



TECHNOLOGIE
STIFTUNG
BERLIN

Open Data in der Berliner Verwaltung

Status quo, Bedarfe und Perspektiven

Bericht

Impressum

© Technologiestiftung Berlin 2018

Grunewaldstraße 61-62 · 10825 Berlin · Telefon +49 30 209 69 990

info@technologiestiftung-berlin.de · technologiestiftung-berlin.de

Autor

Benjamin Seibel

Gestaltung

WEBERSUPIRAN.berlin



Textinhalte, Tabellen und nicht mit dem ©-Symbol gekennzeichnete Abbildungen des Werkes können genutzt und geteilt werden: Creative Commons - Lizenz Namensnennung - Weitergabe unter gleichen Bedingungen 3.0 Deutschland (www.creativecommons.org/licences/by-sa/3.0/de)

Als Namensnennung für Text und Tabellen ist anzugeben:

Benjamin Seibel: „Open Data in der Berliner Verwaltung“, Technologiestiftung Berlin, 2018.

Bei den Abbildungen Seite 6, 21, 24, 28, 33, 39 und 48 ist zusätzlich anzugeben:

Abbildung: WEBERSUPIRAN.berlin

Abbildung Seite 4 © tomeng/istock; Seite 9 © DasMaddin/istock; Seite 23 © kai zhang /istock;

Seite 38 © gopixa/istock

Der Autor weiß um die Bedeutung einer geschlechtergerechten Sprache und befürwortet grundsätzlich den Gebrauch von Parallelformulierungen. Von einer durchgehenden Benennung beider Geschlechter bzw. der konsequenten Verwendung geschlechterneutraler Bezeichnungen wurde im vorliegenden Text dennoch abgesehen, weil die Lesbarkeit deutlich erschwert würde.



Dieses Projekt der Technologiestiftung Berlin wird gefördert von der Senatsverwaltung für Wirtschaft, Energie und Betriebe und der Investitionsbank Berlin aus den Mitteln des Landes Berlin.

Inhalt

	Zusammenfassung/Summary	6
1	Einleitung: Wenn wir wüssten, was wir wissen...	8
2	Datenmanagement und Open Data in der Berliner Verwaltung	13
3	Bedarfe der Verwaltung bei der Bereitstellung von Daten	27
	3.1 Koordination: Wer ist eigentlich zuständig?	28
	Zuständigkeit innerhalb einzelner Verwaltungen	28
	Austausch zwischen Verwaltungen und mit der Open Data-Community	30
	3.2 Orientierung: Was soll überhaupt veröffentlicht werden?	32
	Angebots- oder Nachfrageorientierung?	32
	Systematische Inventur von Datenbeständen auf Verwaltungsseite	33
	Gezielte Ergänzung der Datenveröffentlichung aus Nachfragesicht	35
	3.3 Unterstützung: Wie kann die Veröffentlichung erleichtert werden?	37
	Unterstützung durch eine zentrale Servicestelle	37
	Unterstützung durch geeignete Software-Tools	38
	Auf dem Weg zu einem integrierten Datenmanagement	40
4	Vorgehensmodell	42
	4.1 Vorbereitende Schritte	43
	4.2 Inventur der Datenbestände	45
	4.3 Datenkatalog und Veröffentlichungsplan	47
	4.4 Aufbereitung und Veröffentlichung	49
	4.5 Ausbau und Verstetigung	50
5	Anhang	53
	5.1 Literatur	54
	5.2 Liste der Interviewpartner*innen	56
	5.3 Liste der Bereitsteller im Berliner Open Data-Portal	57
	5.4 Links	58

Zusammenfassung

Das Berliner Open Data-Angebot ist eine Erfolgsgeschichte: Nach rund fünf Jahren Regelbetrieb finden sich im Datenportal des Landes bereits mehr als 1700 Datensätze zur freien Weiternutzung. Bedingt durch das Wachstum des Portals und durch die Anforderungen aus dem neuen Berliner E-Government-Gesetz werden nun weitere Schritte erforderlich: Das Datenangebot der Verwaltung soll stärker systematisiert und qualitativ verbessert werden.

In Interviews mit Beschäftigten der Bezirks- und Hauptverwaltungen hat die Technologiestiftung untersucht, wie eine bedarfsgerechte Weiterentwicklung des Open Data-Angebotes im Land Berlin erfolgen kann. Dabei wurden drei zentrale Ansatzpunkte identifiziert:

- Bedarf nach klaren Zuständigkeiten innerhalb einzelner Verwaltungen und Behörden
- Bedarf nach stärkerer Koordination und strategischer Steuerung der Veröffentlichung
- Bedarf nach praktischer und technischer Unterstützung im Veröffentlichungsprozess

Ausgehend von diesen Bedarfen wird ein Vorgehensmodell skizziert, mit dem die Open Data-Bereitstellung in Berlin weiter verbessert werden kann. Die darin aufgeführten Maßnahmen umfassen:

- Inventur der in der Verwaltung verfügbaren Daten
- Priorisierung und Erarbeitung eines Veröffentlichungsplans
- Redaktionelle Prüfung und kontrollierte Veröffentlichung
- Ausbau und Verstetigung der Open Data-Prozesse

Zur erfolgreichen Umsetzung des Vorgehensmodells wird zudem empfohlen, klare Verantwortlichkeiten für Open Data auf Bezirks- und Landesebene zu benennen. Für die praktische Unterstützung wird die Einrichtung einer übergreifenden „Servicestelle Open Data“ angeregt.

Summary

The story of Open Data and the city of Berlin can largely be seen as a successful one: roughly five years since the city first released its Open Data portal, more than 1,700 datasets are currently available for download and re-use. But given the continued growth of the portal as well as the requirements introduced by the new Berlin E-Government Law, it's clear that further steps are needed to improve the quality of data offerings as well as the systematization of data publishing.


