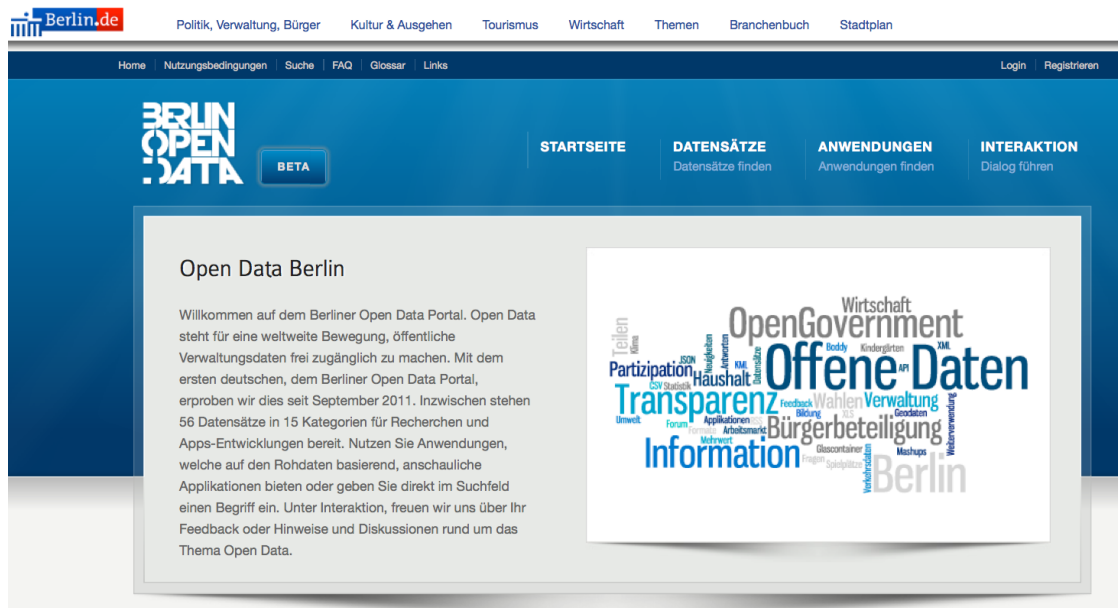


Abbildung 7.1: Startseite des Datenportals daten.berlin.de

sowie die Möglichkeit, Datensätze zu registrieren und Metadaten zu pflegen. Das dabei verwendete Metadatenschema wird in Abschnitt 7.3 hergeleitet und detailliert erläutert. Gemeinsam mit dem Schema wurde dabei eine Schnittstelle entwickelt, die die integrierte Bereitstellung von Metadaten aus dem CMS des Landes Berlin (Imperia) in das Open Data-Register ermöglicht. Eine Übersicht über die Datensätze und Apps zum Start und während der Laufzeit des Pilotsystems findet sich in Abschnitt 7.4. Die Erfahrungen aus dem Einsatz des Datenregisters werden in Abschnitten 7.5 geschildert; in den Abschnitten 7.6 und 7.7 wird dargestellt, wie das Pilotsystem von Datennutzern und Medien aufgenommen wurde. Abschließend werden die Aufgaben während des Pilotbetriebs und die nächsten Schritte dargestellt.

7.1 Datenportal

Die folgenden Grafiken geben einen exemplarischen Eindruck von der Gestaltung des Berliner Open Data-Portals und des Berliner Datenregisters.

Die Startseite des Berliner Open Data-Portals ist in Abbildung 7.1 dargestellt. Über den Reiter Datensätze gelangt man zum Datenkatalog, der eine Übersicht über die Datensätze, nach Kategorien gegliedert, präsentiert (vgl. Abb. 7.2). Durch die Auswahl eines Datensatzes erhält man eine Detailansicht, die zugehörige Metadaten darstellt und einen Link zur eigentlichen Datenquelle enthält (vgl. Abbildung 7.3).

Datensätze

[Startseite](#)

Hauptmenü

- Startseite
- Datensätze
 - Arbeitsmarkt
 - Bildung
 - Demographie
 - Geographie und Stadtplanung
 - Kunst und Kultur
 - Sonstiges
 - Sport und Erholung
 - Tourismus
 - Umwelt und Klima
 - Ver- und Entsorgung
 - Verbraucherschutz
 - Wahlen
 - Wirtschaft
 - Wohnen und Immobilien
 - Öffentliche Verwaltung, Haushalt und Steuern
- Anwendungen
- Interaktion

Die 5 neusten Datensätze

- Verzeichnis der

Hier können Sie nach Datensätzen suchen. Nutzen Sie dazu die Suchmaske, die sich nach Klick auf "Datensätze filtern" öffnet.

+

Datensätze filtern

[Verzeichnis der Weihnachtsmärkte in Berlin und Brandenburg 2011](#)

FREIE LIZENZ

Weihnachtsmärkte-Verzeichnis 2011 für den Raum Berlin-Brandenburg. Für Berlin geordnet nach Bezirken mit Informationen zu Öffnungszeiten und Standort.

Änderungsdatum: 11.11.2011

Kategorie: [Tourismus](#)

Tags: [Weihnachten](#), [Weihnachtsmarkt](#), [Weihnachtsmärkte](#)

[Straßenverzeichnis der Stadtplanung im Bezirk Lichtenberg](#)

KEINE FREIE LIZENZ

Dieses Straßenverzeichnis vom Bezirk Lichtenberg hilft Ihnen dabei den richtigen Ansprechpartner im Stadtentwicklungsamt Fachbereich Stadtplanung zu finden.

Änderungsdatum: 10.11.2011

Tags: [Bauen](#), [Bebauung](#), [Bebauungspläne](#), [Bezirk](#), [Entwurf](#), [Karte](#), [Lichtenberg](#), [Plan](#), [Simplesearch](#), [Stadtentwicklung](#)

Abbildung 7.2: Katalogseite des Datenportals daten.berlin.de

Einwohner am 30.6.2011

[Startseite](#)
[Datensätze](#)

Hauptmenü

- Startseite
- Datensätze
 - Arbeitsmarkt
 - Bildung
 - Demographie
 - Geographie und Stadtplanung
 - Kunst und Kultur
 - Sonstiges
 - Sport und Erholung
 - Tourismus
 - Umwelt und Klima
 - Ver- und Entsorgung
 - Verbraucherschutz
 - Wahlen
 - Wirtschaft
 - Wohnen und Immobilien
 - Öffentliche Verwaltung, Haushalt und Steuern
- Anwendungen
- Interaktion

Die 5 neuesten Datensätze

- Verzeichnis der

Melderechtlich registrierte Einwohner am Ort der Hauptwohnung in Berlin am 30.6.2011, Ortsteile (95), Altersgruppen, Geschlecht(1=männlich, 2=weiblich), Staatsangehörigkeit (D=Deutsch, A=Ausländer).

E-Mail Kontakt: info@statistik-berlin-brandenburg.de

Website: <http://www.statistik-berlin-brandenburg.de/produkte/Produkte-open-data.asp>

Zeitperiode: vom 30.06.2011 bis 01.07.2011

Lizenz: [Creative Commons Namensnennung](#)

Zeitliche Granularität: [Keine](#)

Veröffentlichende Stelle: [Amt für Statistik Berlin-Brandenburg](#)

Geographische Abdeckung: [Berlin](#)

Geographische Granularität: [Ortsteil](#)

veröffentlicht

14 SEP 2011

aktualisiert

06 OCT 2011

2

Ressource:

komprimierte CSV-Datei

URL: http://www.statistik-berlin-brandenburg.de/produkte/opendata/EWR_Ortsteile_2011-06-30.zip

Hash: d6017d0f-5717-487a-91fc-d07565b07252

Format: ZIP:CSV

Sprache: Deutsch

Kategorie: [Demographie](#)

Tags: [Ortsteil](#)

Abbildung 7.3: Detailseite des Datenportals daten.berlin.de

Neben der übersichtlichen Suche nach Datensätzen und deren Darstellung stehen weitere Bereiche bereit:

- Bei »Interaktion« befindet sich der Forenbereich. Hier können registrierte Nutzer Fragen stellen und Feedback hinterlassen sowie Fragen und Feedback anderer Nutzer kommentieren. Darüber hinaus besteht die Möglichkeit, einzelne Datensätze zu kommentieren.
- Unter »Anwendungen« werden Webseiten aufgelistet, die Anwendungen – auch Apps genannt – präsentieren, die auf Basis von offenen Datensätzen entwickelt wurden bzw. diese einbeziehen. Zum Start des Pilotsystems wurden dabei exemplarisch Anwendungen ausgewählt, die beliebige frei zugängliche Daten benutzen. Während des Pilotbetriebs wurden dagegen nur Anwendungen hinzugefügt, die auf Daten des Berliner Datenportals aufbauen.

Das Informationsangebot wird vervollständigt durch die im Projekt erstellte FAQ-Liste (Anhang B), eine Link-Liste (Anhang C) und ein Glossar zu Open Government und Open Data.

7.2 Datenregister

Um die Metadaten zu verwalten, kommt das Programm CKAN der Open Knowledge Foundation zum Einsatz, das an die Berliner Anforderungen angepasst wurde. Dieses Werkzeug ist frei lizenziert und lässt sich durch den konsequenten Opensource-Ansatz flexibel erweitern.

CKAN zeichnet sich durch zwei Leistungsmerkmale aus:

- Es gibt ein Metadatenschema vor, das einerseits gut für Datensätze geeignet ist, andererseits leicht erweitert und angepasst werden kann.
- In CKAN gespeicherte Metadaten können auf verschiedenen Wegen gepflegt und veröffentlicht werden; insbesondere existiert eine benutzerfreundliche Weboberfläche und eine mächtige Programmierschnittstelle, über die andere Programme auf die Metadaten zugreifen können.

Wegen dieser Eigenschaften ist CKAN in der Open Data-Community sehr beliebt und hat sich als Quasistandard etabliert. Das führt auch dazu, dass Datensätze schnell und einfach über verschiedene CKAN-Instanzen synchronisiert und gefördert werden können. Dieses Szenario wird z. B. dann wichtig, wenn es darum geht, die Berliner Open Data-Plattform an die geplante nationale oder europäische Open Data-Plattform anzubinden.

Die Metadaten werden durch Verwaltungsmitarbeiter bei der Datenlieferung im Berliner Datenregister eingetragen und von dort automatisch an das Berliner Open Data-Portal übermittelt. Die Oberfläche des nur mit Nutzeraccount für Verwaltungsmitarbeiter zugänglichen Datenregisters wurde individuell an den Berliner Fall angepasst. Das zentrale Metadatenformular ist in Anhang E dargestellt. Das Formular wurde als CKAN-Erweiterung realisiert und so wie CKAN selbst in der Open Source Community GitHub² veröffentlicht, sodass die Nachnutzung durch andere Kommunen kostenfrei möglich ist.

Das Datenregister ist auf Anwendungsebene gegen den Zugriff durch Dritte gesichert, d. h., Datenbereitsteller müssen sich mit einem durch den Plattformbetreiber zugewiesenen Benutzernamen und Passwort anmelden.

7.3 Metadaten für den Berliner Datenkatalog

Es gibt derzeit keinen allgemeingültigen Standard für Form und Bedeutung von Metadaten für Open Data. Trotzdem sollte Sorge dafür getragen werden, dass die Berliner Datensätze nicht isoliert bleiben, sondern möglichst interoperabel mit verschiedenen Anwendungen und anderen Datenkatalogen sind, z. B. auf nationaler oder europäischer Ebene. Aus diesem Grund wurden die Metadatenschemata der bekannten Datenportale

²<https://github.com/fraunhoferfokus/ckanext-berlin>.

analysiert und eine dazu passende Lösung für Berlin entworfen. Ergänzend wurden dazu passende programmatische Schnittstellen definiert und eine Liste der zu verwendenden Kategorien erstellt.

7.3.1 Analyse existierender Datenkataloge

Um zu einem zukunftsfähigen Metadatenschema zu kommen, wurden folgende Quellen berücksichtigt:

Die Software CKAN, die bei vielen Open Data-Portalen, wie auch in Berlin, zum Einsatz kommt, gibt erste, zentrale Felder vor. Die meisten sind jedoch nicht obligatorisch, außerdem müssen zahlreiche Felder als sogenannte »extras« ergänzt werden³.

```
title: text
url: text
author: text
author_email: text
maintainer: text
maintainer_email: text
notes: text
license_id: license identifier
tags: arbitrary user created strings
group (curated tag)
resource:
  url: url
  format: free text
  description: text
  hash: hash
```

Liste 7.1: CKAN-Standardfelder

Das englische Portal data.gov.uk ist mit mehreren Tausend Datensätzen eine der größten CKAN-Instanzen. Einige zusätzliche Felder haben sich hier bewährt und wurden daher in das Berliner Metadatenschema übernommen. Dazu zählen vor allem solche Angaben, die einen Datensatz räumlich und zeitlich einordnen.

```
date_released
date_updated
geographic_coverage
geographic_granularity
temporal_coverage_from
temporal_coverage_to
temporal_granularity
```

Liste 7.2: Zusätzliche Felder aus der Analyse von data.gov.uk

³http://wiki.ckan.net/Schema_for_Packages.

Des Weiteren wurden das Metadatenschema des Wiener Datenportals⁴ sowie die Studie zu Metadaten der Cooperation OGD Österreich (2011) berücksichtigt. Eingang fanden außerdem die Ergebnisse der europäischen Umfrage zu Open Data aus dem OpenCities Projekt⁵ und die bereits in der Vorstudie erfassten Datensätze (Barnickel et al. 2010).

Durch den Abgleich mit den wichtigsten Initiativen zu Open Data wird eine Integration bzw. Vernetzung zwischen diesen und dem Berliner Datenportal ermöglicht. Dazu wird lediglich jeweils eine bedeutungserhaltendene Abbildungsfunktion benötigt (sogenanntes Semantic Mapping).

7.3.2 Metadatenschema, Kategorien und APIs für Berlin

Zusätzlich zu den übernommenen Einträgen anderer Datenportale wurden folgende Felder für Berlin eingeführt:

<code>relates_to</code> <code>apps</code>
--

Liste 7.3: Zusätzlich eingeführte Felder

Es besteht also die Möglichkeit, formal eindeutig Zusammenhänge zwischen Datensätzen abzubilden sowie Anwendungen anzugeben, die den betreffenden Datensatz verwenden.

Die vollständige Liste der Metadateneinträge des Berliner Datenportals ist im Anhang D unter »Spezifikation Output API« enthalten, da die registrierten Datensätze nach diesem Schema ausgelesen werden können (vgl. CKAN-Schnittstellen⁶).

Teil der Metadaten ist auch eine Prüfsumme für jede Datei eines Datensatzes. Diese erlaubt die Überprüfung der entsprechenden Datei auf Authentizität. Die Prüfsumme kann entweder beim Registrieren eines Datensatzes angeben oder, im Fall des Eintrags über das Imperia-CMS, automatisch beim Hochladen ermittelt werden. Volle Bedeutung erlangt die Prüfsumme jedoch erst, wenn Daten nicht mehr wie heute direkt beim Datenbereitsteller heruntergeladen werden, der den Datensatz auch registriert hat, sondern erst dann, wenn die eigentlichen Daten typischerweise von Dritten ausgeliefert werden, z. B. durch sogenannte Spiegelserver.

Gemäß dem Architekturvorschlag wurde für Datenbereitsteller eine zusätzliche Programmierschnittstelle realisiert. Diese besitzt bedeutungsmäßig die gleichen Felder, verwendet jedoch deutsche Bezeichner, um die Integration mit den Berliner Behörden zu vereinfachen. Diese Schnittstelle einschließlich Angaben zu Datentypen der einzelnen Felder ist ebenfalls in Anhang D dokumentiert (»Spezifikation Input API«), .

⁴<http://data.wien.gv.at>.

⁵<http://opencities.net>.

⁶<http://docs.ckan.org/en/latest/api.html>.

Des Weiteren enthält Anhang D Listen für diejenigen Felder, die nur eine bestimmte Menge an Eingabetexten zulassen. Die Kategorienliste wurde gemeinsam mit den Metadatenschemata konsolidiert (vgl. den vorhergehenden Abschnitt), wobei jede Kategorie mit einem eindeutigen, maschinenfreundlichen Bezeichner versehen wurde. Hinsichtlich der geographischen Abdeckung wurden alle Berliner Bezirke und Ortsteile aufgenommen. Die geographische Granularität bzw. Auflösung erlaubt Konzepte verschiedener Bezugssysteme, insbesondere das der Lebensweltlich orientierten Räume (Bömermann et al., 2006). Für die zeitliche Granularität sind verschiedene Zeitintervalle zugelassen, um von sekundengenauen Sensordaten bis zu Wahldaten allen Datensätzen die richtige Auflösung zuordnen zu können.

7.4 Umfang und Entwicklung bei Datensätzen und Apps

Für den Start des Portals wurden 18 Datensätze aus fünf verschiedenen Kategorien in das Register aufgenommen. Sie stammten vor allem aus dem Bestand des Amts für Statistik Berlin-Brandenburg. Mit Stand vom 5. Dezember 2011 umfasste das Datenportal 56 Datensätze aus 15 Kategorien.⁷ Die erste Quelle für diesen Fortschritt war die Senatsverwaltung für Wirtschaft, Technologie und Frauen: Hier wurden während der ersten Monate des Pilotbetriebs zahlreiche Daten maschinenlesbar veröffentlicht, frei lizenziert und im Datenportal registriert. Außerdem trugen wie geplant die SimpleSearches von BerlinOnline zur Vergrößerung des Bestands bei. Insbesondere der Bezirk Lichtenberg meldete im Pilotzeitraum zahlreiche Datenquellen auf diesem Weg an, die dadurch automatisch auch im Register aufgeführt werden.

Bemerkenswerterweise sind schon in der Pilotlaufzeit mehrere Anwendungen entstanden, die auf Daten aus dem Berliner Datenportal aufbauen. Sie wurden von unbeteiligten Dritten realisiert, die durch das Portal auf die jeweiligen Datensätze aufmerksam wurden. Drei davon wurden zeitnah in die Liste der Anwendungen auf dem Portal aufgenommen. Besonders zu erwähnen ist dabei der Ozon Sonar⁸, eine Visualisierung der Berliner Ozonbelastung, die nur einen Tag nach dem Launch des Portals veröffentlicht wurde.