Using interviews conducted with employees of both the central city administration and individual borough administrations, the Technologiestiftung has evaluated how Berlin's Open Data offerings can be further developed based on the specific needs of data users and publishers. Through these interviews, three key needs emerged:

- A need for clearly delineated responsibilities with respect to data management and data publishing within individual administrations and offices
- A need for a data-publishing protocol that is both better coordinated and more strategically oriented
- A need for more practical and technical assistance with regards to the data publishing process

Based on these identified needs, the Technologiestiftung has developed a procedural model for how the provision of Open Data in Berlin can be improved. The proposed measures include:

- Conducting an inventory of all data held by the city administration
- Prioritization and creation of a data publishing plan
- Increased supervision of the quality and content of published datasets
- Upgrading of Open Data processes, including through increased automation

In order for this model to be successfully implemented, this report recommends that both borough-level and state-level administrations establish clear responsibilities with respect to Open Data. Further, in order to provide government bodies with practical support in the generating and publishing of Open Data, a "Service Point Open Data" should be established.



1
Einleitung:
Wenn wir wüssten,
was wir wissen...

„The cities, countries and even continents that build the best and most open data infrastructure will have an enormous advantage in the 21st century economy.“¹

Die Berliner Verwaltung verfügt über enorme Datenbestände. Viele dieser Daten sind von öffentlichem Interesse. Sie können dabei helfen, die Stadt besser zu verstehen, Politik transparenter zu machen und neue Dienste für Bürger*innen zu entwickeln. Hinter „Open Data“ steht die Idee, dass Verwaltungen weite Teile ihrer Datenbestände für eine freie Weiterverarbeitung durch Dritte zugänglich machen. Dass dieses Vorgehen nicht nur unter Transparenzgesichtspunkten wünschenswert ist, sondern darüber hinaus auch wirtschaftliche Innovation fördert, gilt inzwischen als unstrittig. Zahlreiche Studien haben in den letzten Jahren die großen Potenziale offener Verwaltungsdaten für Gesellschaft und Wirtschaft beleuchtet.²

Bei der Transformation von Metropolregionen in sogenannte „Smart Cities“³ – also in Städte, deren Infrastrukturen und Prozesse digital und damit datenbasiert geregelt werden – kann und sollte Open Data eine zentrale Rolle spielen. Die Verfügbarkeit qualitativ hochwertiger Geo-, Umwelt- oder Verkehrsdaten wird mit der wachsenden volkswirtschaftlichen Bedeutung datengetriebener Innovationen zu einem wichtigen Standortfaktor. Ein möglichst freier Zugang zu diesen Datenbeständen ist bereits heute Grundlage für neue Formen digitaler Teilhabe und Wertschöpfung. Entsprechend forderte jüngst eine OECD-Studie, neben Glasfaser- und Funknetzen auch Daten selbst als essentiellen Bestandteil digitaler Infrastrukturen zu betrachten.⁴

Um sich für die Zukunft zu wappnen, müssen Städte in den Auf- und Ausbau offener Dateninfrastrukturen investieren. Im Land Berlin wurde diese Notwendigkeit früh erkannt und bereits 2012 eine Open Data-Strategie veröffentlicht.⁵ Im gleichen Jahr wurde das von der damaligen Berliner Senatsverwaltung für Wirtschaft, Technologie und Forschung und dem Fraunhofer Institut für offene Kommunikationssysteme (FOKUS) initiierte Berliner Datenportal in den Regelbetrieb überführt. Nach fünf weiteren Jahren ist dieses Portal eine Erfolgsgeschichte: Rund 1.700 Datensätze der Berliner Verwaltung stehen heute unter einer offenen Lizenz zum Download bereit. Innovative Unternehmen, Forschungseinrichtungen und zivilgesellschaftliche Initiativen nutzen diese Daten für die Erstellung von Analysen, Informationsangeboten und weiterführenden Services.⁶

1 <http://theodi.org/what-is-data-infrastructure>

2 Vgl. als Auswahl Technologiestiftung 2014, Seibel 2016, Kuzev 2016, Bitkom 2017

3 Vgl. als Überblick Erbstößer 2014.

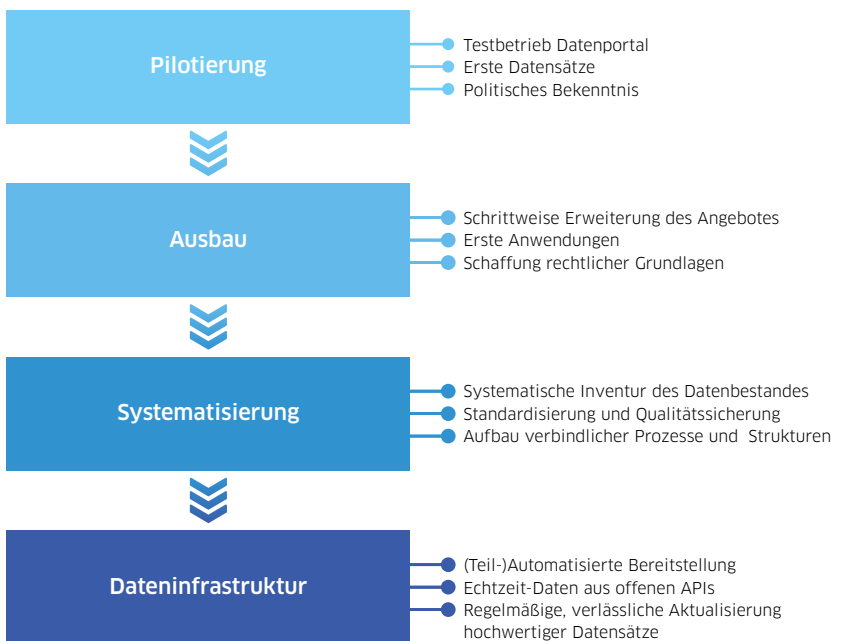
4 www.oecd.org/sti/ieconomy/PolicyNote-DDI.pdf

5 Both/Schieferdecker 2012.

6 <http://daten.berlin.de/anwendungen>

Gleichwohl sind die Potenziale offener Daten in Berlin noch längst nicht ausgeschöpft. Auch wenn die Hauptstadt im bundesweiten Vergleich gut da steht: Ein Blick auf die internationale Open Data-Landschaft verdeutlicht, welche Chancen in einer umfassenden Bereitstellung qualitativ hochwertiger Verwaltungsdaten liegen können.⁷ Um dieses Ziel zu erreichen, ist einerseits eine stärkere strategische Steuerung und Systematisierung der Datenbereitstellung nötig. Andererseits benötigen die Beschäftigten der öffentlichen Verwaltungen eine effektive Unterstützung bei der Integration einer regelmäßigen Datenveröffentlichung in ihre gewohnten Arbeitsabläufe.

Angelehnt an bestehende Open Data-Reifegradmodelle („Maturity Models“) lassen sich beim Aufbau einer Berliner Dateninfrastruktur vier Phasen unterscheiden: Eine Pilotphase, begann 2011/2012 mit der Formulierung einer Open Data-Strategie und dem Testbetrieb des Berliner Datenportals. Seitdem befindet sich das Portal im Regelbetrieb, wobei die Anzahl an verfügbaren Datensätzen kontinuierlich gewachsen ist. Zudem wurde mit der Verabschiedung des Berliner E-Government-Gesetzes und insbesondere mit dem §13 zur Bereitstellung allgemein zugänglicher Datenbestände eine gesetzliche Grundlage für den Aufbau einer offenen behördlichen Dateninfrastruktur geschaffen.⁸



Weiterentwicklung des Berliner Open Data-Angebots (eigene Darstellung)

⁷ Vgl. Wangermann (Hrsg.) 2016.

⁸ Vgl. „Gesetz zur Förderung des E-Government (E-Government-Gesetz Berlin - EGovG Bln) vom 30. Mai 2016. Auch bundespolitisch ist das Thema Open Data durch den Beitritt Deutschlands zur Open Government Partnership und der jüngsten Verabschiedung eines Open Data-Gesetzes des Bundes höher auf die Agenda gerückt.

Der nun anstehende Schritt einer Systematisierung der Open Data-Bereitstellung birgt eigene Herausforderungen, die bislang nur wenig untersucht sind: Wie verhält sich die Anforderung einer regelmäßigen Datenveröffentlichung zur Arbeitsrealität der Berliner Verwaltungen? Wo bestehen gegenwärtig Hindernisse, wo liegen Bedarfe für Unterstützung? Welche Maßnahmen sind kurzfristig nötig, um die Qualität des Datenangebots zu verbessern? Und wie lässt sich langfristig ein Modell von „Data Governance“⁹ etablieren, das den Ansprüchen an eine fortschrittliche Dateninfrastruktur gerecht wird?

Um uns diesen Fragen anzunähern, haben wir über die letzten Monate zahlreiche Interviews mit Beschäftigten der Berliner Bezirks- und Hauptverwaltungen geführt. Wir wollten besser verstehen, wie die tägliche Arbeit mit Daten in öffentlichen Einrichtungen funktioniert und wo die Bedarfe jener Mitarbeiter*innen liegen, die heute schon mit der Erhebung, Aufbereitung und Veröffentlichung von Verwaltungsdaten betraut sind. In den Debatten über die politischen und rechtlichen Rahmenbedingungen offener Verwaltungsdaten sind diese Stimmen aus der Praxis oft zu wenig präsent. Um politische Strategien erfolgreich umzusetzen ist es aber unerlässlich, dieses Expertenwissen stärker einzubeziehen.

Im Rahmen unserer Gespräche wurde schnell deutlich, dass eine umfassende Datenbereitstellung die Beschäftigten vor große operative Herausforderungen stellt. Zwar ist eine Rechtsgrundlage für Open Data geschaffen, die Frage nach der konkreten Umsetzung bleibt aber offen. Der veränderte Umgang mit Daten betrifft die Organisationskultur in Gänze. In der verzweigten Berliner Verwaltungslandschaft ist schon die Identifikation und Koordination der für das Thema relevanten Akteure schwierig. So sind es nicht in erster Linie technische oder rechtliche, sondern vor allem organisatorische und praktische Hürden, die beim Aufbau einer offenen Dateninfrastruktur überwunden werden müssen. Dies bedeutet umgekehrt, dass gerade im Aufbau geeigneter Strukturen und Abläufe die Chance liegt, die Potenziale offener Verwaltungsdaten zu entfalten.

„Wenn wir wüssten, was wir wissen, wären wir einen großen Schritt weiter“ – Dieses Resümee eines unserer Gesprächspartner ist uns besonders im Gedächtnis geblieben.¹⁰ Der Satz verdeutlicht nicht nur eine zentrale Schwierigkeit bei der Datenbereitstellung, sondern zugleich auch, dass die Verwaltungen selbst zu den größten Profiteuren einer offenen Datenkultur zählen können. Ein vereinfachter Zugriff auf offene und aktuelle Verwaltungsdaten wäre nicht zuletzt für die eigenen Beschäftigten mit großen Arbeitserleichterungen verbunden. In der Transformation der Verwaltung hin zu einem zeitgemäßen E-Government kann Open Data eine Schlüsselrolle zukommen.

⁹ Vgl. Lebhart 2016.

¹⁰ Sämtliche Interviews wurden aufgezeichnet und transkribiert, eine Liste der Gesprächspartner*innen findet sich im Anhang. Auf eine namentliche Kennzeichnung einzelner Zitate verzichten wir in Absprache mit den Befragten.

Entscheidend dafür wird sein, die Veränderungen im Umgang mit Daten bedarfsgerecht zu gestalten. Angesichts knapper Ressourcen müssen Wege gefunden werden, die Datenveröffentlichung möglichst effektiv und unkompliziert in bestehende Arbeitsabläufe einzubinden. Wie das gelingen kann, können die Beschäftigten selbst häufig am besten einschätzen. Ausgehend von einem Überblick über den gegenwärtigen Stand der Datenveröffentlichung werden im Folgenden drei Handlungsfelder beschrieben, bei denen im Rahmen unserer Gespräche besonderer Unterstützungsbedarf ersichtlich wurde. Abschließend wird ein Vorgehensmodell vorgeschlagen, mit dem sich die Open Data-Bereitstellung in Berlin systematisieren und ausbauen lässt.

An aerial photograph of Berlin, Germany, showing a mix of modern and historic architecture. The city is viewed from a high angle, with a prominent red brick building featuring a clock tower in the foreground. The sky is blue with scattered white clouds. Overlaid on the image are two large, semi-transparent geometric shapes: a blue triangle pointing downwards and a red triangle pointing upwards, which intersect to form a dark blue diamond shape. The text is centered within this dark blue area.