Auch bei Aggregatoren für Datenkataloge wurde das Berliner Portal registriert:

- <http://datacatalogs.org>
- <http://offenedaten.de>
- <http://datos.fundacionctic.org/sandbox/catalog/faceted>

Dies demonstriert eine erfolgreiche Förderierung demonstriert.

⁷Dabei enthält ein Datensatz bis zu 20 verschiedene Datenquellen bzw. Dateien, z. B. »20 grüne Hauptwege«.

⁸<http://daten.berlin.de/anwendungen/ozon-sonar>.

7.5 Datenregister im Einsatz, Handreichung, Rückmeldung

Der erste Datenbereitsteller, der das angepasste Metadatenformular benutzte, war die Senatsverwaltung für Wirtschaft, Technologie und Frauen. Zum Einstellen von mehreren Datensätzen wurde in Zusammenarbeit mit FOKUS eine Handreichung erstellt. In dieser ist anschaulich erklärt, welche Schritte zur Registrierung eines Datensatzes notwendig sind. Sie ist als Anhang F beigelegt und dient dazu, dass weitere Datenbereitsteller in Zukunft das Register vereinfacht nutzen können.

Außerdem wurden die Erfahrungen und Rückmeldungen des ersten Datenbereitstellers genutzt, um Verbesserungsmöglichkeiten für die Administrationsoberfläche zu finden. Zum Teil wurden diese Vorschläge direkt umgesetzt, teilweise gingen die Vorschläge über den Rahmen eines Prototypen hinaus.

Als empfohlenene Datenformate wurden XLS und CSV festgelegt. Diese stellen hinsichtlich der maschinellen Verarbeitung eine deutliche Verbesserung gegenüber PDF-Dokumenten dar. Außerdem lässt sich eines der beiden Formate direkt aus dem meisten Anwendungsprogrammen der Verwaltungsmitarbeiter exportieren. Damit ist der Aufwand erheblich geringer als beim Erzeugen des optimalen Open Data-Formats RDF. Allerdings lassen sich auch diese Formate nur verwerten, wenn benutzte Schlüssel und Abkürzungen klar dokumentiert sind.

Als Teil der Metadaten eines Datensatzes muss auch ein Name festgelegt werden. Dieser wird u. a. dafür verwendet, dem Datensatz eine weltweit eindeutige URL zuzuweisen. Die Administrationsoberfläche generiert einen Namen automatisch aus dem Titel. Der generierte Name lässt sich jedoch optimieren, etwa durch das Entfernen von Bindewörtern. Insgesamt sollte der Name möglichst kurz sein, aber den Datensatz immer noch sinnvoll identifizieren. Wichtig ist weiter, dass der Name keine Umlaute, Leer- oder Sonderzeichen oder Großbuchstaben enthält. Beinhaltet ein Datensatz nur Daten aus einem Jahr (z. B. Wahlergebnisse von 2011), so sollte die Jahreszahl dem Namen hinzugefügt werden. Sind aber Zahlenreihen (z. B. Beschäftigte seit 2007) eingestellt, so ist das Enddatum wegzulassen, damit bei neuen Jahresscheiben der Datenumfang einfach erweitert werden kann, ohne dass der Namen des Datensatzes geändert werden muss. Bei ständig aktualisierten Daten (z. B. Verkehr oder Umwelt) ist kein Zeitbezug in der URL anzugeben.

Diese Hinweise gelten grundsätzlich auch für die URL des eigentlichen Datensatzes, der weiterhin bei der jeweiligen Behörde abgelegt und gepflegt wird (z. B. `www.berlin.de/behörde/<ggf.-unterseite>/<bezeichnung>`).

7.6 Nutzerzahlen

Abbildung 7.4 illustriert die Besucherzahlen des Datenportals. Es lässt sich also ablesen, dass etwa 2000 Webnutzer das Portal täglich besuchen und dabei im Schnitt zwei bis drei Seiten ansehen.

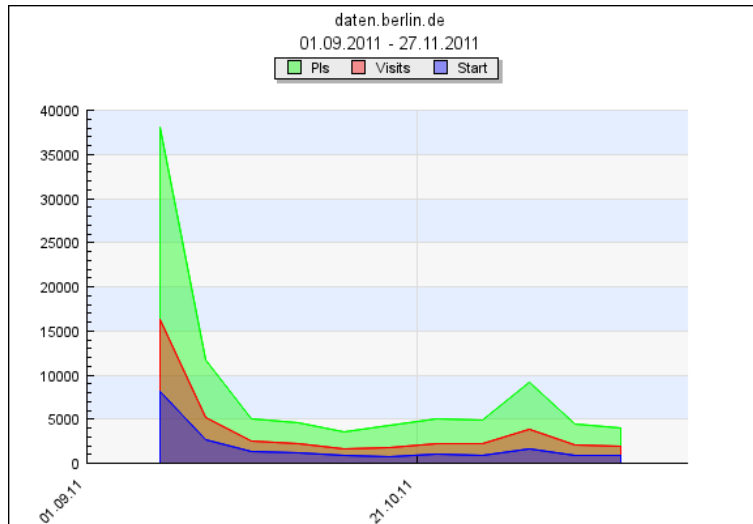


Abbildung 7.4: Nutzungsstatistik des Datenportals daten.berlin.de
 »PI« – kurz für Page Impression: die Anzahl angezeigten Seiten
 »Visitors«: die Anzahl unterschiedlicher Besucher
 »Start«: die Frequenz der Startseite

Wie sich das Interesse der Nutzer auf die verschiedenen Bereiche verteilt, geht aus Tabelle 7.1 hervor. Dargestellt sind die angezeigten Seiten für den Monat November 2011. Das Interesse an den Datensätzen überwiegt stark das an den Anwendungen bzw. der Interaktion, zumal auch die Seite »Kategorie« zu den Datensätzen gezählt werden muss.

Tabelle 7.2 gibt die beliebtesten Datensätze und die zugehörige Aufrufszahl für den Monat November wieder. Mit der Einschränkung, dass es noch kein repräsentatives Angebot an Daten gibt, lässt sich festhalten, dass Daten mit Raumbezug generell von größerem Interesse sind als solche ohne.

Für den Bereich Interaktion lässt sich festhalten: Mit Stand vom 29. November sind 110 Nutzer auf dem Portal registriert. Es gibt 16 Kommentare und 7 Rückmeldungen von Nutzern im Bereich »Interaktion«. Über das Kontaktformular sind 13 Anfragen eingetroffen.

Es ist zu erwarten, dass sich die Anzahl der Besucher mit einem wachsenden Datenangebot weiter erhöht. Erst dann lässt sich auch verlässlich sagen, inwieweit die Nutzungszahlen für einzelne Datensätze mit der Umfrage⁹ der Senatsverwaltung für Wirtschaft, Technologie und Frauen übereinstimmen. Zum jetzigen Zeitpunkt lässt sich nur beobachten, dass für Nutzungszahlen der Raumbezug eine große Rolle spielten, was insoweit mit der Umfrage übereinstimmt, als dort vermehrt Stadtplanungsdaten nachgefragt worden waren.

⁹<http://www.berlin.de/sen/wtf/voting/opendata.php>.

Bereich	angezeigte Seiten
Datensätze	7 343
Startseite	5 548
Kategorie	5 397
Anwendungen	2 802
Interaktion	924
Search	452
Tags	319
Impressum	97
FAQ	92
User	87
Nutzungsbedingungen	86

Tabelle 7.1: Nutzungsstatistik des Datenportals daten.berlin.de nach Bereichen (November 2011)

7.7 Medienresonanz

Die folgende Liste gibt einen Überblick über die Berichterstattung zum Berliner Datenportal:

- Meldung in den Online News des Heise-Verlags: <http://www.heise.de/newsticker/meldung/OpenData-Portal-Berlin-eroeffnet-1342810.html>
- Meldung auf der Seite der Open Knowledge Foundation: <http://okfn.de/2011/09/daten-berlin-de-berlin-macht-sich-auf-zur-open-data-hauptstadt/>
- TV-Beitrag auf TVBerlin: <http://www.tvbvideo.de/video/iLyROoaf2EYH.html>
- Meldung auf EPSIplatform.eu: <http://epsiplatform.eu/content/berlin-launches-data-portal>

7.8 Wartungsaufgaben während des Pilotbetriebs

Während der Pilotlaufzeit fielen verschiedene Wartungsaufgaben an. Anhand der folgenden Liste soll einerseits die Planung des Regelbetriebs präzisiert werden, andererseits dokumentiert sie, wo sich durch Implementierungstätigkeit Wartungsarbeiten automatisieren lassen.

- Aktualisierung der Software auf den Servercomputern
- Korrektur der neu implementierten Funktionalitäten nach Identifikation von Fehlern
- Verschiedene Schnittstellen synchron halten
- Inhalte pflegen, insbesondere Neuigkeiten und Linkliste

Datensatz	Angezeigte Seiten
Verzeichnis-Der-Weihnachtsmärkte-Berlin-Und-Brandenburg-2011	610
Einwohner-Am-30.6.2011	392
GPS-Tracks-Für-Die-Radrouten-Durch-Berlin	258
Ortsteil-Geometrien-Berlin	226
Open-Streetmap-Daten-Für-Berlin	196
Koordinaten-Für-Spazier-Und-Wanderstrecken	160
Luftgüte-Tagesaktuell	152
Glascontainer-Auf-öffentlichem-Straßenland-Charlottenburg-W.	137
Ausschreibungen-Nach-VOL-VOB-VOF	125
Abschlussprüfungen-Hochschulen-2009	121
Haushaltsplan-Berlin-2011	113
Anzahl-Arbeitsloser-Frauen-Und-Männer-Berlin	107
Mietbare-Veranstaltungsräume-Im-Bezirk-Lichtenberg	101

Tabelle 7.2: Nutzungsstatistik des Datenportals daten.berlin.de nach Datensätzen (November 2011)

- auf Nutzerkommentare reagieren
- Anpassung der Lizenzliste: Regelmäßig tauchen Datensätzen auf, die zwar eine offene Lizenz, aber keine der bisher verwendeten Lizenzen haben. Dann müssen entsprechende Einträge angelegt werden.
- Nutzerverwaltung: Anlegen von Nutzern, Zurücksetzen von Passwörtern, Rechtevergaben, manuelle Korrektur, gegebenenfalls Rückgängigmachung von Änderungen
- Support für Datenbereitsteller
- Über Aufnahme von Datensätzen und Anwendungen entscheiden
- Back-up von Datenregister und Inhalten des Datenportals
- Pflege der Kategorienliste: Hinzufügen, Teilen, Zusammenführen von Kategorien

7.9 Fazit und nächste Schritte

Der Prototyp demonstriert deutlich, dass sich für das Land Berlin in verhältnismäßig kurzer Zeit ein Datenportal realisieren ließ, das durchaus über die Grenzen von Berlin hinaus auf großes Interesse bei Bürgern, Unternehmen und auch der Verwaltung selbst stößt. Gleichzeitig zeigt sich, dass weitere Ausbaustufen und ein dedizierter Regelbetrieb sichergestellt werden müssen, um das Open Data-Portal zu einem festen Bestandteil der Berliner eGovernment-Landschaft zu entwickeln. Die wichtigsten Ziele sollten dabei sein: die Steigerung des Datenangebots durch Popularisierung und Qualifizierung zum Datenregister innerhalb der Verwaltung, die überregionale Koordination des Metadatenschemas sowie eine Unterstützung für Linked Data.

Die wichtigste Maßnahme für die Etablierung eines Open Data-Ökosystems ist zweifelsfrei der Ausbau des Angebots. Hier lohnt noch einmal der Vergleich mit der Stadt London, wo etwa zehnmal so viele Datensätze bereitstehen und regelmäßig neue hinzukommen. Dafür ist es nötig, innerhalb der Verwaltung für die mit einem relativ geringen Aufwand verbundenen Vorzüge der Beteiligung an Open Data zu werben, z. B. mit Informationsveranstaltungen und Workshops. Gleichzeitig ist es wichtig, eine Anlaufstelle zu etablieren, wo jegliche Fragen und Probleme der Mitarbeiter zum Thema Open Data gegebenenfalls in Erfahrung gebracht und beantwortet werden können. Außerdem wäre es hilfreich, über einen Grundstock an Dokumentations- und Trainingsmaterial im Stil der Handreichung in Anhang F zu verfügen, mit dessen Hilfe sich Mitarbeiter jederzeit selbständig informieren können. Durch weitergehende Prozess- und Organisationsanalysen lassen sich hier unter Umständen auch Wege finden, wie Daten automatisiert als Open Data veröffentlicht werden können.

Damit Datenkataloge im Allgemeinen und das Berliner Datenregister im Besonderen nicht zu Datensilos werden, sondern Metadaten frei ausgetauscht und gefördert werden können, ist es nötig, einen Metadatenstandard zu setzen. Bis die entsprechende W3C-Arbeitsgruppe¹⁰ eine entsprechende Empfehlung ausspricht, werden mit jedem Datensatz Tatsachen geschaffen. Daher gilt es, sich gemeinsam mit Vorreitern wie der Stadt Wien und der Open Knowledge Foundation als CKAN-Produzent auf Form und Bedeutung von Metadaten zu einigen. Dabei sollen einerseits die bestehenden Datensätze abgebildet werden können, andererseits muss eine bedeutungserhaltende Übertragung auf kommenden Standards realisierbar sein.

In technologischer Hinsicht muss der Prototyp langfristig klar in Richtung Linked Data weitergeführt werden. Das bedeutet, dass ein sogenannter Datastore – genauer: ein Triplestore – aufgebaut werden muss, also eine Speicherschicht, in der jedes Einzeldatum als individueller, maschinenverständlicher Fakt abgelegt und abrufbar ist. Dieses Projekt muss von weiteren Umstellungen in der Berliner Verwaltung hin zu Linked Data sowie der Entwicklung passender Schnittstellen und Werkzeuge für die Entwicklerseite flankiert werden.

¹⁰<http://www.w3.org/egov>.

8

Handlungsempfehlungen

Informationen sind der Rohstoff des 21. Jahrhunderts. Der offene und strukturierte Zugang zu frei verfügbaren Datenbeständen der öffentlichen Hand ist ein wichtiger Beitrag für die Weiterentwicklung der Wissensgesellschaft. Er stärkt das Vertrauen zwischen Politik und Zivilgesellschaft, zwischen Verwaltung, Wirtschaft und Medien. Staatliches Handeln wird transparenter, nachvollziehbarer und überprüfbarer. Gleichzeitig birgt die Öffnung Potenziale für effizienteres Handeln und Qualitätsgewinn in sich. Zudem tragen der freie Zugang zu und die proaktive Bereitstellung von Informationen und Daten wesentlich zu dem Ziel bei, die Zusammenarbeit und den Zusammenhalt gesellschaftlicher Gruppen zu stärken. Neue Angebote aus der Wirtschaft können entstehen und einen Beitrag zur Weiterentwicklung des Gemeinwesens leisten.

Die im Folgenden zusammengestellten Handlungsempfehlungen sind in politische, organisatorische, rechtliche und technische Handlungsempfehlungen gegliedert. Sie sind Resultat der Studie und der Erkenntnisse aus dem Pilotbetrieb des Berliner Datenportals. Hinsichtlich der zeitlichen Dimension wird eine Einteilung in kurzfristige (K, ca. 1 Jahr), kurz- bis mittelfristige (KM), mittelfristige (M, ca. 2-3 Jahre), mittel- bis langfristige (ML) und langfristige (L, ca. 5 Jahre) Handlungsempfehlungen vorgenommen. In Kombination mit dem Ausbau des Berliner Datenportals führen die hier vorgeschlagenen Maßnahmen zum abschließenden Stufenplan für nächste Schritte zu offenen Daten im Land Berlin.

8.1 Politische Handlungsempfehlungen

8.1.1 Politisches Bekenntnis zu offenen Daten in und aus Berlin

Der Zielkonflikt zwischen Beiträgen zur Haushaltssanierung durch den Verkauf von Datenbeständen, dem Berliner Informationsfreiheitsgesetz und der Offenlegung entsprechend der EU-PSI-Richtlinie (siehe Abschnitt 6.2) und dem deutschen Informationsweiterverwendungsgesetz (IWG) (siehe Abschnitt 6.1.2) ist durch ein politisches Bekenntnis des Landes Berlin zu offenen Daten und eine gesetzliche Handlungsgrundlage aufzulösen. Eine Ressortzuordnung für Belange offener Daten in Berlin und die Ernennung eines Open Data-Verantwortlichen für Berlin sind dabei wesentlich. Den Datenbearbeitern in Berlin sind klare Richtlinien an die Hand zu geben, die eine Identifizierung und die Bereitstellung offener Datensätze ermöglichen.

→ **K1.** Es ist ein klares politisches Bekenntnis zu offenen Daten notwendig: Ein Prinzip der generellen Veröffentlichung nicht personenbezogener und sicherheitsunkritischer Daten im öffentlichen Interesse ist im Berliner Informationsfreiheitsgesetz zu verankern.

→ **K2.** Das politische Bekenntnis zu offenen Daten in und aus Berlin bildet die Grundlage für konkrete Maßnahmen und ist daher zur Beförderung der anderen Handlungsempfehlungen sehr kurzfristig umzusetzen.

→ **K3.** Die Finanzierung des Berliner Datenportals, die den Betrieb des Portals, die Bereitstellung der Datensätze und die Durchführung der für Open Data relevanten Prozesse in Berlin umfasst, ist sicherzustellen.

→ **K4.** Es sind Richtlinien zur Auswahl von Daten für das Berliner Datenportal hinsichtlich der wirtschaftlichen, sicherheitstechnischen und datenschutzbezogenen Eignung zu erarbeiten und in den Verwaltungsvorschriften zu verankern. Es ist notwendig, einen Kriterienkatalog zur Auswahl von Daten für das Berliner Datenportal zu erarbeiten, der mögliche Datensätze in ihren Eigenschaften charakterisiert, sodass eine Prüfung der Datensätze auf Eignung möglich ist.

→ **K5.** Werkzeuge und Schulungen sind für die Datenaufbereiter und die Redakteure des Datenportals bereitzustellen, damit effizient und zeitnah qualitativ hochwertige Datenangebote erstellt werden können.

8.1.2 Generelle Informationsstrategie

Die Open Data-Strategie des Landes Berlin ist im Kontext einer generellen Informationsstrategie des Landes Berlin zu sehen, die zu entwickeln ist und in die die Open Data-Strategie eingebunden werden sollte.

→ **M1.** Die Open Data-Strategie des Landes Berlin ist ein wesentliches Element einer zu entwickelnden generellen Informationsstrategie und sollte deshalb nicht singulär, sondern im Kontext dieser generellen Strategie vorangetrieben werden.