2 Datenmanagement und Open Data in der Berliner Verwaltung

„Open Data“ mag für öffentliche Verwaltung ein relativ junges Tätigkeitsfeld sein, die elektronische Datenverarbeitung ist es nicht: Schon am 10. Juli 1969 eröffnete in Westberlin das „Landesamt für elektronische Datenverarbeitung“ (LED), um andere Behörden mit Rechenleistung zu unterstützen. An Stelle großer Mainframe-Computer traten im Laufe der 1980er Jahre die Desktop-PCs, von denen heute mehr als 80.000 in der Berliner Verwaltung im Einsatz sind. Mit diesem Paradigmenwechsel einher ging eine Dezentralisierung der Verfahren zur Datenverarbeitung und -speicherung, die in wesentlichen Teilen bis heute fortbesteht.

Die Problematik dieser Entwicklung wird etwa dort ersichtlich, wo die Berliner Verwaltung mehr als 3.000 physikalische Server betreibt, die aufgrund der unterschiedlich gewachsenen Systemarchitekturen größtenteils dezentral und eigenständig durch die jeweiligen Behörden und Ämter verwaltet werden.¹¹ Zwar wurden inzwischen wichtige Maßnahmen zur Konsolidierung in die Wege geleitet¹², zum jetzigen Zeitpunkt bleibt die Berliner IT-Landschaft jedoch ungewöhnlich heterogen. Für dieselben Aufgaben kommt in verschiedenen Ämtern mitunter unterschiedliche Software zum Einsatz. Ein übergreifendes, koordiniertes Datenmanagement ist unter diesen Umständen eine besondere Herausforderung.

Die Digitalisierung der öffentlichen Verwaltung ist ein hochkomplexes Unterfangen, das schrittweise erfolgt, während gleichzeitig ein laufender Betrieb gewährleistet bleiben muss. Mit der Zeit sind so zahlreiche hybride Formen der Informationsverarbeitung entstanden: Je nach Vorgang spielt die papierbasierte Dokumentation nach wie vor eine große Rolle, vieles wird auch mit handelsüblicher Office-Software erledigt. Mehr als 300 IT-Fachverfahren sind zudem für spezielle Anwendungen und Dienstleistungen der Verwaltung im Einsatz. Teils finden durchaus komplexe Datenbanken und Plattformen Verwendung, die jedoch auf klar spezifizierte Einsatzfelder begrenzt sind.

Aus der Sonderstellung Berlins als Stadtstaat und der damit einhergehenden zweistufigen Verwaltungsstruktur ergeben sich zusätzliche Herausforderungen. Neben den Senatsverwaltungen, die als Hauptverwaltungen genuin staatliche Aufgaben übernehmen, bearbeiten die Bezirke mit einer gewissen operativen Autonomie jene Felder, die im politischen System der Bundesrepublik üblicherweise den Kommunen zugeordnet sind. Diese seit jeher umstrittene Dezentralisierung ist Stärke und Schwäche zugleich: Stärke, weil sie in einer polyzentrischen Metropole wie Berlin ein gewisses Maß an Flexibilität und kommunaler Selbstverwaltung ermöglicht. Schwäche, weil sie die Koordination erschwert und ein landesweit abgestimmtes Vorgehen in vielen Fällen bremst.

¹¹ Vgl. IT-Bestands- und Planungsübersicht – Bericht 2016

¹² Ein Überblick über geplante Veränderungen findet sich im Dokument „IKT-Architektur für das Land Berlin“ auf www.berlin.de/sen/inneres/moderne-verwaltung/informationstechnik/it-infrastruktur

Bezirke

Bei Bürgerangelegenheiten und sozialraumbezogenen Vorgängen sind oft die Bezirksverwaltungen und die ihnen zugehörigen Fachämter der Ausgangspunkt im behördlichen Datenfluss. Zu den Hauptgeschäftsfeldern der Bezirke gehören das Meldewesen, die Kinder- und Jugendarbeit, das Umwelt- und Grünflächenmanagement, Bildung, Kultur und Soziales, sowie das Planungs- und Vermessungswesen. Auch die Liegenschaftsverwaltung und die Bauaufsicht über Straßen und Infrastruktur liegt in der Verantwortung der Bezirksämter. Über den konkreten Zuschnitt und die Verteilung der Ressorts entscheiden die einzelnen Ämter innerhalb gewisser Grenzen selbst.

The screenshot shows a web application for applying for a German ID card (Personalausweis). The interface is in German and includes the following elements:

- Header:** "Antrag Personalausweis | Großstadt | Hauptstandort".
- Applicant Information:** "Burk, Alexander (16)", born 01.01.2000 in Lindeholm.
- Application Details:** "T246V37C84 (Antrag begonnen) 11.03.2016 - 10.03.2022", issued by "STADT GROßSTADT PERSONAUSWEISBEHÖRDE", address: "AW in Baroper Str. 251, 99995 Großstadt OT Buckow-Nord seit 01.01.2016".
- Preview of ID Card:** Shows the front and back of the ID card. The front card displays: "BUNDESREPUBLIK DEUTSCHLAND", "PERSONAUSWEIS", "BEZIRK", "ALEXANDER", "01.01.2000", "BERLIN", "LINDEHOLM", "10.03.2022". The back card displays: "99995 GROßSTADT BUCKOW-NORD BAROPER STR. 251", "11.03.2016", "STADT GROßSTADT PERSONAUSWEISBEHÖRDE".
- Form Fields:** "es wurde eine ausländische Staatsangehörigkeit", "Namensnachweis", "Geburtsnachweis", "Geschlecht" (männlich).
- Buttons:** "Kontrollblatt drucken", "aus Dateisystem", "aus Dateisystem", "Speed Capture".
- Footer:** "Fingerabdrücke aufnehmen", "Kein", "Der PIN-Brief wird an die Behörde versandt", "Das Informationsmaterial wurde an den Antragsteller übergeben", "Die Erklärungen zum Personalausweis (Formular) mündlich abgegeben".

Beispiel: VOIS-Fachverfahren der Bürgerämter

Bezirksverwaltungen sind in erster Linie Dienstleister für die Berliner Bürger*innen. Sie bearbeiten Anträge, betreiben Volkshochschulen, zahlen Unterstützungen aus oder planen Bebauungen. Entsprechend vielfältig sind die bei diesen Tätigkeiten anfallenden Daten, wobei die Verwaltung selbiger zunächst den jeweils zuständigen Fachabteilungen obliegt. IT-seitig kommen bei vielen Vorgängen Fachverfahren zum Einsatz, die der Bearbeitung eines klar spezifizierten Vorgangs, etwa der Anmeldung eines Wohnsitzes, dienen. Mitunter sind zuständige Hauptämter wie das Landesamt für Bürger- und Ordnungsangelegenheiten (LABO) direkt an diese Verfahren angeschlossen. In einigen Fällen erlauben Fachverfahren auch einen strukturierten Datenexport in maschinenlesbare Formate, so dass eine externe Weiterverarbeitung möglich ist.

Beschaffung und Weitergabe von Daten

Häufig bleibt die Datenerhebung in den Bezirken aber auch „Handarbeit.“ Informationen, die in erster Linie der bezirklichen Selbstverwaltung und Planung dienen, werden von den in den jeweiligen Ämtern zuständigen Personen eigenständig erfasst und verwaltet. Die hier zum Einsatz kommenden Verfahren der Speicherung und Verarbeitung sind bislang wenig standardisiert. Allerdings werden Daten von gesamtstädtischer Relevanz regelmäßig an die zuständigen Senatsverwaltungen gemeldet, wo dann auch eine Vereinheitlichung erfolgen kann. So existieren auf dem Online-Portal Berlin.de stadtweite Informations- und Suchangebote, die aus bezirklichen Daten gespeist, aber von Senatsverwaltungen betreut werden.

Auch wenn also durchaus punktuell Daten zwischen Bezirks- und Hauptverwaltungen ausgetauscht werden, verbleibt ein Großteil der bezirklichen Datenbestände bei den jeweils zuständigen Fachämtern. Selbst innerhalb einzelner Verwaltungen fehlt es dabei häufig an einem Überblick, welche Daten in den einzelnen Bereichen existieren. Dies führt mitunter dazu, dass Daten mehrfach und mit unterschiedlichem Aktualisierungsstand vorliegen. Um hier Abhilfe zu schaffen, haben erste Bezirke mit der Konzeption und dem Aufbau eigener (Geo-)Dateninfrastrukturen begonnen. So sollen intern Redundanzen abgebaut und mehr Transparenz über verfügbare Datensätze geschaffen werden.

Bezirkliche Datenkoordination

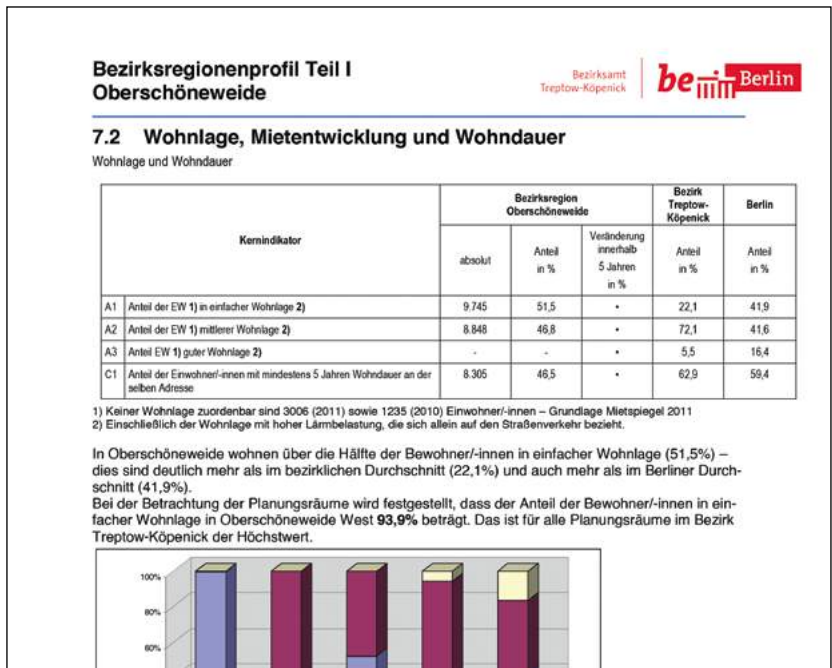
Eine wichtige Rolle in diesem Zusammenhang spielt die Position der „bezirklichen Datenkoordination“, die in mittlerweile sechs Bezirken¹³ als Teil der „Organisationseinheit Soziale Planungscoordination“ (OE SPK)¹⁴ eingerichtet wurde. Die Datenkoordination übernimmt nach Bedarf die Beschaffung, Zusammenführung, Aufbereitung und Bereitstellung von Daten aus den einzelnen Fachämtern und fungiert so als übergreifendes Bindeglied. Sie unterstützt zudem bei der Erstellung datenbasierter Bezirksregionenprofile, sowie beim Aufbau eines gemeinsamen Datenpools, in dem Fachdaten zusammengeführt und verwaltungsübergreifend abgestimmt werden.

Die Einrichtung einer Datenkoordination ist perspektivisch für alle Bezirke vorgesehen, wurde aber vornehmlich aus Ressourcengründen noch nicht überall umgesetzt. Ein Fachartikel fasst die Aufgaben dieser noch recht jungen Position wie folgt zusammen: „Ziel der Datenkoordination ist die Bearbeitung verwaltungsinterner Fragestellungen sowie Anfragen aus dem Bezirksamt. Dazu zählt die Aufbereitung aktueller Einwohnerdaten nach unterschiedlichen Merkmalen auf Ebene der LOR und Statistischen Blöcke. Weitere Aufgaben

13 Treptow-Köpenick, Tempelhof-Schöneberg, Pankow, Mitte, Lichtenberg, Marzahn-Hellersdorf.

14 Der Aufbau einer „Organisationseinheit Sozialraumorientierte Planungscoordination“ (OE SPK), der seit 2011 im Bezirksverwaltungsgesetz vorgeschrieben ist, zielt auf ein fachübergreifendes und kleinräumiges, an den Bedürfnissen und der Teilhabe der Bürger*innen orientiertes Planen und Handeln. Zum Konzept der Sozialraumorientierung in der Berliner Stadtplanung vgl. Senatsverwaltung für Stadtentwicklung (2009).

sind die Leistung von Hilfestellungen bei der Bearbeitung statistischer Angaben in bezirklichen Förderanträgen sowie die Bereitstellung umfangreichen Datenmaterials zu Themen wie beispielsweise Ausländerinnen und Ausländer oder Migration und Flüchtlinge. Die Datenkoordination fungiert dabei als eine Art Statistik-Stelle und ist für die technische Bearbeitung und Aufarbeitung des Datenmaterials zuständig.“¹⁵