8.1.3 Politisches Bekenntnis zum Berliner Datenportal

Zusätzlich zu der vorliegenden Studie ist bereits ein Berliner Datenportal für offene Daten in und aus Berlin unter <http://daten.berlin.de> freigeschaltet worden. Über 100 Datenquellen aus Behörden wurden bereits eingestellt. Das Berliner Datenportal diene als Vorbild für die zweistufige Ausschreibung des Bundesministeriums des Innern zu offenen Daten in Deutschland und bietet Datensätze als Grundlage für den Apps4Deutschland-Wettbewerb, der im Oktober 2011 gestartet wurde.

→ **K6.** Der Übergang in den Regelbetrieb des Berliner Datenportals ist zu sichern. Dazu ist die Fachaufsicht einer Dienststelle festzulegen, die auch die Registrierung von Mitarbeitern für das Datenregister vornimmt, ferner ist der Regelbetrieb auf dem Stadtportal langfristig sicherzustellen. Landesredakteure sind mit der Aufbereitung der Berliner Datensätze sowie mit der Moderation auf dem Datenportal zu beauftragen. Ein Betriebskonzept mit festgelegten Rollen und Verantwortlichkeiten ist zu erarbeiten.

8.1.4 Gesamtverantwortlicher für offene Daten und Ansprechpartner zu den Daten

Für die Arbeit zu offenen Daten in Berlin sind die Verantwortlichkeiten zu regeln und Verantwortliche zu benennen.

→ **K7.** In Ergänzung zu den Ansprechpartnern je Datensatz ist

1. ein Open Data-Verantwortlicher für Berlin zu bestimmen und
2. eine ressortübergreifende Arbeitsgruppe für den Ausbau des Berliner Datenportals einzurichten.

8.1.5 Stabile und langfristige Bereitstellung der Daten

Eine erfolgreiche Nutzung der Angebote des Datenportals durch die Wirtschaft wie auch die Erhöhung der Transparenz staatlichen Handelns basieren wesentlich auf einer stabilen und langfristigen Bereitstellung der Berliner Daten. Das Angebot ist so zuverlässig und qualitativ hochwertig umzusetzen, dass ein nachhaltiges Handeln beispielsweise auf wirtschaftlichem, gesellschaftlichem und journalistischem Gebiet unter Nutzung dieser Daten ermöglicht wird.

→ **KM1.** Die Bereitstellung der Daten muss in offenen, maschinenlesbaren, wohldefinierten und dokumentierten Formaten und Datenkatalogen über ein zentrales und benutzerfreundliches Berliner Datenportal erfolgen, das festgelegte Qualitätskriterien für die Datenangebote erfüllt.

8.1.6 Ausbau des Datenportals

Für die aktive Bereitstellung von offenen Verwaltungsdaten (außer dem Datenschutz und der Sicherheit unterliegenden Daten) für jeden (Bürger, Unternehmen, Medien, Verwaltung etc.) zur Information und zunehmenden Transparenz in der Stadt, die eine Kombination und Integration mit Daten aus anderen Quellen erlauben, werden folgende Maßnahmen vorgeschlagen. Ausnahmen müssen begründet werden.

→ **KM2.** Inhaltlicher Ausbau: Es sind kontinuierlich weitere maschinenverarbeitbare Datensätze aus der Verwaltung und von anderen Stellen auf dem Berliner Datenportal bereitzustellen.

→ **M2.** Organisatorischer Ausbau: Die notwendigen rechtlichen Grundlagen und Anpassungen der Verwaltungsvorschriften in Berlin sind zu schaffen.

→ **ML1.** Technischer Ausbau: Es sollten weitere Projekte zur Entwicklung von Mehrwertdiensten und Werkzeugen für das Angebot offener Daten in Berlin durchgeführt werden.

8.1.7 Föderation des Berliner Datenportals mit anderen Portalen

Berlin leistet mit dem Open Data-Konzept und dem Datenportal Pionierarbeit. Nur wenige europäische Metropolen verfügen heute bereits über ein solches Angebot. Die verschiedenen europäischen Einzelaktivitäten führen jedoch ohne gemeinsame Vereinbarungen oder gar Standards zu Insellösungen, die eine zukünftige Verflechtung für integrierte Angebote behindern.

→ **ML2.** Berlin sollte bei nationalen und europäischen Gremien einfordern, dass die besonders drängende Frage der Vereinheitlichung von Formaten, Metadaten und Lizenzen einer Abstimmung und Standardisierung zugeführt wird, und sich in diese Prozesse selbst aktiv einbringen.

8.2 Organisatorische Handlungsempfehlungen

8.2.1 Anschlussfähigkeit der Open Data-Strategie

Das Land Berlin hat bereits wesentliche Vorarbeiten zur Bereitstellung von Verwaltungsinformationen wie den Datenpool beim Amt für Statistik Berlin-Brandenburg oder die

Geodateninfrastruktur der Senatsverwaltung für Stadtentwicklung geleistet. Der weitere Ausbau des Berliner Datenportals sollte mit existierenden Vorhaben koordiniert werden.

→ **KM3.** Der Ausbau des Berliner Datenportals ist mit existierenden Vorhaben zur Bereitstellung von Verwaltungsinformationen des Landes Berlins zu koordinieren.

8.2.2 Richtlinien zur Auswahl der Daten

Bei der Auswahl möglicher Datensätze für das Berliner Datenportal treten in den Institutionen Schwierigkeiten bei der Bestimmung der Datensätze im Rahmen der rechtlichen Randbedingungen auf. Auch wenn nicht personenbezogene und nicht sicherheitskritische Datensätze zur Veröffentlichung anstehen, kann eine Einschätzung der Sachlage im konkreten Fall schwer sein. Datenschutzprobleme werden befürchtet.

→ **K8.** Es ist zu klären, welche Grenzen bezüglich einer datenschutzmäßigen Unbedenklichkeit beachtet werden müssen – beispielsweise bei organisatorischen Informationen der Institutionen wie Namen, Adressen oder Terminen.

8.2.3 Wohldefinierte Prozesse für das Datenportal

Für den Betrieb, den Ausbau und die Wartung des Berliner Datenportals sind wohldefinierte Prozesse zu beschreiben, die Verantwortlichkeiten auch in Interaktion mit den Nutzern des Datenportals festlegen.

→ **K9.** Es ist ein redaktioneller Verantwortlicher zu bestimmen, der die Redaktion für das Berliner Datenportal leitet und koordiniert.

→ **K10.** Unter den Redakteuren des Berliner Datenportals ist ein Verantwortlicher zu bestimmen, der die Verantwortung für die Interaktion mit den Nutzern trägt und zeitnah die Reaktion auf die Nutzeranfragen und -vorschläge koordiniert.

→ **M3.** Die Redaktion des Berliner Datenportals ist durch einen zuverlässigen technischen Betrieb zu untermauern, der neben Zuverlässigkeit und Barrieriefreiheit des Angebots auf die IT-Sicherheit, das Back-up und die Archivierung des Angebots achtet.

8.2.4 Weiterbildungsmaßnahmen zu offenen Daten

Entsprechend dem heutigen Auftrag veröffentlichen Institutionen bislang Daten für genau bestimmte Zwecke mit den entsprechenden Restriktionen und dazu passenden (Insel-)Strukturen. Open Data erfordern jedoch einen Paradigmenwechsel an verschiedenen Stellen: Alles ist öffentlich, was nicht ausdrücklich als geheim gekennzeichnet ist. Alle Daten, die keiner berechtigten Datenschutz- oder Sicherheitsbeschränkung unterliegen, werden proaktiv, in vollem Umfang und zeitnah veröffentlicht.

→ **M4.** Zur Umsetzung dieses Paradigmenwechsels sind passende Weiterbildungsmaßnahmen aufzubauen und anzubieten, um die sich daraus ergebenden Vorteile für die Institutionen zu vermitteln und organisiertes Handeln im Umgang mit offenen Daten zu ermöglichen.

8.2.5 Vielfältige Interaktionen mit den Nutzern des Berliner Datenportals

Das Berliner Datenangebot profitiert von der vielfältigen Interaktionen mit den Nutzern – sowohl aus der Zivilgesellschaft, den Organisationen und Unternehmen der Stadt als auch aus der Verwaltung. Beispielsweise können in einem Ideenportal Überlegungen geäußert und kommentiert werden, sowohl zum Portal selbst als auch zu konkreten oder gewünschten Datensätzen. Auf einem Wunschzettel können Nutzer angeben, welche Datensätze und welche Anwendungen gewünscht werden. Die Nutzergemeinde sollte über vorgeschlagene Datensätze abstimmen und diese priorisieren können. In Foren können »Unterhaltungsfäden« zu verschiedenen Themen erstellt werden. Dabei können neben der bloßen Diskussion Nutzer sich auch gegenseitig helfen, und zwar in der Form, dass eine einmal diskutierte Lösung danach unbeteiligten Nutzern ebenfalls zur Verfügung steht. Auch und gerade im Forenbereich ist Moderation notwendig. Hier können Redakteure doppelte Themen zusammenfassen, aber auch auf ähnliche Themen verweisen. Außerdem sind Mechanismen umzusetzen, mit denen entsprechend der Nutzungsordnung unangemessene Beiträge entfernt werden können.

→ **M5.** Die bestehenden Interaktionsmöglichkeiten wie Kommentare und Feeds sollten um weitere Möglichkeiten wie beispielsweise ein Ideenportal oder einen Wunschzettel erweitert werden.

8.2.6 Vorschlagsmöglichkeiten für Datensätze

Des Öfteren werden bereits jetzt ausgewählte Datensätze wie Geodaten oder Daten des öffentlichen Nahverkehrs von den Nutzern des Berliner Datenportals nachgefragt.

→ **M6.** Das Angebot auf dem Berliner Datenportal ist kontinuierlich auf mögliche und nötige Erweiterungen zu prüfen und entsprechend auszubauen.

→ **KM4.** Ein bereits heute klar erkennbarer Bedarf besteht an einem frei verfügbaren Raumbezugssystem für Berlin. Es ist zu klären, inwiefern ein solches über das Berliner Datenportal zur Verfügung gestellt werden kann.

8.2.7 Publiizierte Daten in das Datenportal stellen

Die Berliner Behörden und nachgeordneten Einrichtungen veröffentlichen regelmäßig umfangreiches Zahlenmaterial sowie Daten und Fakten in Form von Pressemitteilungen, Berichten und Zusammenstellungen aus den verschiedensten Amtsbereichen. Auch die

Beantwortungen parlamentarischer Anfragen enthalten vielfach interessante Datensätze. Das Material wird durch Medien, Unternehmen und Bürger zur Kenntnis genommen, aufbereitet und diskutiert.

→ **K11.** Pressemitteilungen, Berichte und andere Veröffentlichungen mit Zahlenmaterial sind um maschinenlesbare Tabellen, die auf den Originaldaten beruhen, zu ergänzen und mit der Meldung oder dem Bericht auf der Website der betreffenden Behörde verfügbar zu machen. Die Pressestellen sorgen für den Eintrag im Datenregister des Berliner Datenportals.

8.2.8 Beschlüsse und Protokolle in das Datenportal stellen

Neben Basisdaten haben auch Beschlüsse und Protokolle der Berliner Verwaltung und der Landespolitik eine Relevanz im Rahmen offener Daten und sollten über das Datenportal verfügbar gemacht werden.

→ **M7.** Mittelfristig sollen alle Beschlüsse und Protokolle von Senats-, Stadtrats-, Parlaments- und Ausschusssitzungen in einem offenen Format wie beispielsweise Office Open XML (OOXML) oder OASIS Open Document Format (ODF) veröffentlicht werden.

8.2.9 Daten landeseigener Unternehmen in das Datenportal stellen

Das Portal sollte nicht nur Basisdaten der Verwaltungseinrichtungen des Landes Berlin bereitstellen, sondern ebenso Basisdaten von Unternehmen und Organisationen und insbesondere die Daten landeseigener Unternehmen, Stiftungen etc.

→ **M8.** Landeseigene Unternehmen, Stiftungen etc. sollen mittelfristig Basisdaten ins Datenportal einstellen.

8.3 Rechtliche Handlungsempfehlungen

8.3.1 Einheitliche und einfache Lizenz- und Nutzungsbestimmungen

Es wird die Festlegung von einheitlichen und einfachen, auf der Open Definition basierenden Lizenz- und Nutzungsbestimmungen empfohlen, die eine Weiterverarbeitung, Weiterverwendung und Weiterverbreitung der Daten des Landes Berlin durch jeden und für jegliche Zwecke, auch kommerzielle, befördern. Minimale Kosten sollten nur in Ausnahmefällen (beispielsweise für die technische Bereitstellung) erhoben werden.

→ **K12.** Datensätze sollten unter Creative Commons License mit Namensnennung (CC BY) bereitgestellt werden. Ausnahmen von dieser Grundregel sollten begründet werden.

→ **M9.** Für Datensätze, bei denen die kommerzielle Nutzung eingeschränkt werden sollte, sind ein Lizenzvorschlag und ein Nachnutzungsangebot zu entwickeln, die beispielsweise anteilige Kostendeckung im Fall der kommerziellen Nutzung zulassen. Dies ist auf nationaler und europäischer Ebene abzustimmen.

8.3.2 Direkter Zugang zu den Lizenz- und Nutzungsbestimmungen

Zur Unterstützung der Weiterverwendung offener Daten ist eine einfache Vermittlung der jeweiligen Nutzungsrechte essenziell.

→ **KM5.** Die Lizenz- und Nutzungsbestimmungen werden gut sichtbar bereitgestellt und begleitend erläutert. Nutzerseitige Nachfragen werden zeitnah und rechtssicher beantwortet.

8.3.3 Nachvollziehbarkeit der Datenquellen

Die Datenangebote sind eindeutig bezüglich der Urheberschaft und zudem – soweit möglich – hinsichtlich der erfolgten Aggregationen der Daten durch Dritte zu kennzeichnen.

→ **M10.** Eine Lizenz mit Namensnennung für die Datensätze auf dem Berliner Datenportal ist essentiell.

→ **ML3.** Es ist zu untersuchen, ob eine Verfolgbarkeit von Bearbeitungsschritten für aus den Basisdaten abgeleitete Daten und Informationen angeboten werden kann.

8.4 Technische Handlungsempfehlungen

8.4.1 Standards und Sicherheit des Berliner Datenportals

Erst die Bereitstellung offener Daten in bestimmten vereinheitlichten Formaten sowie entlang verschiedener inhaltlicher Kriterien ermöglicht die einfache Weiterverwendung in neuen Anwendungen, die z. B. eine Visualisierung oder anderweitige Aufbereitung der Daten vornehmen.

→ **M11.** Es sind geeignete technische Standards für Open Data zu bestimmen und gegebenenfalls nach Standards und Architekturen für E-Government-Anwendungen (SAGA) zu klassifizieren.

→ **M12.** Es sind relevante Standards zur Gewährleistung der notwendigen Sicherheit des Berliner Datenportals festzulegen und nachhaltig anzuwenden.

8.4.2 Kompatibilität und Integrierbarkeit der Daten

Folgende Maßnahmen werden zur Sicherung einer optimalen Kompatibilität und Integrierbarkeit mit offenen Daten anderer Angebote durch Verwendung internationaler Standards empfohlen, beispielsweise des World Wide Web Consortium (W3C), International Organization for Standardization (ISO) und Comité Européen de Normalisation (CEN).

→ **L1.** Auf dem Berliner Datenportal sind ausschließlich Daten in maschinenverarbeitbaren Formaten anzubieten, die zudem nationalen, europäischen oder internationalen Standards genügen.

→ **L2.** Die Datenangebote des Berliner Datenportals sind um semantisch höherwertige Formate wie Simple Query Language (SQL), RDF etc. anzureichern und mit geeigneten Werkzeugen zur Sichtung, Analyse, Auswertung und Präsentation semantischer Daten zu unterstützen.

→ **ML4.** Es ist je ein Metadatenschema zur Charakterisierung der Datensätze und zur Charakterisierung der Applikationen für das Berliner Datenportal festzulegen. Diese sind in Deutschland, im deutschsprachigen Raum und in Europa abzustimmen.

8.4.3 Qualitätssicherung der Daten

Es sind Maßnahmen zur Qualitätssicherung der Daten und gegebenenfalls der weiterverarbeiteten Daten zu schaffen. Die Datenqualität sollte mittels eines dokumentierten Verfahrens bewertbar sein. Dabei sind beispielsweise die Nachvollziehbarkeit der Quellen, die Aktualität, Korrektheit und Präzision der Basisdaten und der Weiterverarbeitungsschritte zu betrachten.

→ **ML5.** Es ist ein Qualitätsmanagement- und Qualitätssicherungsprozess als Teil des Betriebskonzepts für das Berliner Datenportal festzulegen und konsequent anzuwenden. Eine hohe Qualität des Datenangebots muss kontinuierlich gesichert werden.

→ **M13.** Es sind verschiedene Feedbackmöglichkeiten für die Nutzer des Berliner Datenportals und der Datensätze vorzuhalten, die es beispielsweise erlauben, in einen regen Austausch bezüglich des aktuellen und zukünftigen Datenangebots des Portals zu treten.

→ **L3.** Für die Berliner Datenangebote auf dem Portal ist eine Fehlerkultur zu entwickeln, die die Möglichkeit von Fehlern einplant und u. a. wohldefinierte Korrekturmethode und Korrekturvorschriften umfasst.

8.4.4 Dienste und Werkzeuge für den Umgang mit den Daten

Für eine praktikable, effiziente und barrierefreie Nutzung der Daten durch die breite Öffentlichkeit sind Dienste und Werkzeuge auf dem Portal bereitzustellen.

→ **ML6.** Das Angebot auf dem Portal ist um Dienste und Werkzeuge für die Aggregation, Auswertung und Visualisierung der Basisdaten sowie daraus abgeleiteter Daten und Informationen zu ergänzen.

8.5 Stufenplan

Der Stufenplan fasst die Handlungsempfehlungen in Phasen zusammen und gibt Empfehlungen für nächste Schritte zur Verstetigung und zum Ausbau des Datenangebots in Berlin:

- Kurzfristig (ca. 1 Jahr): Verankerung von Open Data in Gesetzen und Verwaltungsvorschriften des Landes Berlin und Überführung des Berliner Datenportals in den Regelbetrieb. Dazu sind insbesondere die regulatorischen, organisatorischen, technischen und rechtlichen Rahmenbedingungen zu schaffen. Diese Stufe umfasst die kurzfristigen Handlungsempfehlungen K1 bis K12 und die kurz- bis mittelfristigen KM1 bis KM5.
- Mittelfristig (ca. 2 bis 3 Jahre): Nachhaltiger Ausbau der Datenangebote des Landes Berlin und Ergänzung um Werkzeuge, höherwertige Dienste, Schulungen etc. Diese Stufe umfasst die mittelfristigen Handlungsempfehlungen M1 bis M13 und die mittel- bis langfristigen ML1 bis ML6.
- Langfristig (ca. 3 bis 5 Jahre): Abstimmung und Integration der Berliner Datenangebote mit Angeboten in Deutschland, im deutschsprachigen Raum und in Europa. Diese Stufe umfasst die Handlungsempfehlungen L1 bis L3.

Es ist zu beachten, dass einige der Handlungsempfehlungen in den Übergängen zwischen kurz- und mittelfristig bzw. mittel- und langfristig liegen und dementsprechend je nach Bedarf in ihrer zeitlichen Umsetzung angepasst werden können.

Abkürzungen

AfS	Amt für Statistik Berlin-Brandenburg
API	Application Programming Interface, dt. etwa Programmierschnittstelle
ASN.1	Abstract Syntax Notation One
BDSG	Bundesdatenschutzgesetz
BerIDSG	Datenschutzgesetz Berlin
BerIFG	Informationsfreiheitsgesetz Berlin
BODDy	Berlin Open Data Day
CC BY	Creative Commons License mit Namensnennung
CEN	Comité Européen de Normalisation, Europäisches Komitee für Normung
CKAN	Comprehensive Knowledge Archive Network
CMS	Content Management System, dt. System zum Verwalten von Webinhalten
CRUD	Create Read Update Delete, bezeichnet vollumfängliche Bearbeitungsmöglichkeit
CSV	Comma Separated Value, einfaches Format für Tabellen
ETSI	European Telecommunications Standards Institute
GeoZG Bln	Geodatenzugangsgesetz Berlin
GPS	Global Positioning System
HTML	Hypertext Markup Language , die Auszeichnungssprache des WWW
IDL	Interface Definition Language
IKT	Informations- und Kommunikationstechnologie
IETF	Internet Engineering Task Force

ITDZ	IT-Dienstleistungszentrum Berlin
i.S.d.	im Sinne des
ISO	International Organization for Standardization
ITU	International Telecommunication Union
IWG	Informationsweiterverwendungsgesetz, Gesetz über die Weiterverwendung von Informationen öffentlicher Stellen
JSON	Javascript Object Notation, ein Datenaustauschformat
LoR	Lebensweltlich orientierte Räume
NRO	Nicht-Regierungs-Organisation
ODbL	Open Database License
ODF	OASIS Open Document Format
ODP	Open Data-Plattform
OGL	Open Government License
OOXML	Office Open XML
PDF	Portable Document Format
PSI	Public Sector Information, EU-Direktive zu PSI
RDF	Resource Description Framework, populäres Format für Linked Data
Ref-E EGovG Bln	Referenzentwurf E-Governmentgesetz Berlin
REST	Representational State Transfer, ein Schnittstellentyp
RStV	Rundfunkstaatsvertrag
SenBWF	Senatsverwaltung für Bildung, Wissenschaft und Forschung
SAGA	Standards und Architekturen für E-Government-Anwendungen
SPARQL	Simple Protocol and RDF Query Language, Anfragesprache für Linked Data, bzw. RDF
SQL	Simple Query Language, Anfragesprache Datenbanken
TKG	Telekommunikationsgesetz
TMG	Telemediengesetz
UIG	Umweltinformationsgesetz
URL	Uniform Resource Locator, dt. etwa einheitlicher Ressourcenbezeichner, insb. für Webadressen

VIG	Verbraucherinformationsgesetz
VwVfG	Verwaltungsverfahrensgesetz
W3C	World Wide Web Consortium
XLS	Excel Spreadsheet, Speicherformat für Microsoft Excel
XML	Extensible Markup Language
ZLB	Zentral- und Landesbibliothek

Glossar

Abgrenzung Open Government – E-Government

Open Government umfasst jede Aktivität in öffentlichen Verwaltungen, die Transparenz, Teilhabe und Zusammenarbeit fördert, unabhängig davon, ob sie unter Einsatz moderner Informations- und Kommunikationstechnik (IT) erfolgt. Soweit Open Government unter Einsatz moderner Informations- und Kommunikationstechnik (IT) betrieben wird, handelt es sich hierbei auch um eine besondere Ausprägung des E-Government.

Blog, Weblog

Tagebuchartige Internetseite; die Einträge (Postings, Posts) können meist kommentiert werden.

Forum

Eine Plattform (ein »virtueller Platz«), auf der Nutzer Wissen, Erfahrungen, Meinungen und Gedanken austauschen und archivieren können.

Government 2.0

Die Begriffe Open Government und Government 2.0 werden häufig synonym verwendet. Government 2.0 ist ein Sammelbegriff für den Einsatz von Web 2.0-Technologien in der internen und externen Kommunikation von Verwaltungen. Government 2.0 ist nicht zwangsläufig mit Open Government verbunden.

Kosten/Geldleistungspflichten

Der Begriff Kosten stellt eine betriebswirtschaftliche Größe dar; für die hiesige Begriffsdefinition wird der Überbegriff Geldleistungspflichten verwendet. Geldleistungspflichten für die Nutzung von Daten können in Form von Gebühren (öffentlich-rechtlich) oder Entgelten (privatrechtlich) entstehen. Die Weiterverwendung der Daten wird besonders gefördert, wenn den Nutzern keine Geldleistungspflichten entstehen. Ausnahmen können

in haushaltsrechtlichen und haushaltspolitischen Vorgaben begründet sein. Ein hierfür einzurichtendes Entgeltsystem sollte möglichst einfach sein.

Linked Data

Als Linked Data werden strukturierte Informationen bezeichnet, die nicht isoliert interpretiert werden müssen, sondern – wo möglich – auf externe allgemein anerkannte Objekt verweisen. Linked Data ist eine Variante des Sematic Webs, bei der eingeräumt wird, das Begriff u. U. zwar keine allgemeingültige Bedeutung haben, wohl aber einen Bedeutungszusammenhang. Beispiel: Ein Feld in einer Tabelle mit der Zeichenkette »Bezirk« zu bezeichnen wäre nicht Linked Data und hat unterschiedliche Bedeutungen, z. B. in Deutschland und Österreich. Wird dagegen eine URL gewählt wie [http://de.dbpedia.org/resource/Bezirk__\(Berlin\)](http://de.dbpedia.org/resource/Bezirk__(Berlin)), ist eindeutig, was gemeint ist, und man spricht hier von Linked Data.

Lizenzierung

Nutzungsbestimmungen (Lizenzen) legen fest, welche Nutzung unter welchen Voraussetzungen zulässig ist. Klare, einfach handhabbare und einheitliche Nutzungsbedingungen fördern die Weiterverwendung und Weiterverbreitung. Diese Nutzungsbedingungen sind unabhängig von etwaigen Entgeltregelungen. Orientierung können die Standard-Lizenzverträge des Creative Commons geben.

Maschinenlesbar

Grundsätzlich sind alle von Software interpretierbaren Daten maschinenlesbar. Im Zusammenhang mit Open Data werden darunter vor allem solche Datenformate verstanden, die eine Weiterverarbeitung ermöglichen. Die zu Grunde liegende Datenstruktur und entsprechende Standards müssen öffentlich zugänglich sein und sollten vollständig publiziert und kostenfrei erhältlich sein.

Mashup

Einfach Webanwendung, in der verschiedene Datenquellen kombiniert werden.

Microblogging

Bloggen mit Kurznachrichten (z. B. mit dem Dienst Twitter).

Ontologie

Ein Begriffssystem, bei dem Konzepte und deren Beziehungen in einem eindeutigen Vokabular festgelegt werden.

Open (Government) Data

Als Open (Government) Data wird die öffentlich verfügbare Bereitstellung von Datenbeständen der öffentlichen Hand, in der Regel in Form von Rohdaten zur Nutzung, insbesondere durch Weiterverwendung und Weiterverbreitung, bezeichnet. Ausgenommen hiervon sind personenbezogene Daten sowie Daten, die anderweitig schutzwürdig sind (z. B. sicherheitsrelevante Daten, Betriebs- und Geschäftsgeheimnisse).

Open Data kann von Bedeutung für die drei Teilbereiche von Open Government sein. Die Bereitstellung von Daten kann Transparenz erhöhen (z. B. Verwendung von Haushaltsmitteln wird sichtbar), Teilhabe fördern (z. B. durch verbesserte Informationsbasis für eine Mitwirkung im Rahmen eines Bürgerhaushalts) und neue Kooperationsformen ermöglichen (z. B. Entwicklung von Apps auf der Basis von Geodaten). Eine entscheidende Bedeutung hat Open Data für den Teilbereich Transparenz – wobei Transparenz auch ohne Open Data denkbar ist. Der besondere Mehrwert von Open Data liegt vor allem darin, durch Verknüpfung von maschinenlesbaren Daten neue Informationszusammenhänge aufzuzeigen.

Open Data hat damit auch Potential z. B. in den Bereichen Innovations- und Wirtschaftsförderung.

Open Data-Kriterien

Der Diskussion um Open Data liegen die zehn sogenannten Open Data-Kriterien¹ zu Grunde.

- Vollständigkeit
- Primärquelle/Rohdaten
- Zeitnahe Bereitstellung/Echtzeit
- Einfachheit des Zugangs
- Diskriminierungsfreiheit
- Dauerhafte Verfügbarkeit
- Maschinenlesbarkeit

¹Kriterien wurden 2007 von der Nichtregierungsorganisation Sunlight-Foundation als Maximalforderungen an die Bereitstellung von Daten aufgestellt und seither weiterentwickelt.

- Offene Standards/offene Formate
- Lizenzfreiheit
- Kostenfreiheit

Für die Umsetzung von Open Data erscheinen die nachfolgend aufgeführten, an die Open Data-Kriterien angelehnten Aspekte besonders bedeutsam; die Definitionen berücksichtigen die besonderen Rahmenbedingungen in Deutschland.

Open Government

Open Government beschreibt die weitere Öffnung von Regierung und Verwaltung gegenüber Bürgerinnen und Bürgern, Wirtschaft und Wissenschaft. Damit sollen Transparenz, Teilhabe und Zusammenarbeit gefördert werden. Diese Öffnung erfolgt verstärkt mittels moderner Informations- und Kommunikationstechniken (unter anderem Soziale Medien). Die verschiedenen Teilbereiche von Open Government können im Rahmen von Open Government-Projekten unterschiedlich stark ausgeprägt sein. Die Grenzen können fließend sein.

Open Government ist dabei als Oberbegriff der drei Teilbereiche Transparenz, Teilhabe und Zusammenarbeit zu verstehen, die unten definiert werden.

Rohdaten (Primärquelle)

Rohdaten liegen vor, solange die erhobenen Daten nicht akkumuliert und noch nicht bewertet wurden. Im Einzelfall können die Daten bereits bearbeitet sein, um sie handhabbar zu machen oder um rechtliche Vorgaben (z. B. Datenschutz, Statistikgeheimnis) zu erfüllen.

SimpleSearch

»SimpleSearch-Anwendungen sind einfache Mini-Applikationen, die aus einem tabellari-schen Datenbestand mit relativ wenig Aufwand eine Applikation machen kann. Sie kann mit dem SimpleSearch-Baukasten für Imperia eingerichtet werden.

So können z. B. Excel-Tabellen oder andere strukturierte Daten schnell in eine durchsuchbare Datenbank mit Webfrontend verwandelt werden.«²

²<http://support.berlin.de/wiki/index.php/SimpleSearch>.

Soziale Netzwerke

Plattformen zur Pflege sozialer Beziehungen: Menschen können sich untereinander virtuell verbinden (vernetzen) und somit Kontakte knüpfen bzw. pflegen. Soziale Netzwerke ermöglichen üblicherweise auch den Austausch von Nachrichten über einen integrierten Mailedienst.

Teilhabe, Partizipation, Mitwirkung

Teilhabe im Sinne von Open Government bedeutet verbesserte Mitwirkung an staatlichen Entscheidungsprozessen. Wesentliches Ziel ist es, die Gesellschaft stärker in Entscheidungen von Politik und Verwaltung einzubeziehen. Die Form der Mitwirkung kann dabei unterschiedlich stark ausgeprägt sein, von einer Befragung der Bürgerinnen und Bürger zu einem Themenbereich, über eine (E-)Konsultation zu einem Gesetzgebungsvorhaben bis hin zu einer Mitgestaltung der Haushaltsaufstellung im Rahmen des sog. Bürgerhaushalts. Diese Mitwirkungsformen sollen die Akzeptanz von Entscheidungen erhöhen.

Transparenz

Unter Transparenz ist die weitere Offenlegung von Informationen, Meinungsbildungs-, Abwägungs- und Entscheidungsprozessen und die Verfügbarkeit von Daten der öffentlichen Hand zu verstehen. Wesentliches Ziel ist es, Grundlagen von Entscheidungen, die Verwendung von finanziellen Mitteln, die Verantwortlichkeit der Entscheidungsträger usw. offen zugänglich zu machen, um das Handeln von Politik und Verwaltung verständlicher und nachvollziehbarer darzustellen sowie fundierte gesellschaftliche Diskussionen zu fördern.

Web 2.0-Werkzeuge, Soziale Medien

Die Werkzeuge des Web 2.0 zeichnen sich dadurch aus, dass sie eine Kommunikation von Vielen zu Vielen erlauben. Sie sind leicht zu bedienen und interaktiv. Jeder Nutzer kann mitgestalten. Gängige Werkzeuge sind Wikis³, Weblogs⁴, Soziale Netzwerke⁵. Der Begriff Soziale Medien wird weitgehend synonym gebraucht.

³<http://de.wikipedia.org/wiki/Wiki>.

⁴<http://de.wikipedia.org/wiki/Weblog>.

⁵[http://de.wikipedia.org/wiki/Soziales_Netzwerk_\(Internet\)](http://de.wikipedia.org/wiki/Soziales_Netzwerk_(Internet)).

Wiki

Internetseite, die Nutzer lesen und ändern können, ermöglicht gemeinsames Bearbeiten von Texten.

Zusammenarbeit, Kooperation, Kollaboration

Zusammenarbeit beschreibt das vertiefte Zusammenwirken von staatlichen Stellen untereinander sowie von Staat mit Bürgerinnen und Bürgern, Wirtschaft und Wissenschaft bei der Erledigung von Aufgaben (z. B. E-Meldesysteme). Ziel ist es, staatliche Aufgaben in der Servicequalität zu verbessern, das weit verbreitete Wissen in der Gesellschaft zu nutzen und gegebenenfalls den Staat von Aufgaben zu entlasten.

Abgrenzung der Teilbereiche: Die einzelnen Teilaspekte sind in den jeweiligen Open Government-Vorhaben unterschiedlich stark ausgeprägt, z. T. sind mehrere Aspekte berührt, z. T. bedingen sie einander (Beispiel: Im Rahmen eines E-Meldedienstes ist die Mitarbeit der Bürgerinnen und Bürger gefragt (Kooperation); die von der Verwaltung aus der Meldung gezogenen Konsequenzen werden transparent gemacht). Eine klare Abgrenzung ist nicht immer möglich.

Quellenverzeichnis

Aktionsplattform Open Government Data (Hrsg.). (18.5.2011). Berliner Open Data Agenda. <http://berlin.opendataday.de/agenda/>

Almirall, E. (2010). *Open Innovation Mechanisms in Smart Cities* (Description of Work No. 270896). Competitiveness and Innovation Framework Programme.

Bangert, D. (2010). Planungsraumbezogene Informationsbereitstellung für Monitoring und Analyse (PRISMA) als Umsetzungsbaustein der Rahmenstrategie Soziale Stadtentwicklung Berlin – Auf dem Weg vom sektoralen zum vernetzten Verwaltungshandeln. In Beuth Hochschule für Technik Berlin (Hrsg.), *Innovative Entwicklungen für Freizeit-anlagen der Bildung und Erholung BAER2FIT-Projekt der Beuth Hochschule für Technik Berlin* (S. 107-109). Berlin. http://projekt.beuth-hochschule.de/fileadmin/projekt/baer2fittagung/BAER2FIT_Tagungsband_2010.pdf

Barnickel, N., Flügge, M., Schieferdecker, I., & Hager, S. (2010). Vorstudie zur City Data Cloud Berlin. Fraunhofer Institut für Offene Kommunikationssysteme. http://www.fokus.fraunhofer.de/de/motion/_media/CDC_Vorstudie.pdf

BDSG. Bundesdatenschutzgesetz in der Fassung der Bekanntmachung vom 14. Januar 2003 (BGBl. I S. 66), das zuletzt durch Artikel 1 des Gesetzes vom 14. August 2009 (BGBl. I S. 2814) geändert worden ist.

BerlDSG. Gesetz zum Schutz personenbezogener Daten in der Berliner Verwaltung (Berliner Datenschutzgesetz). In der Fassung der Bekanntmachung vom 17. Dezember 1990. (GVBl.1991 S. 16, ber. S. 54), BRV 205-1. Zuletzt geändert durch Art. I Viertes ÄndG vom 2. 2. 2011 (GVBl. S. 51).

BerlIFG. Gesetz zur Förderung der Informationsfreiheit im Land Berlin (Berliner Informationsfreiheitsgesetz) vom 15. Oktober 1999 (GVBl. S. 561), zuletzt geändert durch Gesetz vom 8. Juli 2010 (GVBl. S. 358).

Berliner Senatsverwaltung für Wirtschaft, Technologie und Frauen (Hrsg.). (2010). Online Voting zu Open Data – Auswertung der Online-Umfrage auf Berlin.de. http://www.berlin.de/projektzukunft/fileadmin/user_upload/pdf/pr%C3%A4sentationen/auswertung_Online_Voting_zu_Open_Data.pdf

Berliner Senatsverwaltung für Wirtschaft, Technologie und Frauen. (24.2.2011). Projekt Zukunft: Pressemitteilung: Apps4Berlin-Wettbewerb kürt Sieger.