Bezirksregionenprofil Oberschöneeweide (Auszug)

Die bislang gemachten Erfahrungen mit der Datenkoordination sind in den Bezirken durchweg positiv. Aus der amtsübergreifenden Perspektive können Mehrwerte sichtbar gemacht werden, die aus der Zusammenführung und Analyse von Daten entstehen. Die bereits aktiven Datenkoordinator*innen tauschen sich regelmäßig zu ihrem Vorgehen aus. Neben der Harmonisierung und Qualitätsprüfung von Datenbeständen befasst sich die Gruppe derzeit auch mit der Erstellung eines Leitfadens, der die Aufgaben der Datenkoordination weiter spezifiziert. Für den Aufbau einer berlinweiten Dateninfrastruktur leisten die bezirklichen Datenkoordinator*innen derzeit wichtige Grundlagenarbeit.

Senatsverwaltungen

Als Leitungsorgane der Berliner Landesregierung fungieren die Berliner Senatskanzlei sowie die gegenwärtig zehn Senatsverwaltungen. Teil der Hauptverwaltung sind zudem die der Senatsverwaltung nachgeordneten Landes-

oberbehörden (u.a. Landesamt für Bürger- und Ordnungsangelegenheiten, Landesamt für Gesundheit und Soziales, Landesdenkmalamt, Landesverwaltungsamt, Polizei und Feuerwehr) und die unter Aufsicht der Senatsverwaltung stehenden Eigenbetriebe. Aufgaben, denen eine gesamtstädtische Bedeutung zukommt, sind der Hauptverwaltung vorbehalten. Dazu gehören ministerielle Leitungsaufgaben wie Planung, Steuerung und Aufsicht, sowie zahlreiche weitere Aufgaben von bezirksübergreifender Relevanz.

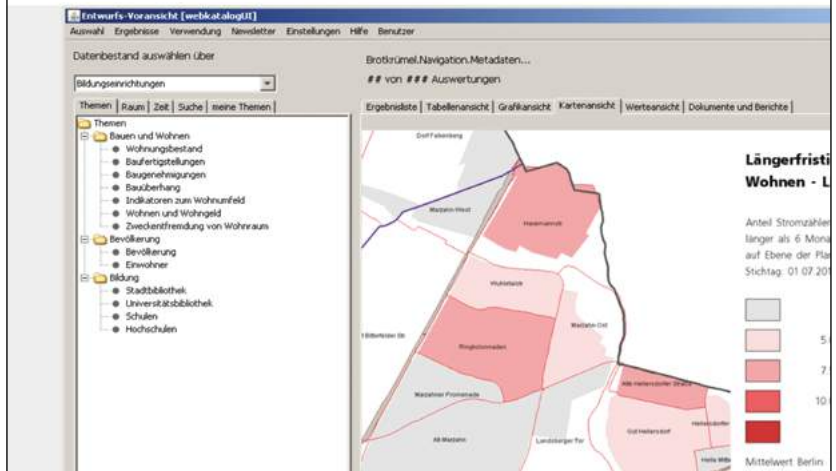
Im Rahmen dieser Tätigkeiten fragen die Senatsverwaltungen regelmäßig Datenbestände von Bezirken, Landesämtern und Betrieben an und werten diese aus. Verglichen mit den Bezirken werden also weniger Primärdaten erhoben. Gemäß der ministeriellen Aufgaben werden vielmehr bestehende Daten aus verschiedenen Quellen zusammengeführt und zu Zwecken der Berichterstattung, Planung und Steuerung aufbereitet. Die Beschaffung der Daten geschieht teils bereits automatisiert über die direkte Anbindung an Fachverfahren (etwa im Bereich des Meldewesens). In vielen Fällen werden Daten aber auch manuell bei den Beschäftigten der bezirklichen Fachämter angefragt oder von externen Angeboten bezogen (z.B. Bundesagentur für Arbeit, statistische Landes- und Bundesämter, externe Dienstleister).

Eine der bezirklichen Datenkoordination vergleichbare, ressortübergreifende Position zum Management der internen Datenbestände existiert auf Ebene der Hauptverwaltungen bislang nicht. In jenen Verwaltungen, die stärker datengestützt arbeiten, konzentrieren sich die Verantwortlichkeiten in entsprechend spezialisierten Abteilungen oder Referaten (etwa in der Abteilung III „Geoinformation“ der Senatsverwaltung für Stadtentwicklung und Wohnen oder im Referat IA „Gesundheits- und Sozialinformationssysteme“ der Senatsverwaltung für Gesundheit, Pflege und Gleichstellung). In den meisten Fällen aber ist das Auswerten und Aufbereiten von Daten Teil der alltäglichen Verwaltungsarbeit und findet themenbezogen in den jeweiligen Referaten statt.

Bereitstellung intern und extern

Die von den Hauptverwaltungen und nachgeordneten Behörden erfassten Datenbestände werden in vielen Fällen für eine Weiternutzung aufbereitet, wobei sich zwischen verwaltungsinternen und öffentlichen Angeboten unterscheiden lässt. Zu den verwaltungsinternen Angeboten zählt etwa der „Abgestimmte Datenpool“ des Amts für Statistik oder das „Planungsraumbezogene Informationssystem für Monitoring und Analyse“ (PRISMA) der Senatsverwaltung für Stadtentwicklung und Wohnen. Diese Datenbanksysteme sind aus dem Intranet der Berliner Verwaltung zugänglich und erlauben es den Beschäftigten, aktuelle Daten und Auswertungen zu beziehen und für ihre eigene Arbeit zu nutzen.

PRISMA - Kartenansicht



Planungsraumbezogenes Informationssystem für Monitoring und Analyse (PRISMA)

Daneben existieren verschiedene Portalangebote von Verwaltungen und Ämtern, die auch öffentlich zugänglich sind. Genannt seien an dieser Stelle das Geodatenportal „FIS-Broker“ der Senatsverwaltung für Stadtentwicklung,¹⁶ das „Gesundheits- und Sozialinformationssystem“ (GSI) der Senatsverwaltung für Gesundheit, Pflege und Gleichstellung¹⁷ oder das „Statistische Informationssystem Berlin Brandenburg“ (StatIS-BBB) des Amtes für Statistik.¹⁸ Die Inhalte dieser Portale machen einen Großteil des bestehenden Open Data-Angebotes im Land Berlin aus. Gleichwohl handelt es sich hierbei um individuelle Lösungen der jeweiligen Verwaltungen und Ämter, die teils älter sind als die Open Data-Strategie des Landes. Entsprechend heterogen sind diese hinsichtlich (Meta-)Datenstrukturen, Nutzerführung und Erscheinungsbild.