<http://www.berlin.de/projektzukunft/presse/pressearchiv/detailseite/datum/2011/02/24/pressemitteilung-apps4berlin-wettbewerb-kuert-sieger>

Bömermann, H., Jahn, S., Nelius, K. (2006) Lebensweltlich orientierte Räume im Regionalen Bezugssystem (Teil 1) in: Berliner Statistik, Monatsschrift 8/06

Bundesministerium des Innern (Hrsg.). (2010). Regierungsprogramm – Vernetzte und transparente Verwaltung. http://www.verwaltung-innovativ.de/cln_117/SharedDocs/Publikationen/Pressemitteilungen/regierungsprogramm__vernetzte__und__transparente__verwaltung,templateId=raw,property=publicationFile.pdf/regierungsprogramm_vernetzte_und_transparente_verwaltung.pdf

Büchting, H., Heussen, B. (Hrsg.). Beck'sches Rechtsanwalts-Handbuch. 10. Auflage, München 2011, Verlag C.H.Beck

Bündnis 90/Die Grünen, & SPD Baden-Württemberg (Hrsg.). (2011). Der Wechsel beginnt. Koalitionsvertrag zwischen Bündnis 90/Die Grünen und SPD Baden-Württemberg. Baden-Württemberg 2011-2016. <http://www.gruene-bw.de/fileadmin/gruenebw/dateien/Koalitionsvertrag-web.pdf>

Cooperation OGD Österreich. (2011). Abschlussbericht der Arbeitsgruppe Metadaten. http://www.ref.gv.at/uploads/media/OGD-Metadaten_1-0_2011024.pdf

Dekkers, M., Polman, F., te Velde, R., & de Vries, M. (2006). MEPSIR - Measuring European Public Sector Information Resources. Final Report of Study on Exploitation of public sector information-benchmarking of EU framework conditions. (European Commission, Hrsg.).

Drescher, S. (16.9.2011). Wir brauchen eine Kanzlerinnen-Initiative. *politik-digital.de*. <http://politik-digital.de/wir-brauchen-eine-kanzlerinnen-initiative>

EU Presidency (Hrsg.). (18.11.2009). Ministerial Declaration on eGovernment. <http://www.egov2009.se/wp-content/uploads/Ministerial-Declaration-on-eGovernment.pdf>

European Institute for Public Administration (Hrsg.). (2011). EPSA 2011 Best Practice Certificate Recipients. <http://www.epsa2011.eu/files/Final%20Results/43%20BPC%20Recipients%20table%20final.pdf>

Florini, A. (2007). The right to know: transparency for an open world. Columbia University Press.

Fornfeld, M., Boele-Keimer, G., Recher, S., & Fanning, M. (2009). Assessment of the Reuse of Public Sector Information (PSI) in the Geographical Information, Meteorological Information and Legal Information Sectors. (MICUS Management Consulting GmbH, Hrsg.).

GeoZG. Gesetz über den Zugang zu digitalen Geodaten. (2009). BGBl. I S. 278. <http://www.gesetze-im-internet.de/bundesrecht/geozg/gesamt.pdf>

GeoZG Bln. Gesetz über den Zugang zu digitalen Geodaten im Land Berlin – Geodatenzugangsgesetz Berlin v. 12. 12. 2009, GVBl. Vom 12.12.2009 S. 682.

Glock, C., & Broens, M. (2008). Wissensmanagement in der öffentlichen Verwaltung mithilfe von Social-Software-Applikationen. *Verwaltung & Management – Zeitschrift für allgemeine Verwaltung*, 4(5), S. 273-276.

Government 2.0 Netzwerk Deutschland e. V., Open Data Network e. V., & Open Knowledge Foundation Deutschland e. V. (Eds.). (13.7.2011). Informationsblatt “Apps für Deutschland” – Der Wettbewerb für offene Daten, Anwendungen und Ideen. http://apps4deutschland.de/docs/Infoblatt_apps4de.pdf

Grabow, B., Knipp, R., & Schneider, S. (2009). *ServiceStadt Berlin 2016[202F?]: Anforderungen an die zukünftige Ausgestaltung der Leistungs- und Serviceangebote im Land Berlin unter sich verändernden Rahmenbedingungen*. Difu-Impulse[202F?]; 2009,3. Berlin: Dt. Inst. für Urbanistik.

HM Treasury (Hrsg.). (2011). Autumn Statement 2011. http://cdn.hm-treasury.gov.uk/autumn_statement.pdf.

Hoeren, T., Sieber, U. (Hrsg.). Handbuch Multimedia-Recht. 28. Ergänzungslieferung April 2011. München. Verlag C.H. Beck.

Holtkamp, L. (2006). Partizipative Verwaltung – hohe Erwartungen, ernüchternde Ergebnisse. In J. Bogumil, W. Jann, & F. Nullmeier (Hrsg.), *Politik und Verwaltung* (S. 185-207). Wiesbaden: Verlag für Sozialwissenschaften.

IFG. Gesetz zur Regelung des Zugangs zu Informationen des Bundes (Informationsfreiheitsgesetz). Vom 5. September 2005. (BGBl. I S. 2722). FNA 201-10.

IWG. Gesetz über die Weiterverwendung von Informationen öffentlicher Stellen (Informationsweiterverwendungsgesetz). Vom 13. Dezember 2006 (BGBl. I S. 2913) FNA 772-3.

Kilian, W., Heussen, B. (Hrsg.). Computerrechts-Handbuch. 29. Ergänzungslieferung Februar 2011. München. Verlag C.H. Beck.

Klessmann, J., Schulz, S. E., Hoffmann, C., Warnecke, T., Graudenz, D., & Krug, B. (2010). Vom Open Government zur Digitalen Agora – Die Zukunft offener Interaktionen und sozialer Netzwerke im Zusammenspiel von Politik, Verwaltung, Bürgern und Wirtschaft. ISPRAT.

Kubicek, H. (2007). *Electronic Democracy and Deliberative Consultation on Urban Projects – “Putting E-Democracy into Context”* (Report for the Congress of Local and Regional Authorities). http://www.ifib.de/publikationsdateien/Creative_final.pdf

Kubicek, H. (2008). Next generation FoI between information management and web 2.0. *Proceedings of the 2008 international conference on Digital government research*, dg.o '08 (S. 9–16). Digital Government Society of North America. <http://portal.acm.org/citation.cfm?id=1367832.1367838>

- Land Brandenburg, & Land Berlin (Hrsg.). (12.12.2008). Masterplan für den Aufbau der Geodateninfrastruktur Berlin/Brandenburg (GDI-BE/BB). http://geoportal.brandenburg.de/fileadmin/user_upload/unterlagen/be_bb/masterplan_v_1.pdf
- Landmann/Rohmer, Band 1. Umweltrecht, 62. Ergänzungslieferung 15. Juli 2011. München. Verlag C.H. Beck. Hrsg.: Beckmann, M., Durner, W., Mann, T., Röckinghausen, M.
- Leupold, A, Glossner, S. (Hrsg.). Münchener AnwaltsHandbuch IT -Recht. 2 .Auflage. München, Verlag C.H. Beck.
- Linnert, K., Kubicek, H., & Sommer, I. (18.1.2011). Bremer Empfehlung zu Open Government Data – Ein elektronischer Weg zu besserem Informationszugang und mehr Transparenz in der öffentlichen Verwaltung. http://www.finanzen.bremen.de/sixcms/media.php/13/2011-01-17_BremerErklaerung.pdf
- Loewenheim, U. (Hrsg.). Handbuch des Urheberrechts. 2 .Auflage. München 2011, Verlag C.H. Beck.
- Newbery, D., Bently, L., & Pollock, R. (2008). Models of Public Sector Information Provision via Trading Funds. (Office for Public Sector Information (OPSI), United Kingdom, Hrsg.). <http://www.opsi.gov.uk/advice/poi/models-psi-via-trading-funds.pdf>
- NRW SPD, & Bündnis 90 / Die Grünen NRW (Hrsg.). (2010). Gemeinsam neue Wege gehen – Koalitionsvertrag zwischen der NRW SPD und Bündnis 90 / Die Grünen NRW. http://www.gruene-nrw.de/fileadmin/user_upload/landesverband/gruene-nrw/aktuelles/2010/koalitionsvertrag/Koalitionsvertrag_Rot-Gruen_NRW_2010-2015.pdf
- Office of Fair Trading (Hrsg.). (2006). The commercial use of public information (CUPI). http://www.oft.gov.uk/shared_oftr/reports/consumer_protection/oft861.pdf
- Open Knowledge Foundation. (2010). Open Definition – Defining the Open in Open Data, Open Content and Open Services. *Open Definition*. <http://www.opendefinition.org/okd/>
- Piotrowski, S. J., & Van Ryzin, G. G. (2007). Citizen Attitudes Toward Transparency in Local Government. *The American Review of Public Administration*, 37(3), 306-323. doi:10.1177/0275074006296777
- Pollock, R. (2009). *The Economics of Public Sector Information* (Cambridge Working Papers in Economics). Faculty of Economics, University of Cambridge. <http://econpapers.repec.org/RePEc:cam:camdae:0920>
- Ref-EGovG Bln. Entwurf eines Gesetzes über E-Government- und Organisationsfragen. http://www.berlin.de/imperia/md/content/verwaltungsmodernisierung/publikationen/110620_stand_19._sts_a_synopse_egovgesetz.pdf?start&ts=1308814816&file=110620_stand_19._sts_a_synopse_egovgesetz.pdf

Relly, J. E., & Sabharwal, M. (2009). Perceptions of transparency of government policymaking: A cross-national study. *Government Information Quarterly*, 26(1), 148-157. doi:10.1016/j.giq.2008.04.002

Richtlinie 2007/2/EG des Europäischen Parlaments und des Rates vom 14.3.2007 zur Schaffung einer Geodateninfrastruktur in der Europäischen Gemeinschaft. (2007). *RICHTLINIE 2007/2/EG*. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:108:0001:0014:DE:PDF>

RStV. Staatsvertrag für Rundfunk und Telemedien (Rundfunkstaatsvertrag) in der Fassung des dreizehnten Staatsvertrags zur Änderung rundfunkrechtlicher Staatsverträge (13. RÄStV) vom 30.10. / 04.11. / 20.11.2009. Inkrafttreten: 01. April 2010.

Schenk, F., Kruse, F., & Klenke, M. (2010). From Simple Data Sources to a Complex Information System: Integrating Heterogeneous Data Models into an Information Infrastructure for the Public Administration. Presented at the 24th International Conference on Informatics for Environmental Protection in cooperation with Intergeo 2010, Integration of Environmental Information in Europe, Bonn. http://www.kst.portalu.de/pub/praes/2010/praes_enviroinfo_bonn_sk_10-2010.pdf

Schmidt, M. G. (2008). Beteiligungszentrierte Demokratietheorien. *Demokratietheorien* (4. Aufl., S. 236-253). Wiesbaden: Verlag für Sozialwissenschaften.

Sebastopol Group. (12.7.2007). 8 Principles of Open Government Data. http://resource.org/8_principles.html

Senatsverwaltung für Inneres und Sport. (2010). Modernisierungsprogramm Service-Stadt Berlin. 3. Bericht an den Hauptausschuss des Abgeordnetenhauses von Berlin. http://www.berlin.de/imperia/md/content/verwaltungsmodernisierung/publikationen/101105_3_bericht_an_ha_zum_31.12.10_internetfassung.pdf?start&ts=1297418712&file=101105_3._bericht_an_ha_zum_31.12.10_internetfassung.pdf

Senatsverwaltung für Stadtentwicklung Berlin (Hrsg.). (2009). Ergebnisbericht 2009. Teil B. Handbuch zur Sozialraumorientierung Grundlage der integrierten Stadt(teil)entwicklung Berlin. http://www.stadtentwicklung.berlin.de/soziale_stadt/sozialraumorientierung/download/TeilB_Ergebnisbericht_RSSE09.pdf

Shadbolt, N., O'Hara, K., Salvadores, M., & Alani, H. (2011). eGovernment. *Handbook of Semantic Web Technologies* (S. 840-900). Springer-Verlag. <http://eprints.ecs.soton.ac.uk/21711/>

Spanish EU Presidency. (19.4.2010). Granada Ministerial Declaration on the European Digital Agenda. http://www.eu2010.es/export/sites/presidencia/comun/descargas/Ministerios/en_declaracion_granada.pdf

SPD Berlin. (2011). Berlinprogramm 2011-2016. (Sozialdemokratische Partei Deutschlands (SPD), Landesverband Berlin, Hrsg.). http://www.spd-berlin.de/w/files/spd-parteitage/berlinprogramm_2011-2016.pdf

SPD Berlin & CDU Berlin (Hrsg.). (2011). Koalitionsvereinbarung zwischen der Sozialdemokratischen Partei Deutschlands, Landesverband Berlin und der Christlich Demokratischen Union, Landesverband Berlin. Vorlage für die Beratung des Landesparteitages der SPD am 21. November 2011. <http://www.tagesspiegel.de/downloads/5859422/1/Koalitionsvereinbarung%20zwischen%20SPD%20und%0CDU>.

Sunlight Foundation. (8.11.2010). Ten Principles for Opening Up Government Information. *Sunlight Foundation*. <http://sunlightfoundation.com/policy/documents/ten-open-data-principles/>

UIG. Umweltinformationsgesetz. (2004). BGBl. I S. 3704. http://www.gesetze-im-internet.de/bundesrecht/uig_2005/gesamt.pdf

von Lucke, J., & Geiger, C. (2010). *Open Government Data – Frei verfügbare Daten des öffentlichen Sektors* (Gutachten für die Deutsche Telekom AG zur T-City Friedrichshafen). Deutsche Telekom Institute for Connected Cities, Zeppelin University Friedrichshafen.

VwVfG. Verwaltungsverfahrensgesetz. In der Fassung der Bekanntmachung vom 23. Januar 2003 (BGBl. I S. 102) FNA 201-6. Zuletzt geändert durch Art. 2 Abs. 1 G zur Modernisierung von Verfahren im patentanwaltl. Berufsrecht vom 14. 8. 2009 (BGBl. I S. 2827).

Anhang

A Interviewleitfaden der Organisationsanalyse

Interview-Leitfaden

Berliner Open Data Strategie

Hintergrundinformationen

Im Rahmen der BODS werden eine Reihe von Interviews mit Verantwortlichen beteiligter Behörden durchgeführt. Das Ziel dieses Interview-Leitfadens ist eine Sammlung von Fragen anhand derer wir den Status Quo und die Erwartungen bezüglich einer Öffnung der Datensätze der Berliner Verwaltung bestimmen können. Dabei werden die Fragen uns eine fundierte Einschätzung der Möglichkeiten zur Einführung elektronischer Prozesse zur Erfassung, Verarbeitung und öffentlichen Freigabe von Datensätzen ermöglichen. Falls einzelne Fragen oder komplette Themenkomplexe nicht zutreffen, so brauchen diese nicht beantwortet zu werden.

Allgemeine Informationen zum Interviewpartner

Name:

Kontaktdaten:

Organisationszugehörigkeit:

Position/Aufgabenbereich:

Themenkomplexe der Fragen

Allgemeine Informationen zu Datensätzen

- Welche Datensätze stellen Sie öffentlich zur Verfügung?
- Welche Datensätze würden Sie potentiell zur Verfügung stellen wollen?
- Könnten Sie diese Datensätze bitte jeweils mit wenigen Sätzen charakterisieren?

Strategische Rahmenbedingungen in Berlin

- Welche Vor- und Nachteile, bzw. Chancen und Risiken sehen Sie bei der Öffnung Ihrer Datensätze in Bezug auf folgende vier Bereiche?
 - Zivilgesellschaft
 - Verwaltung
 - Wirtschaft
 - Medien
- Aus welchen Gründen stellen Sie Daten zur Verfügung?
 - Rechtliche Verpflichtungen
 - Anderweitige Motivation
- Gibt es in Ihrem Haus ein strukturiertes Vorgehen zur strategischen Bereitstellung von Datensätzen?
 - Gibt es eine schriftlich fixierte Strategie?
- Welche Gesetze, Verordnungen, Arbeitsanweisungen oder sonstige Regelungen haben einen Einfluss auf Ihren Umgang mit Datensätzen?
 - Gibt es eine Gebührenordnung für die Bereitstellung von Datensätzen?
- Könnten Sie diese bitte beschreiben, bzw. uns Material dazu zur Verfügung stellen?

Identifikation und Struktur von Datensätzen

- Können Sie die Anzahl öffentlicher Datensätze aus Ihrem Haus, sowie deren Größe, schätzen?
 - Wird sich dies in absehbarer Zukunft ändern?
- Wie werden Datensätzen bei Ihnen gespeichert (z.B. Excel oder Datenbanken)?
 - In welchen Formaten liegen die Datensätze vor?
 - Welche Metadaten erfassen sie?
- Welche Werkzeuge verwenden sie zur Erhebung, Verwaltung und Bereitstellung von Datensätzen?
 - Konvertieren sie häufig Datensätze zwischen verschiedenen Formaten?
 - Gibt es andere Werkzeuge, die Sie gerne benutzen würden?
- Haben Ihre Datensätze Raumbezug?
 - In welcher Granularität (Stadt, Bezirk, Straße, Punkt)
- Wie wichtig ist die Aktualität der bereitgestellten Daten?
 - Was sind typische Zeitrahmen für die Aktualität der Daten?
 - Wie weit gehen die Datensätze zurück?
 - Geht es nur um aktuelle Daten oder auch um historische Datensätze?
 - In welchem Zyklus findet eine Veröffentlichung von Daten statt?
 - Wie oft aktualisieren Sie Daten?

Architektur- und Plattformfragen

- Speichern sie Daten innerhalb Ihres Hauses oder benutzen sie zentrale Systeme?
 - Dürften die Daten außerhalb Ihres Hauses gespeichert werden?
- Wie stellen Sie Datensätze zur Verfügung?
 - Welche Medien benutzen Sie dazu?
 - Ist Ihnen bekannt wie Datensätze aus Ihrem Haus weiterverarbeitet werden?
- Welche Technologien / Produkte benutzen Sie intern?
- Wie erfolgt die Zugangs- und Zugriffskontrolle auf die verschiedenen IT-Systeme, die zur Verarbeitung und Publikation von Datensätzen verwendet werden?

Nutzergruppen

- Welche Gruppen schätzen Sie als hauptsächliche Nutzer der von ihnen veröffentlichten Datensätze ein?
- Warum verwenden diese Nutzergruppen Ihre Datensätze?
- Wie wird typischerweise mit den Datennutzern kommuniziert (z.B. Feedback)?
- Ist neben der Veröffentlichung auch eine interne Verwendung der Datensätze beabsichtigt?

Prozesse

- Welcher Ihrer Prozesse der Erhebung, Verarbeitung oder Veröffentlichung von Daten ist am repräsentativsten für die Arbeit in Ihrem Haus?
- Bitte beschreiben Sie diesen Prozess.
- Wie viele Mitarbeiter/innen sind an Erstellung und Bereitstellung der Daten beteiligt?
 - In welcher Rolle geschieht dies jeweils und mit welcher Verantwortung?
- Könnten Sie schätzen ob diese Mitarbeiter/innen grundlegende Kenntnisse besitzen im Umgang mit
 - Office Produkten?
 - Dem World Wide Web?
 - Content Management Systemen, Wikis und Blogs?
 - Datenbanksystemen?
 - Email?

- Sozialen Medien (z.B. Facebook und Twitter)
 - Chat-Programmen?
 - Moderation von Diskussionsforen?
- Könnten Sie uns Ansprechpartner nennen mit denen Sie im Rahmen dieses Prozesses zusammenarbeiten?
- Wie könnte dieser Prozess durch Einsatz von Internettechnologie verbessert werden?
- Durchlaufen zu veröffentlichende Datensätze einen Freigabeprozess?
 - Wird dieser durch technische Systeme (z.B. Workflow-Portale) unterstützt?
- Gibt es in Ihrem Haus eine Dokumentation der Datenerhebungs- und Datenverarbeitungsschritte?
- Wie sichern sie die Datenqualität, Konsistenz und Richtigkeit der Daten?
 - Wie nehmen sie nachträglich Korrekturen an Datensätzen vor?
 - Wie sieht die Qualitätssicherung für Dokumente aus, die auf Webseiten (z.B. Berlin.de) veröffentlicht werden?
- Gibt es mehrere Kopien von Datensätzen in Ihrem Prozess/System?
 - Welche Kopie/Instanz definiert die „richtige“ Version eines Datensatzes?
- Beinhaltet Ihr Datenverarbeitungsprozess einen Schritt der Datenarchivierung?

Erfahrungswerte und Erwartungen

- Welche Schwierigkeiten haben Sie derzeit bei der Veröffentlichung Ihrer Datensätze?
 - Wie könnten auf diese Schwierigkeiten eingegangen werden?
- Haben Sie Erfahrungen mit ähnlichen Projekten in der Vergangenheit?
 - Haben Sie bereits Projekte zur Veröffentlichung von Datensätzen durchgeführt?
 - Wie waren Ihre Erfahrungen hierbei?
- Haben Sie generelle Hinweise zur Öffnung von Daten der Berliner Verwaltung?
- Welche Punkte würden Sie gerne adressiert sehen, wenn für Berlin ein Konzept für offene Daten erarbeitet wird?
- Gibt es Dinge, die mit den bisherigen Fragen nicht angesprochen wurden?

B FAQ – Häufig gestellte Fragen

→ **An wen kann ich mich bei Fragen, Kommentaren, Ideen oder Beschwerden wenden?**

Im Regelfall können Sie direkt unter den entsprechenden Inhalten wie Datensätzen oder Ankündigungen kommentieren. Für Themen, die bisher noch nicht auf dem Portal diskutiert werden, haben wir die Bereiche Fragen bzw. Feedback vorgesehen. Dort können Sie als angemeldeter Nutzer des Portals Inhalte erstellen und so mit uns in Kontakt treten.

Für Fragen, die nicht öffentlich diskutiert werden können oder sollen, können Sie unter Kontaktformular benutzen.

→ **Entstehen mir durch die Bereitstellung Kosten?**

Die Bereitstellung der Daten erfolgt in der Regel für die Nutzer dieses Angebots kostenlos.

Nur in Ausnahmefällen (beispielsweise für die technische Bereitstellung) könnten minimale Gebühren erhoben werden.

→ **Existieren in Deutschland Angebote um z.B. auf Daten auf Bundesebene zuzugreifen?**

In dem Portal <http://offenedaten.de> können Behörden, Verwaltungen, andere Institutionen sowie jeder Bürger Informationen zu Daten bereitstellen. Daraus ergibt sich ein Katalog, der kostenlos nach offenen Daten durchsucht werden kann.

Existierende Angebote der öffentlichen Verwaltung in Deutschland sind bislang vorwiegend Fachangebote. So stellen die Statistikämter auf allen föderalen Ebenen Daten zur Verfügung. Ebenso tun dies viele Umweltbehörden. Explizite Portale für offene Daten, wie dieses, gibt es bislang kaum in Deutschland.

→ **Gibt es Beispiele für die Weiterverwendung von Daten der Berliner Verwaltung?**

Beispiele für weiterverarbeitete Datensätze für verschiedene Bereiche können Sie sich unter Anwendungen ansehen.

→ **Gibt es ein Forum oder eine Plattform für Softwareentwickler, um Ideen, Applikationen oder die Weiterverwendung von Ideen zu diskutieren?**

Sie können entweder Fragen stellen oder Feedback abgeben. In beiden Bereichen gibt es die Möglichkeit über die verschiedenen Themen zu diskutieren.

→ **Ich bin Mitarbeiter einer Berliner Behörde und würde gerne weitere Daten über www.daten.berlin.de zur Verfügung stellen.**

Wenn Sie Zugriff auf das Content Management System des Landes haben, gibt es verschiedene Möglichkeiten, Datensätze zu publizieren. Bei Fragen hilft Ihnen Ihr CVD, oder der Support von BerlinOnline weiter.

→ **Ich möchte die Daten kommerziell weiterverwenden. Ist dies möglich?**

Eine kommerzielle Nutzung der bereitgestellten Daten ist abhängig von den individuellen Nutzungsbedingungen der einzelnen Datensätze. Generell wird angestrebt, eine kommerzielle Nutzung zu ermöglichen. Sollte das bei einem Sie interessierenden Datensatz nicht der Fall sein, sind wir gerne Bereit nach einer Lösung zu suchen.

→ **In welchen Formaten sind die Daten vorhanden?**

Die Daten sind größtenteils in den Formaten XLS, HTML oder PDF verfügbar. Aber auch auf Kartenmaterial (KML) und webbasierte Datenbanken kann derzeit über das Berliner Datenportal zugegriffen werden.

→ **Mir sind Datensätze der Berliner Verwaltung bekannt, die hier nicht verzeichnet sind, aber anderweitig bereits verfügbar.**

Hinweise zu weiteren Datensätzen sind stets willkommen. Im Bereich Fragen bzw. Feedback können Sie uns auf solche Datensätze aufmerksam machen. Was ist das Berliner Datenportal

Im Berliner Datenportal stellt das Land Berlin Datensätze der öffentlichen Verwaltung zur Verfügung. So soll Verwaltungsmitarbeitern, Bürgern, Unternehmen und Wissenschaftler die Möglichkeit gegeben werden, über einen zentralen Einstiegspunkt auf Daten und Informationen der Berliner Verwaltung zuzugreifen und diese weiterzuverwenden, so dass durch neue Ideen sowie Kombination und Analyse neue Erkenntnisse aus den vorhandenen Daten gewonnen werden können.

Dieses Pilotprojekt befindet sich z.Z. noch im Aufbau.

→ **Was ist die Idee hinter Open Data?**

Informationen sind der Rohstoff des 21. Jahrhunderts. Der offene und strukturierte Zugang zu freiverfügbaren Datenbeständen der öffentlichen Hand ist ein wichtiger Beitrag für die Weiterentwicklung der Wissensgesellschaft. Er stärkt das Vertrauen zwischen Politik und Zivilgesellschaft, zwischen Verwaltung, Wirtschaft und Medien. Staatliches Handeln wird transparent, nachvollziehbar und überprüfbar. Gleichzeitig birgt die Öffnung Potentiale für effizienteres Handeln und Qualitätsgewinn in sich.

→ **Welche Daten werden nicht veröffentlicht?**

Personenbezogene und sensitive Daten werden aus datenschutzrechtlichen Gründen sowie aus Gründen der Sicherheit nicht veröffentlicht.

→ **Welche Datensätze sind über www.daten.berlin.de verfügbar?**

Die Verwaltung des Landes Berlin stellt Informationen zu vorhandenen Datensätzen aus verschiedenen Bereichen (z.B. Umwelt, Bildung, Statistik) bereit. Es ist eine Vielzahl an unterschiedlichen Daten vorhanden. So finden Sie im Berliner Datenportal neben Statistiken und Jahresberichten, beispielsweise auch Karten der Stadt Berlin, Wahlergebnisse oder Datenbanken, die es Ihnen ermöglichen, über Suchkriterien passende Einrichtungen zu finden.

→ **Welche Programme kann ich nutzen, um die Daten weiter zu verarbeiten?**

Natürlich steht es Ihnen frei, welche Programme Sie für die Weiterverarbeitung verwenden wollen. Falls Sie aber noch nach passenden Anwendungen suchen, finden Sie unter auf diesem Portal auch Anwendungen, mit denen sich die Daten weiterverarbeiten lassen.

→ **Wer ist für das Berliner Datenportal verantwortlich?**

Das Berliner Datenportal ist Bestandteil von Berlin.de, dem offiziellen Internetauftritt des Landes Berlin. Der Betrieb erfolgt in Public-Private-Partnership mit der BerlinOnline Stadtportal GmbH & Co. KG. Die inhaltliche Verantwortung liegt derzeit bei der Senatsverwaltung für Wirtschaft, Technik und Frauen.

→ **Werden zukünftig weitere Datensätze zur Verfügung gestellt?**

Natürlich gestaltet es sich als schwierig, alle vorhandenen Daten auf einen Schlag zu sichten und zu veröffentlichen. Das Land Berlin mit seinen öffentlichen Einrichtungen bemüht sich sukzessive und schnellstmöglich so viele Daten wie möglich bereitzustellen.

→ **Wie ist das Berliner Datenportal entstanden?**

Der Berliner Senat für Wirtschaft, Technologie und Frauen hat bei Fraunhofer FOKUS im Frühjahr 2011 die Berliner Open Data Studie beauftragt. Neben der Untersuchung strategischer, organisatorischer und rechtlicher Fragestellungen lag der Schwerpunkt auf der Entwicklung und pilothaften Realisierung einer technischen Architektur für ein Berliner Open Data Portal.

→ **Wo finde ich die Nutzungsbedingungen einzelner Datensätze?**

Hinweise zu den individuellen Nutzungsbedingungen werden in den Metainformationen zu jedem einzelnen Datensatz gegeben. Detaillierte Beschreibungen der Nutzungsbedingungen eines Datensatzes finden sich in der Regel an der Stelle, von der der Datensatz bezogen werden kann.

→ **Wo finde ich mehr Informationen zu Open Data Projekten?**

Weitere Informationen zum Thema Open Data finden sie unter anderem unter:

Einen Überblick über verschiedene Open Data Projekte erhalten Sie unter: <http://www.opendata-showroom.org>

Desweiteren fand am 18.05.2011 in der Berliner Senatsverwaltung für Wirtschaft, Technologie und Frauen der Berlin Open Data Day statt. Präsentationen und Ergebnisse finden sie unter: <http://berlin.opendataday.de/>

Die Berliner Open Data Agenda finden Sie unter: <http://berlin.opendataday.de/agenda>

→ **Wo kann ich Ideen für die Weiterverwendung von Datensätzen einbringen?**

Eigene Ideen, welche weiteren Datensätze bereitgestellt werden oder wie Datensätze weiterverwendet werden sollten können Sie unter Feedback oder Fragen einstellen.

C Linksammlung

Open Data Agenda

- Berlin Open Data Agenda
<http://berlin.opendataday.de/agenda>

Daten-Seiten

- Haushalt des Bundes
<http://bund.offenerhaushalt.de/>
- Aktuelle Messwerte der Radioaktivität vom Bundesamt für Strahlenschutz
<http://odlinfo.bfs.de/>
- Daten zu Bildung und Forschung aus dem Bundesministerium für Bildung und Forschung (BMBF)
<http://www.bmbf.de/daten-portal/index.html>
- Radioaktivität weltweit von privaten Messstationen (mit Zeitreihen)
http://you-measure.com/index.php?v_ort=Brandenburg

Daten-Seiten, nicht-offen

- Aktuelles Wetter und Vorhersagen von der Freien Universität Berlin
<http://wind.met.fu-berlin.de/wind/main.php>
- Georeferenzierung von Datensätzen der öffentlichen Hand am Beispiel eines kommunalen Haushalts
http://www.corp.at/archive/CORP2010_15.pdf

Ereignisse

- Berlin Open Data Day am 18.05.2011
<http://berlin.opendataday.de/>

Open Data-Portale

- Das Meta-Portal für offene Daten
<http://ckan.net/>
- Datenportal der Regierungsdaten der britischen Regierung
<http://data.gov.uk/>
- Datenportal der Regierungsdaten der Londoner Regierung
<http://data.london.gov.uk/>
- Datenportal der Stadt Wien
<http://data.wien.gv.at/>
- Datenportal der Stadt Linz
<http://www.data.linz.gv.at>
- Meta-Portal für offene Daten in Europa
<http://publicdata.eu>
- API Leipzig: Standardisierte Schnittstelle für den Zugriff auf öffentliche Daten
<http://www.apileipzig.de/>
- Datenportal der Stadt Barcelona
<http://www.bcn.cat/opendata/>
- Register für offene Datenbestände in Deutschland mit Suchfunktion
<http://www.offenedaten.de/>
- Übersicht zu Open Data Projekten in Berlin und Brandenburg
<http://www.openberlin.net/>
- Datenportal der Stadt Hongkong
<http://www.gov.hk/en/theme/psi/welcome/>
- Datenportal der Weltbank
<http://data.worldbank.org/>
- Datenportal der Vereinten Nationen
<http://data.un.org/>

Organisationen

- Creative Commons: Benutzerfreundliche Lizenzen für offene Inhalte
<http://de.creativecommons.org/>
- Open Knowledge Foundation, Stiftung zur Förderung des freien Wissens
<http://okfn.org/>

Projekte

- Das Tool für die zentrale Registrierung, Pflege und Föderation von Datensätzen
<http://ckan.org/>
- Offene Geodaten, die kollaborativ verbessert werden („Wikipedia“ der Landkarten)
<http://www.openstreetmap.org/>

D Metadaten-Schema und API

Schnittstellenbeschreibung zum Open-Data-Register v1.2

Das Open-Data-Register verfügt über RESTful-Dienstschnittstellen zur Registrierung von neuen Metadatenätzen und zum Export der Metadatenätze an das Open-Data-Portal.

Spezifikation Input-API.....	2
Spezifikation Output-API.....	7

Schnittstellenüberblick

- Open-Data-Register **Input-API** - Die Input API dient als Schnittstelle zwischen Metadatenformular im Imperia-CMS von berlin.de und dem Open Data Register zum Eintragen von Metadaten zu einem Datensatz.
- Open-Data-Register **Output -API** - Die Output API dient als Schnittstelle zwischen dem Open Data Register und dem Open Data Portal zum Export der Metadaten in das Portal.

Versionierung

- Es wird jeweils nur eine verwendbare Version der Input- und Output-API bereitgestellt. Die aktuelle Version ist in diesem Dokument beschrieben und wird kontinuierlich weitergepflegt.

Hinweise zu Kontakt/Author/Maintainer:

Input API	CKAN API	Pflicht	Bedeutung
veroeffentlichende_stelle	author	ja	Die Abteilung, bzw. das Amt das den Datensatzveröffentlicht.
veroeffentlichende_person	extras:username	nein	Name der Person, die für die Veröffentlichung des Datensatzes verantwortlich ist.
kontakt_email	maintainer_email	ja	Email-Adresse, an die man sich zu diesem Datensatz wenden kann, ggf. info@amt-xyz.de
kontakt_name	maintainer	nein	Name der Person, die für den Datensatz verantwortlich ist.

Spezifikation Input-API

- Location / URL: datenregister.berlin.de/metadatenatz
- Method : post
- HTTP-Header: Authorization, Wert: <API-Key>
- Request: {name: name_string,..., veroeffentlichungsdatum: veroeffentlichungsdatum_date, ..., "ressourcen: [{url: http://www..., format: format_string}, {...}],...} → Der detaillierte Aufbau der JSON-Struktur richtet sich nach den Bezeichnern in Tabelle 1.

Vorschlag für Berliner Datenportal	Pflicht	FOKUS Input-API	Typ
Name	x	titel	String
Eindeutiger Name		name	String (nur 'a-z0-9' und '-_')
Beschreibung		kurzbeschreibung	String
Kategorie	x	kategorie	Liste von Strings, s.u.
Schlagwörter		schlagwoerter	Liste von Strings
Veröffentlichende Stelle	x	veroeffentlichende_stelle	String
Veröffentlichungsdatum	x	veroeffentlichungsdatum	String: JJJJ-MM-TT
Änderungsdatum		aenderungsdatum	String: JJJJ-MM-TT
Ressourcen (in versch. Formaten, URL, Beschreibung inkl. Sprache und Link zur Dokumentation)	x	ressourcen	JSON-Dictionary
URL	x	url	String: URL
Format	x	format	String: Dateinamenerweiterung, MIME-Type, oder „api/{spec-type-or-mime-type-of-standard-response}“, Beispiele: „XLS“, „text/html“, „api/json“
Beschreibung	x	beschreibung	String
Sprache	x	sprache	String: zwei Buchstaben (ISO 639-1)
Link zur Dokumentation		link_dokumentation	String: URL
SHA1-Prüfsumme		sha1_hash	String
Kontaktinformation Name		kontakt_name	String
Kontaktinformation Email	x	kontakt_email	String: Email
Webadresse zu weiteren Informationen über den Datensatz		webadresse	String: URL
Geographische Abdeckung		geographische_abdeckung	String, s.u.
Geographische Granularität (Räumliche Auflösung innerhalb des Datensatzes)		geographische_granularitaet	String, s.u.
Zeitraum von		zeitraum_von	String: JJJJ-MM-TT
Zeitraum bis		zeitraum_bis	String: JJJJ-MM-TT
zeitliche Granularität		zeitliche_granularitaet	String: JJJJ-MM-TT
Lizenz	x	lizenz	String, s.u.