16 www.stadtentwicklung.berlin.de/geoinformation/fis-broker

17 www.gsi-berlin.info

18 www.statistik-berlin-brandenburg.de/datenbank/inhalt-datenbank.asp

GSI SYSTEM Senatsverwaltung für Gesundheit, Pflege und Gleichstellung Berlin

Start Sitemap Kontakt Newsletter Hilfe Impressum

Gesundheits- und Sozialinformationssystem

Einfache Suche

GESUNDHEITSDATEN

Die dreispaltige Tabelle listet für die Kategorie Gesundheitsdaten alle Bereiche mit ihren Themen und Unterthemen auf und ermöglicht den schnellen Zugriff auf alle Titel der zugehörigen Dokumente, deren Anzahl im Klammerausdruck angegeben wird.

Bereich (Dokumente)	Thema (Dokumente)	Unterthema (Dokumente)
Ausbildung im Gesundheitswesen (81)	Auszubildende im Gesundheitswesen (46)	Auszubildende im Gesundheitswesen (46)
		Weiterbildung (1)
	Erlaubnis zum Führen einer Berufsbezeichnung ausgewählter Gesundheitsfachberufe (11)	Erlaubnis zum Führen einer Berufsbezeichnung ausgewählter Gesundheitsfachberufe (11)
	Erteilte Approbationen und Berufserlaubnisse (16)	Erteilte Approbationen und Berufserlaubnisse (16)
	Weiterbildung (8)	Weiterbildung (8)
Ausgaben und Finanzierung (194)	Gesundheitsausgaben nach Ausgabenträgern und Leistungsarten (73)	Gesetzliche Krankenversicherung (GKV) (23)
		Gesundheitsausgaben nach Ausgabenträgern und Leistungsarten (1)
		SGB XII (BSHG), BVG und LPIFG (PflegeG) (36)
		weitere Ausgabenträger (13)
	Indikatoren zur Sozialversicherung (24)	Beitragsätze, Grundlohnsommen und ... (4)

Erweiterte Suche
Überblick
Hinweise
Aktuell!
RSS-Feed
Ordner
Gesundheitsdaten
_GBE-Indikatoren
Erläuterungen
Sozialdaten
Sozialstatistiken
Berichte
Publikationsreihen
Stichwort
Info zur Suche
FAQ (häufige Fragen)
Abkürzungen

ServiceStadt Berlin
BSUN OPEN MTA

Gesundheits- und Sozialinformationssystem Berlin (GSI)

Amt für Statistik

Die vielleicht wichtigste Datenbezugsquelle für die Berliner Verwaltungen ist das Amt für Statistik Berlin Brandenburg (AfSBB).

statistik Berlin Brandenburg StatIS-BBB Das Statistische Informationssystem Berlin Brandenburg

Katelog Meine Tabellen Tabellenansicht Diagrammansicht Kartenansicht Benutzerdefinierte Gruppen Suche

Datenquelle: Binnenwanderung im Land Berlin bis 2013 Tabelle heruntergeladen: Excel 2007 (xlsx) max 16.384 Spalten x 65.000 Zeilen

Daten abrufen Tabelle leeren Tabelle speichern Tabelle drucken Papierkorb

Jahre nach Herkunftsgebiet nach Zielgebiet

Filter: Standardansicht: Anzahl
Schichten: 2013 2
X Jahre

Zielerfassung: 308 (12 x 12 x 2) insgesamt, 144 (12 x 12 x 1) angezeigt.

Zielgebiet	01 Mitte	02 Friedrichshain-Kreuzberg	03 Prenzlau	04 Charlottenburg-Wilmersdorf	05 Spandau	06 Steglitz-Zehlendorf	07 Tempelhof-Schöneberg	08 Neukölln	09 Treptow-Köpenick	10 Marzahn-Hellersdorf
Herkunftsgebiet										
01 Mitte	-	2.984	3.821	2.420	1.326	1.269	1.548	1.847	684	731
02 Friedrichshain-Kreuzberg	2.794	-	2.853	1.021	371	820	2.006	3.828	1.226	668
03 Prenzlau	3.097	2.506	-	841	378	544	805	1.021	813	989
04 Charlottenburg-Wilmersdorf	2.425	1.127	1.329	-	1.945	2.555	2.408	1.021	541	291
05 Spandau	882	431	520	1.379	-	534	582	387	390	385
06 Steglitz-Zehlendorf	1.110	625	499	1.885	558	-	2.578	803	322	213
07 Tempelhof-Schöneberg	1.782	1.427	827	2.291	721	2.988	-	2.476	529	371
08 Neukölln	1.885	2.765	1.038	1.029	613	1.067	2.662	-	1.599	472
09 Treptow-Köpenick	503	727	719	326	233	262	498	982	-	1.096
10 Marzahn-Hellersdorf	486	435	831	253	234	198	318	260	1.047	-
11 Lichtenberg	976	1.156	1.861	460	311	303	457	509	1.258	2.635
12 Reinickendorf	1.988	425	1.090	875	998	470	598	478	259	381