Veröffentlichende Person		veroeffentlichende_perso n	String
Sonstiges		sonstiges	String
Bezieht sich auf Datensatz		bezieht_sich_auf	String: URL
Tools/Apps		apps	String: URL

Tabelle 1 Metadatenfelder Input-API

Die möglichen Werte für das Feld **Kategorie**:

Name	Id
Wirtschaft	wirtschaft
Transport und Verkehr	transport
Umwelt und Klima	umwelt
Ver- und Entsorgung	verentsorgung
Geographie und Stadtplanung	geo
Gesundheit	gesundheit
Verbraucherschutz	verbraucher
Öffentliche Sicherheit	sicherheit
Wohnen und Immobilien	wohnen
Bildung	bildung
Öffentliche Verwaltung, Haushalt und Steuern	verwaltung
Gesetze und Justiz	justiz
Demographie	demographie
Arbeitsmarkt	arbeit
Wahlen	wahl
Sozialleistungen	sozial
Kunst und Kultur	kultur
Sport und Erholung	erholung
Tourismus	tourismus
Sonstiges	sonstiges

Tabelle 2 Kategorien

Die möglichen Werte für das Feld **Geographische Abdeckung** sind „Berlin“, die Bezirke und Orsteile:

Berlin	Falkenberg	Konradshöhe	Rosenthal
Charlottenburg-Wilmersdorf	Falkenhagener Feld	Köpenick	Rudow
Friedrichshain-Kreuzberg	Fennpfuhl	Kreuzberg	Rummelsburg
Lichtenberg	Französisch Buchholz	Lankwitz	Schmargendorf
Marzahn-Hellersdorf	Friedenau	Lichtenberg	Schmöckwitz

Mitte	Friedrichsfelde	Lichtenrade	Schöneberg
Neukölln	Friedrichshagen	Lichterfelde	Siemensstadt
Pankow	Friedrichshain	Lübars	Spandau
Reinickendorf	Frohnau	Mahlsdorf	Staaken
Spandau	Gatow	Malchow	Stadttrandsiedlung Malchow
Steglitz-Zehlendorf	Gesundbrunnen	Mariendorf	Steglitz
Tempelhof- Schöneberg	Gropiusstadt	Marienfelde	Tegel
Treptow-Köpenick	Grünau	Märkisches Viertel	Tempelhof
Adlershof	Grunewald	Marzahn	Tiergarten
Alt- Hohenschönhausen	Hakenfelde	Mitte	Waidmannslust
Alt-Treptow	Halensee	Moabit	Wannsee
Altglienicke	Hansaviertel	Müggelheim	Wartenberg
Baumschulenweg	Haselhorst	Neu- Hohenschönhausen	Wedding
Biesdorf	Heiligensee	Neukölln	Weißensee
Blankenburg	Heinersdorf	Niederschöneweide	Westend
Blankenfelde	Hellersdorf	Niederschönhausen	Wilhelmsruh
Bohnsdorf	Hermisdorf	Nikolassee	Wilhelmstadt
Britz	Hohenschönhausen	Oberschöneweide	Wilmersdorf
Buch	Johannisthal	Pankow	Wittenau
Buckow	Karlshorst	Plänterwald	Zehlendorf
Charlottenburg	Karow	Prenzlauer Berg	
Charlottenburg-Nord	Kaulsdorf	Rahnsdorf	
Dahlem	Kladow	Reinickendorf	

Tabelle 3 Geographische Abdeckung

Die möglichen Werte für das Feld **Geographische Granularität**:

Berlin
Bezirk
Ortsteil
Prognoseraum
Bezirksregion
Planungsraum
Block
Einschulbereich
Kontaktbereich
PLZ
Stimmbezirk
Quartiersmanagement
Wohnanlage
Wahlkreis
Hausnummer
GPS-Koordinaten

Tabelle 4 Geographische Granularität

Die möglichen Werte für das Feld **Zeitliche Granularität**:

Keine
5 Jahre
Jahr
Quartal
Monat
Woche
Tag
Stunde
Minute
Sekunde

Tabelle 5 Zeitliche Granularität

Die möglichen Werte für das Feld **Lizenz**:

license_id	License Title
apache	OSI Approved::Apache Software License
apache2.0	OSI Approved::Apache License, 2.0
bsd-license	OSI Approved::New and Simplified BSD licenses
ca-tosl1.1	OSI Approved::Computer Associates Trusted Open Source License 1.1

cc-by	OKD Compliant::Creative Commons Attribution
cc-by-sa	OKD Compliant::Creative Commons Share-Alike
cc-nc	Non-OKD Compliant::Creative Commons Non-Commercial (Any)
cc-zero	OKD Compliant::Creative Commons CCZero
eclipse-1.0	OSI Approved::Eclipse Public License
gfdl	OKD Compliant::GNU Free Documentation License
gpl-2.0	OSI Approved::GNU General Public License (GPL)
gpl-3.0	OSI Approved::GNU General Public License version 3.0 (GPLv3)
mit-license	OSI Approved::MIT license
mozilla	OSI Approved::Mozilla Public License 1.0 (MPL)
mozilla1.1	OSI Approved::Mozilla Public License 1.1 (MPL)
notspecified	Other::License Not Specified
odc-odbl	OKD Compliant::Open Data Commons Open Database License (ODbL)
odc-pddl	OKD Compliant::Open Data Commons Public Domain Dedication and Licence (PDDL)
W3C	OSI Approved::W3C License

Tabelle 6 Lizenzen

- Response: Standard HTTP status codes werden als Rückgabewerte verwendet.

Code	Name
200	OK
201	OK and new object created (referred to in the Location header)
301	Moved Permanently
400	Bad Request
403	Not Authorized
404	Not Found
409	Conflict (e.g. name already exists)
500	Service Error

Tabelle 5 Response

Spezifikation Output-API

Für die Kommunikation zwischen Open-Data-Register und Open-Data-Portal wird die API der CKAN-Registry auf der das Open-Data-Register basiert verwendet. Die CKAN-API Dokumentation ist zu finden unter: <http://docs.ckan.org/en/latest/api.html>. Insbesondere findet die Model-API Verwendung <http://docs.ckan.org/en/latest/api.html#model-api>. Für die Abbildung der Berliner Metadaten auf die internen Felder der CKAN-Registry (CKAN-Output) kann folgende Tabelle als Abbildungsreferenz herangezogen werden.

Metadaten Berliner Datenportal	Pflicht	CKAN Output
Name	x	title
Eindeutiger Bezeichner (nur 'a-z0-9' und '-_')	x	name
Beschreibung		notes
Kategorie	x	group
Schlagwörter		tags
Veröffentlichende Stelle	x	author
Veröffentlichungsdatum	x	extras:date_released
Änderungsdatum		extras:date_updated
Ressourcen (in versch. Formaten, URL, Beschreibung inkl. Sprache und Link zur Dokumentation)	x	resources
URL	x	url
Format	x	format
Beschreibung Sprache [Link zur Dokumentation]	x	description
SHA1-Prüfsumme		hash
Kontaktinformation Name		maintainer
Kontaktinformation Email	x	maintainer_email
Webadresse zu weiteren Informationen über den Datensatz		url
Geographische Abdeckung		extras:geographical_coverage
Geographische Granularität		extras:geographical_granularity
Zeitraum von	x	extras:temporal_coverage-from
Zeitraum bis	x	extras:temporal_coverage-to
zeitliche Granularität		extras:temporal_granularity
Lizenz	x	license_id
Veröffentlichende Person (nicht zum anzeigen)		extras:username
Sonstiges		extras:misc
Bezieht sich auf Datensatz		extras:relates_to
Tools/Apps		extras:apps

Tabelle 6 Abbildungsreferenz Berliner Metadaten auf CKAN

E Metadaten-Formular

Metadaten zum Datensatz

- * kennzeichnet ein Pflichtfeld.

Titel*

Kurzer bezeichnender Freitext (max. 256 Zeichen).

Soll eindeutig auf den Inhalt schließen lassen. Die Beschreibung erfolgt weiter unten.

Z.B. Wahl zum Berliner Abgeordnetenhaus 2006

Name *

Der Name wird fast immer korrekt automatisch ausgefüllt. Er wird später nicht geändert, auch wenn sich der Titel ändert und ist praktisch nur für Entwickler bedeutend.

Zwei oder mehr Zeichen, nur Kleinbuchstaben, 'a-z', '0-9', '-' und '_'. Soll trotzdem verständlich sein für Menschen. Eine spätere Änderung ist nicht vorgesehen.

Z.B. abgeordnetenhaus-berlin-wahl-2006

Veröffentlichende Stelle *

Möglichst genau die Verantwortliche Organisationseinheit

Z.B. Umweltamt Charlottenburg-Wilmersdorf

Kontakt-E-Mail *

Ansprechpartner zum Datensatz sind unter dieser Email zu erreichen.

Kontaktinformation

Name Persönlicher Ansprechpartner zum Datensatz (inhaltliche Verantwortung).

Veröffentlichende Person

Person, die lediglich für Veröffentlichung zuständig ist.

Webadresse

Die URL, unter der der Datensatz näher beschrieben wird (URL der Behörde).

z.B. <http://www.statistik-berlin-brandenburg.de/Statistiken/>

Kurzbeschreibung

Beschreibung des Datensatzes (1000 Zeichen).

Soll kurz gehalten sein, insbesondere sollte der erste Satz zusammenfassen, was genau enthalten ist.

Das sog. **Markdown Format** kann genutzt werden.

Lizenz*

Lizenz, unter der der Datensatz steht.

Derzeit sind zwei Möglichkeiten vorgesehen:

Dateien

URL der Datei*	Format*	Beschreibung*	Sprache*	Link zur API-Doku	
http://www.berlin.de/imperia/	xls	excel	de		
http://www.berlin.de/imperia/	xls	Industriebeschäftigte Teilberei	de		

Die Dateien, die die Daten enthalten, bzw. Adresse der API.

Es können mehrere Dateien angegeben werden, z.B. falls mehrere Formate vorliegen.

 Zeile hinzufügen

URL: WWW-Link, unter dem die Datei direkt heruntergeladen werden kann. Alternativ kann die URL auf einen API-Server zeigen, etwa einen SPARQL-Endpoint oder einen JSON-P-Dienst.

Format: Dateinamenerweiterung, z.B. xls, csv, sql, xml, kmz

Beschreibung: Weitere Angaben, die klären, welche Rolle die Datei innerhalb des Datensatzes spielt. Z.B. spezielles Format der Datei, Datei bezieht sich auf eine spezielle Region oder Zeitraum oder liefert zusätzliche Informationen wie Schlüssellisten oder eine Legende.

Sprache: Zwei-Buchstaben für die Sprache der Datei, z.B. de, en, tr (ISO 639-1).

API-Doku: Ggf. Verknüpfung zu Dokumentation von Format oder API.

Interne Notizen zur Bearbeitung der Metadaten dieses Datensatzes.

Autor: admin

Vorschau

Speichern

Wichtig: By submitting content, you agree to release your contributions under the open license specified on the [Lizenzseite](#). Please **refrain** from editing this page if you are **nicht** happy to do this.

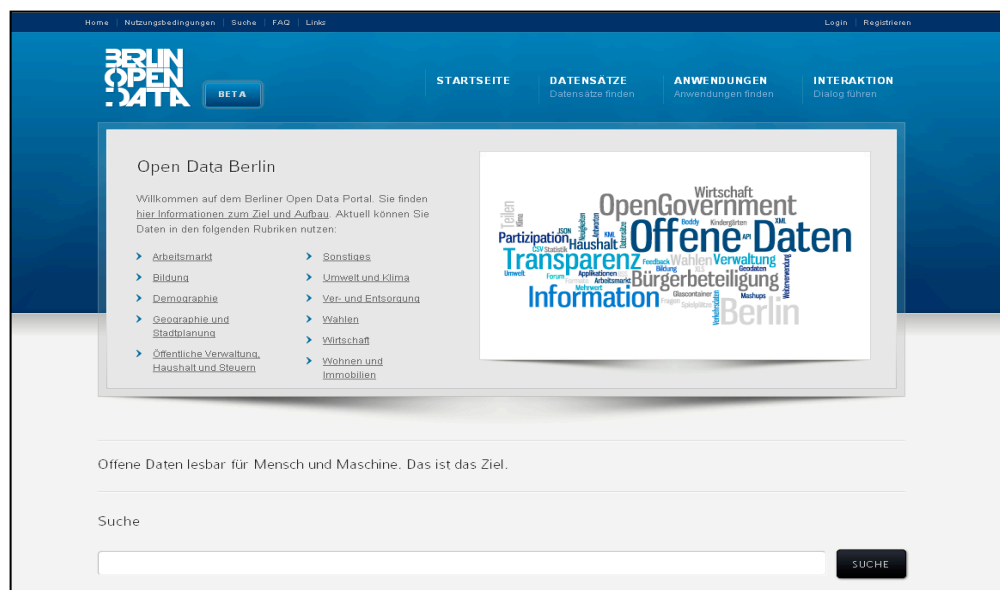
Datenregister Berlin
Powered by CKAN v1.4.2a

Ein Service von Fraunhofer FOKUS.

F Handreichung Datenregister

Leitfaden für das Einstellen und Pflegen von Datensätzen im Berliner Datenportal >> daten.berlin.de <<

Das Land Berlin hat Anfang 2011 durch einen Beschluss des Staatssekretärsausschusses für die Verwaltungsmodernisierung festgelegt, erste Schritte auf dem Weg der öffentlichen Verwaltung hin zur offenen Verwaltung zu gehen, um mehr Transparenz im Verwaltungs-, aber auch im politischen Handeln für Bürger, Institutionen und Medien zu schaffen. Zur Umsetzung dessen, soll der Öffentlichkeit nun ein besserer Zugang zu Datenbeständen der öffentlichen Hand bereitgestellt werden.



Bereits Mitte September 2011 begannen die Senatsverwaltung für Wirtschaft, Technologie und Frauen sowie für Inneres und Sport in Zusammenarbeit mit dem Fraunhofer Institut eine Erprobungsphase des Datenportals www.daten.berlin.de. **Um eine fortlaufende Erweiterung und Verbesserung des Portals bezüglich der Vielfalt zur Verfügung gestellter Daten zu erlangen, sind alle Kolleginnen und Kollegen eingeladen, ihre Daten auch digital über das neue Portal verfügbar zu machen.** Wie Sie dies bewerkstelligen können, soll ein Leitfaden im Folgenden kurz erläutern.

Geeignete Bereiche einstellbarer Daten:

Transport und Verkehr	Wohnen und Immobilien	Wirtschaft
Geographie und Stadtplanung	Demographie	Tourismus
Öffentliche Sicherheit	Sozialleistungen	Bildung
Umwelt- und Klima	Gesetze und Justiz	Wahlen
Verbraucherschutz	Kunst und Kultur	Arbeitsmarkt
Gesundheit, Sport	Öffentliche Verwaltung,	Ver- und
und Erholung	Haushalt und Steuern	Entsorgung

Vorgehensweise:

1. Die Ihnen vorliegenden Daten sollten ein maschinenlesbares Format aufweisen. Dies bedeutet, es muss sich um offene Formate wie Open Document, CSV-, KMZ-, XML- oder ähnliche Datei-Formate handeln, Word- oder PDF- Formate sind nicht möglich, Excel mit Einschränkung. Auch Grafikformate, wie JPEG, Gif oder Tiff, zählen nicht dazu.
2. Sie können mehrere Dateien zu einem Datensatz hinzufügen. Bitte legen Sie nach Möglichkeit jedoch nur eine Tabelle in jeder Datei ab, und verwenden Sie nicht mehrere Blätter innerhalb einer Excel-Tabelle.
3. Anschließend können Sie ihre Daten, weiterhin dezentral, auf der Internetseite Ihrer Behörde online verfügbar machen lassen. Ein konkretes Beispiel hierfür:

http://www.berlin.de/imperia/md/content/sen-wirtschaft/metadaten_itk/ikt_kreativwirtschaft_erwerbstaetige_2005_2008.csv

Ein hausinterner Internet- Redakteur wird Ihnen dabei möglicherweise zur Verfügung stehen.

4. Fordern Sie nun (einmalig) (z. Zt.) bei opendata@fokus.fraunhofer.de Ihre persönlichen Zugangsdaten an. Bitte geben Sie dabei Namen, Organisationseinheit und Telefonnummer an.

5. Nach Empfang Ihrer Zugangsdaten melden Sie sich mit Ihren Zugangsdaten auf <http://datenregister.berlin.de> an („Anmeldung“ oben rechts).
6. Wählen Sie „Metadaten hinzufügen“, es öffnet sich ein neues Formular. Füllen Sie dieses aus. Felder mit einem Sternchen müssen ausgefüllt werden.

7. Ausfüllen des Formulars:

- a. Was Sie im Titel eingeben, wird später auch im Datenportal so sichtbar werden. Das Feld „**Name**“ muss nicht ausgefüllt werden, da es automatisch den Titel übernimmt. Falls Sie Datensätze einstellen wollen, welche fortlaufend weiter geführt werden, ist es sinnvoll im Titel keine Zeiträume anzugeben, dies kann weiter unten in der Kurzbeschreibung erfolgen. So kann eine Aktualisierung Ihres Datensatzes auf der Homepage Ihrer Behörde vorgenommen werden, ohne dass im Titel des Datensatzes im Datenportal auch eine Aktualisierung vorgenommen werden muss.
- b. Unter „**Veröffentlichende Stelle**“ geben Sie Ihre Behörde an.
Bspw. Senatsverwaltung für Wirtschaft, Technologie und Frauen.
- c. Unter „**Kontakt-Email**“ wird die Email-Adresse des Bearbeiters eingetragen, der die Daten pflegt und für Rückfragen zur Verfügung steht.
- d. Unter „**Kontaktinformation Name**“ kann ein Ansprechpartner angegeben werden, der inhaltlich für den Datensatz verantwortlich ist.
- e. Unter „**Veröffentlichende Person**“ kann der Bearbeiter eingetragen werden, der die Daten erhoben hat, sie verwaltet, pflegt und für Rückfragen zur Verfügung steht.
- f. Unter „**Webadresse**“ ist die URL der Behörde einzutragen, unter welcher der Datensatz abgelegt wurde.
- g. Was unter „**Kurzbeschreibung**“ angegeben wird, wird später im Portal unter dem Titel ihres eingestellten Datensatzes sichtbar sein. Wie beispielsweise im Folgenden für den Datensatz „Bevölkerungszahlen Berlin“ :

Hier können zBsp. der betreffende Zeitraum, sowie Details zur Art und Weise der Datenerhebung und auch die Quelle der Daten angegeben werden.

h. Unter „Lizenz“ ist eine Angabe bezüglich der zu beachtenden Nutzungsbedingung Ihrer Daten zu treffen.

Informationen zur Lizenz:

Datensätze des Portals daten.berlin.de unterliegen, soweit nicht anders gekennzeichnet, der Lizenz „Creative Commons Namensnennung 3.0 Deutschland“. Creative Commons ist eine Non-Profit Organisation, die an nationale Rechtsprechung angepasste, standardisierte Lizenzverträge für die Veröffentlichung und Verbreitung digitaler Medieninhalte anbietet.

CC-Lizenzverträge bieten ein breites Spektrum an möglichen Kombinationen verschiedener Bedingungen an:

(siehe auch <http://de.creativecommons.org/was-ist-cc/>)

Namensnennung:

Es muss der Name des Autors/Rechteinhabers in der von ihm festgelegten Weise genannt werden.

Namensnennung - KeineBearbeitung:

Es muss der Name des Autors/Rechteinhabers in der von ihm festgelegten Weise genannt werden. Dieses Werk bzw. dieser Inhalt darf nicht bearbeitet, abgewandelt oder in anderer Weise verändert werden.

Namensnennung – NichtKommerziell:

Es muss der Name des Autors/Rechteinhabers in der von ihm festgelegten Weise genannt werden. Dieses Werk bzw. dieser Inhalt darf nicht für kommerzielle Zwecke verwendet werden.

Namensnennung – NichtKommerziell – KeineBearbeitung:

Es muss der Name des Autors/Rechteinhabers in der von ihm festgelegten Weise genannt werden. Dieses Werk bzw. dieser Inhalt darf nicht für kommerzielle Zwecke verwendet werden. Dieses Werk bzw. dieser Inhalt darf nicht bearbeitet, abgewandelt oder in anderer Weise verändert werden.

Namensnennung – NichtKommerziell – Weitergabe unter gleichen Bedingungen:

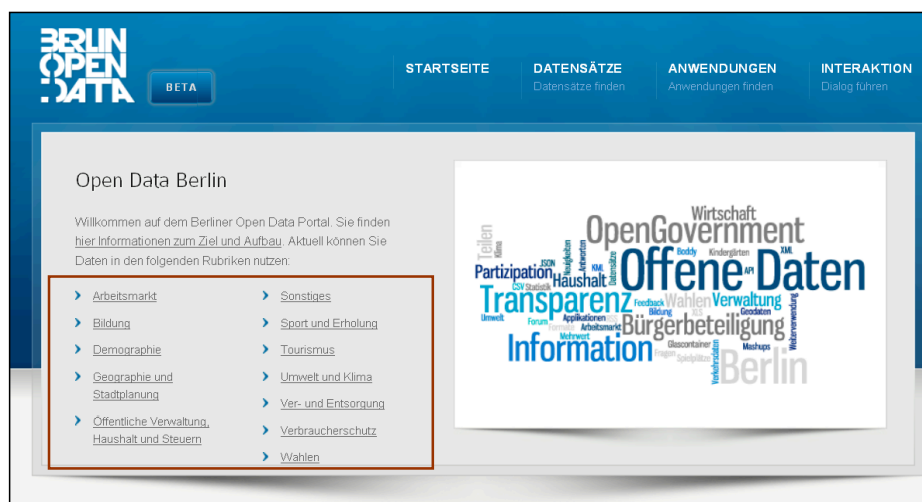
Es muss der Name des Autors/Rechteinhabers in der von ihm festgelegten Weise genannt werden. Dieses Werk bzw. dieser Inhalt darf nicht für kommerzielle Zwecke verwendet werden. Dieses Werk bzw. dieser Inhalt darf nicht bearbeitet, abgewandelt oder in anderer Weise verändert werden. Wird das lizenzierte Werk bzw. der lizenzierte Inhalt bearbeitet oder in anderer Weise erkennbar als Grundlage für eigenes Schaffen verwendet, dürfen die daraufhin neu entstandenen Werke bzw. Inhalte nur unter Verwendung von Lizenzbedingungen weitergeben, die mit denen dieses Lizenzvertrages identisch oder vergleichbar sind.

Namensnennung – Weitergabe unter gleichen Bedingungen:

Es muss der Name des Autors/Rechteinhabers in der von ihm festgelegten Weise genannt werden. Wird das lizenzierte Werk bzw. der lizenzierte Inhalt bearbeitet oder in anderer Weise erkennbar als Grundlage für eigenes Schaffen verwendet, dürfen die daraufhin neu entstandenen Werke bzw. Inhalte nur unter Verwendung von Lizenzbedingungen weitergeben, die mit denen dieses Lizenzvertrages identisch oder vergleichbar sind.

- i. Unter „**Veröffentlichungsdatum**“ geben Sie den Zeitpunkt der Veröffentlichung Ihrer Daten an. (zBsp.: das Datum an dem der Datensatz auf der Website ihrer Behörde verfügbar gemacht wurde)

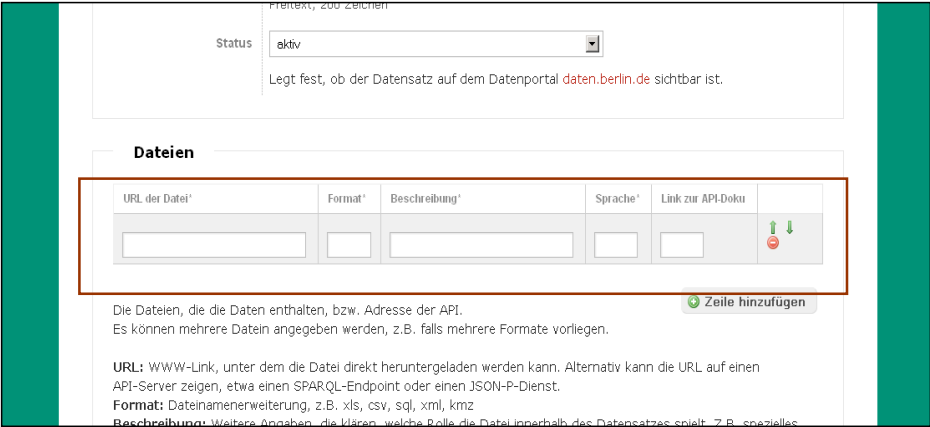
- j. Falls die Daten bereits früher veröffentlicht wurden, kann unter „Aktualisierungsdatum“ ein Datum angegeben werden, an dem die Daten aktualisiert oder erweitert wurden.
- k. Wählen Sie unter „Kategorie“ eine oder mehrere passende Bereiche für Ihren Datensatz aus! Später ist diese auch die Kategorie unter der der Datensatz im Portal zu finden ist. Siehe hier:



- l. Unter „Schlagwörter“ können die Begriffe ausgewählt werden, die in der Suchfunktion des Datenportals später bei Eingabe einer der Begriffe, Ihren Datensatz anzeigen lassen werden. Wählen Sie nur einige passende Begriffe aus.
Bsp.: Datensatz „Bevölkerungszahlen Berlin“ → Bevölkerung, Bevölkerungsstand, Bevölkerungszahlen.
- m. Welchen Zeitraum der Inhalt Ihrer Daten abdeckt, geben Sie unter „Zeitraum von bis“ an.
- n. Unter „Zeitliche Auflösung“ können Sie angeben, in welchen zeitlichen Intervallen die Daten angegeben werden.
- o. Unter „Geographische Abdeckung“ geben Sie an ob sich der Datensatz auf ganz Berlin bezieht oder bspw. einen bestimmten Bezirk.
- p. Die „Geographische Auflösung“ gibt an nach welchen örtlichen Kriterien die Daten erhoben wurden. Wurde zum Beispiel ein Datensatz pro Spalte in Ortsteile

eingeteilt, ist die geographische Auflösung „Ortsteil“. Werden in einem Datensatz Angaben über verschiedene Bezirke gemacht, ist die geographische Auflösung „Bezirk“. Weiterhin möglich sind unter Anderem auch Postleitzahl, Wahlkreis, Wohnanlage etc.

- q. Wenn Ihnen bekannt ist, dass bereits geeignete Anwendungen existieren, die auf den Datensatz aufbauen, geben Sie unter **„Anwendungen“** die entsprechende URL, unter der diese abgelegt ist, an. Möchten Sie Angaben machen, die nicht sichtbar sein sollen können Sie dies unter „Sonstiges“ tun.
- r. Unter **„Status“** aktivieren Sie den eingestellten Datensatz, somit wird er anschließend im Datenportal auch sichtbar gemacht. Wenn Sie den Datensatz allerdings noch nicht sichtbar machen wollen, weil Sie ihn bspw. noch ergänzen oder weiter bearbeiten möchten, wählen Sie **„deaktiviert“** und „Abbrechen“!
- s. Unter dem Abschnitt **„Dateien“** ist zunächst die URL der Datei einzugeben. Kopieren Sie hierfür den konkreten Link einer Ihrer zuvor eingestellten Dateien auf der Seite ihrer Behörde. Weiterhin ist das **„Format“** anzugeben wie bspw. xls; csv; odt; xml. Unter „Beschreibung“ kann (bspw. gut geeignet beim Einstellen mehrerer Dateien zusammen unter einem Titel) ein Untertitel der die Unterschiede der verschiedenen Dateien klarstellt angegeben werden.



Freitext, 200 Zeichen

Status:

Legt fest, ob der Datensatz auf dem Datenportal daten.berlin.de sichtbar ist.

Dateien

URL der Datei*	Format*	Beschreibung*	Sprache*	Link zur API-Doku

[Zeile hinzufügen](#)

Die Dateien, die die Daten enthalten, bzw. Adresse der API.
Es können mehrere Dateien angegeben werden, z.B. falls mehrere Formate vorliegen.

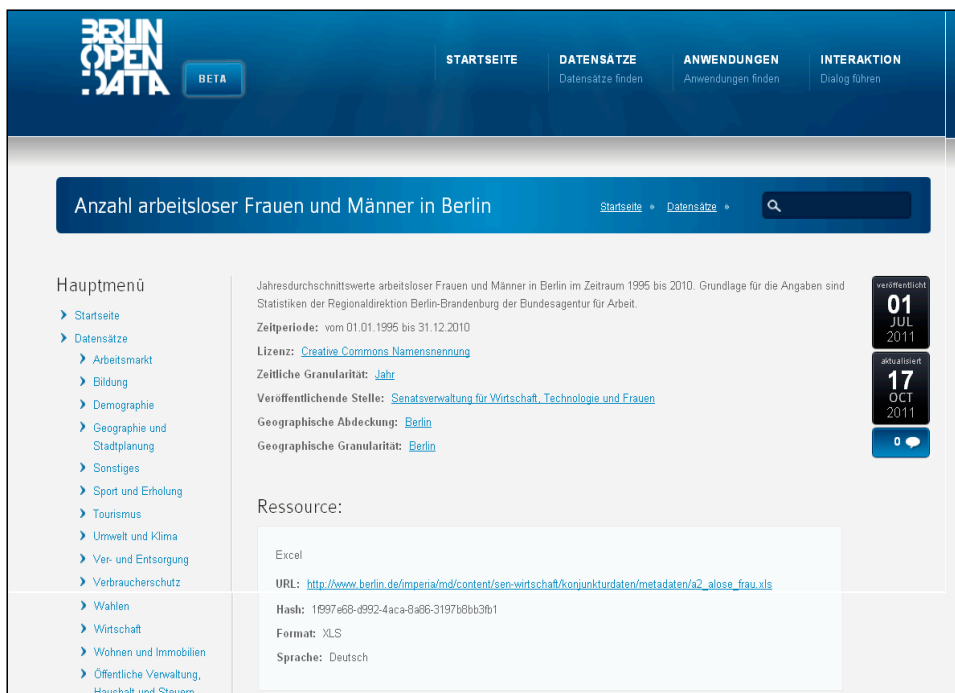
URL: WWW-Link, unter dem die Datei direkt heruntergeladen werden kann. Alternativ kann die URL auf einen API-Server zeigen, etwa einen SPARQL-Endpoint oder einen JSON-P-Dienst.

Format: Dateinamenerweiterung, z.B. xls, csv, sql, xml, kmz

Beschreibung: Weitere Angaben, die klären, welche Rolle die Datei innerhalb des Datensatzes spielt. Z.B. spezielles

8. Klicken Sie nun „Vorschau“ oder „Speichern“.

9. Ggf. können Sie später die Angaben korrigieren oder ergänzen, indem Sie per „Suche“ über den Titel oder Namen Ihren Datensatz anzeigen und auf „Bearbeiten“ klicken (zweiter Reiter oberhalb der Titelzeile neben Ansicht).
10. Ihr Datensatz sollte nun im Datenportal nach ca. einer Stunde sichtbar werden und beispielsweise so aussehen:



The screenshot shows the Berlin Open Data portal interface. At the top, there's a navigation bar with 'STARTSEITE', 'DATENSÄTZE', 'ANWENDUNGEN', and 'INTERAKTION'. Below this, the dataset title 'Anzahl arbeitsloser Frauen und Männer in Berlin' is displayed. A sidebar on the left lists various categories under 'Hauptmenü'. The main content area provides details about the dataset, including its time period (1995-2010), license (Creative Commons), and a download link for an Excel file. A 'Ressource' section also lists the URL, hash, format, and language.

Hauptmenü

- Startseite
- Datensätze
 - Arbeitsmarkt
 - Bildung
 - Demographie
 - Geographie und Stadtplanung
 - Sonstiges
 - Sport und Erholung
 - Tourismus
 - Umwelt und Klima
 - Ver- und Entsorgung
 - Verbraucherschutz
- Wahlen
- Wirtschaft
- Wohnen und Immobilien
- Öffentliche Verwaltung, Haushalt und Steuern

Anzahl arbeitsloser Frauen und Männer in Berlin

Jahresdurchschnittswerte arbeitsloser Frauen und Männer in Berlin im Zeitraum 1995 bis 2010. Grundlage für die Angaben sind Statistiken der Regionaldirektion Berlin-Brandenburg der Bundesagentur für Arbeit.

Zeitperiode: vom 01.01.1995 bis 31.12.2010

Lizenz: [Creative Commons Namensnennung](#)

Zeitliche Granularität: [Jahr](#)

Veröffentlichende Stelle: [Senatsverwaltung für Wirtschaft, Technologie und Frauen](#)

Geographische Abdeckung: [Berlin](#)

Geographische Granularität: [Berlin](#)

Ressource:

Excel

URL: http://www.berlin.de/imperia/md/content/sen-wirtschaft/sonjunkturdaten/metadaten/a2_alose_frau.xls

Hash: 1f97e68-d992-4aca-8a96-319716bb3fb1

Format: XLS

Sprache: Deutsch

veröffentlicht: 01 JUL 2011

aktualisiert: 17 OCT 2011

Damit haben Sie erfolgreich einen Datensatz im Register eingetragen.

Hinweise

Möchten Sie **einen Datensatz** einstellen, welcher **aus mehreren Dateien** unter einer Überschrift zusammengefasst werden soll, ist es sinnvoll das Formular nur einmalig auszufüllen und im Abschnitt „Dateien“ jeweils eine Zeile für eine weitere Datei hinzuzufügen.

Beispiel: Datensatz „20 grüne Hauptwege“

Status

aktiv

Legt fest, ob der Datensatz auf dem Datenportal daten.berlin.de sichtbar ist.

Dateien

URL der Datei ¹	Format ¹	Beschreibung ¹	Sprache ¹	Link zur API-Doku	
http://www.stadtentwicklung.berlin.de/geo/raumplanung/20_gruene_hauptwege/kmz/Spreeweg/Berliner_Urstromtal.kmz	kmz	Spreeweg/Berliner Urstromtal	de		 
http://www.stadtentwicklung.berlin.de/geo/raumplanung/20_gruene_hauptwege/kmz/Spandauer_Weg.kmz	kmz	Spandauer Weg	de		 
http://www.stadtentwicklung.berlin.de/geo/raumplanung/20_gruene_hauptwege/kmz/Hellgenseer_Weg.kmz	kmz	Hellgenseer Weg	de		 
http://www.stadtentwicklung.berlin.de/geo/raumplanung/20_gruene_hauptwege/kmz/Lubarser_Weg.kmz	kmz	Lubarser Weg	de		 
http://www.stadtentwicklung.berlin.de/geo/raumplanung/20_gruene_hauptwege/kmz/Nord-Sud-Weg.kmz	kmz	Nord-Süd-Weg	de		 
http://www.stadtentwicklung.berlin.de/geo/raumplanung/20_gruene_hauptwege/kmz/Lindenberger_Korridor.kmz	kmz	Lindenberger Korridor	de		 
http://www.stadtentwicklung.berlin.de/geo/raumplanung/20_gruene_hauptwege/kmz/Ahrensfelder_Weg.kmz	kmz	Ahrensfelder Weg	de		 
http://www.stadtentwicklung.berlin.de/geo/raumplanung/20_gruene_hauptwege/kmz/Honower_Weg.kmz	kmz	Hönower Weg	de		

SimpleSearch

Wenn Sie bereits SimpleSearch aus dem Baukasten von Imperia einsetzen, so finden Sie dort eine einfache Exportfunktion in das Datenregister. Hinweise hierzu unter:

http://support.berlin.de/wiki/index.php/SimpleSearch_OpenData

Im Auftrag des Landes Berlin untersuchte das Fraunhofer-Institut für Offene Kommunikationssysteme FOKUS im Rahmen des E-Government-Programms »ServiceStadt Berlin« die organisatorischen, rechtlichen und technischen Grundlagen für Open Data in der Hauptstadt. Hierbei wurde eine Konzeptstudie erarbeitet und in einem Online-Pilotsystem umgesetzt, dem ersten Open Data-Portal in Deutschland. Diese Studie dokumentiert neben den Rahmenbedingungen das Konzept zur Umsetzung, den Prototyp sowie erste Erfahrungen aus dem Pilotbetrieb.

Fraunhofer FOKUS, das Fraunhofer-Institut für Offene Kommunikationssysteme, entwickelt und erforscht am Standort Berlin Kommunikations- und Informationstechnologien. Im Zentrum der Forschungsaktivitäten steht die Entwicklung von IT-Infrastrukturen und Netzwerken zur domain- und organisationsübergreifenden Kommunikation.

Mit mehr als 20 Jahren Erfahrung stellt Fraunhofer FOKUS einen bedeutenden IT-Forschungspartner für die Telekommunikations-, Automobil- und Medienindustrie sowie für die öffentliche Verwaltung dar. Auf wissenschaftlicher Ebene gilt Fraunhofer FOKUS als renommierter Kooperationspartner in Forschungsprogrammen der Europäischen Kommission sowie nationaler Fördergeber.

ISBN 978-3-8396-0368-0