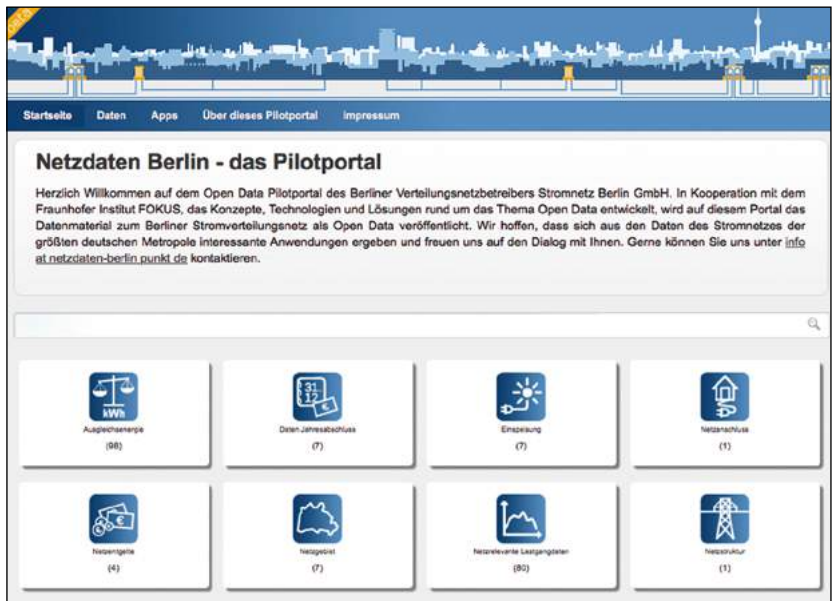
Statistisches Informationssystem Berlin-Brandenburg (StatIS-BBB)

Als zentraler Dienstleister des Landes auf dem Gebiet der amtlichen Statistik erhebt das AfSBB insgesamt über 250 Bundes- und EU-Statistiken sowie 25

Landesstatistiken aus allen Bereichen des gesellschaftlichen Lebens. Zur Anfertigung der Statistiken greift das Amt direkt auf Datenbestände anderer Verwaltungen zu. Bei der Erstellung von Bevölkerungsstatistiken wird etwa ein Datenexport des Fachverfahrens verwendet, das die Bürgerämter mit dem Landesamt für Bürger- und Ordnungsangelegenheiten (LABO) verbindet. Das AfSBB übernimmt die Aufbereitung (dazu gehört, wo nötig, auch eine algorithmische Anonymisierung personenbezogener Daten) und Auswertung und übermittelt die Resultate an verschiedene Stellen zurück. Über den „abgestimmten Datenpool“ des AfSBB können die Bezirke etwa die zur Anfertigung der Bezirksregionenprofile benötigten Indikatoren beziehen. Zudem bietet das AfSBB ein umfangreiches Datenportal an, das auch öffentlich zugänglich ist. Die dort verfügbaren Daten sind zum Großteil in maschinenlesbaren Formaten verfügbar und stehen unter einer offenen CC-BY-3.0-Lizenz.

Landeseigene Betriebe und Landesbeteiligungen

Auch die unter der Aufsicht der Senatsverwaltung stehenden Ver- und Entsorger (Berliner Wasserbetriebe, Berliner Stadtreinigung, Berliner Verkehrsbetriebe) verfügen, ebenso wie die Vattenfall-Tochter Stromnetz Berlin und der Verkehrsverbund Berlin Brandenburg (VBB), über umfassende Datenbestände zur Versorgungslage, Diensten und öffentlichen Infrastrukturen.



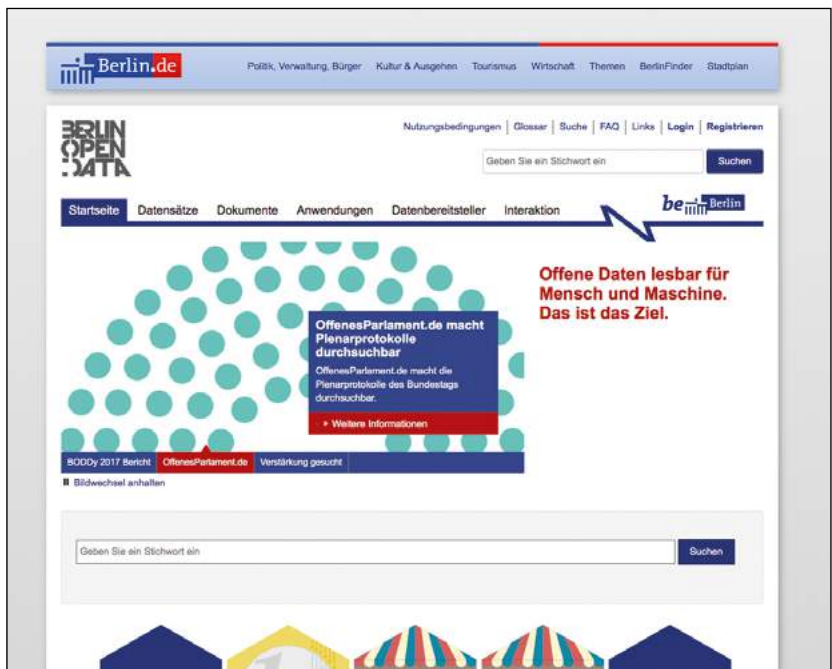
Netzdaten Berlin

Gerade mit Blick auf die Schnittstelle von Stadtentwicklung und Digitalisierung liegen auch hier große Potenziale für den Aufbau einer offenen Dateninfrastruktur. Vorreiter in diesem Bereich ist die Stromnetz Berlin GmbH, die unter „netzdaten-berlin.de“ bereits seit 2012 ein eigenes Open Data Portal betreibt. Auch die vom VBB publizierten Fahrplandaten des öffentlichen Personennah-

verkehrs gehören seit Jahren zu den nachgefragtesten Datensätzen der Berliner Open Data-Landschaft. Die Berliner Stadtreinigung und die Berliner Wasserbetriebe haben erste Daten unter freier Lizenz veröffentlicht, sind aber noch nicht zu einer regelmäßigen Veröffentlichung übergegangen. Eine Fortführung und ein Ausbau dieser Aktivitäten wird gegenwärtig im Rahmen des InfraLAB-Expertenkreises „Shared Data“ diskutiert.

Das Berliner Datenportal

Mit dem Angebot „daten.berlin.de“ betreibt die Senatsverwaltung für Wirtschaft, Energie und Betriebe (zuvor Senatsverwaltung für Wirtschaft, Technologie und Forschung) seit mehr als fünf Jahren das offizielle Open Data Portal des Landes Berlin. In diesem Portal werden die von verschiedenen Verwaltungen und Behörden dezentral bereitgestellten Datenbestände an zentraler Stelle katalogisiert und mit entsprechenden Zusatzinformationen (Metadaten) versehen. Im Portal werden nicht die Daten selbst, sondern lediglich eine Verknüpfung hinterlegt, während die Speicherung und Pflege der Daten weiterhin lokal von den zuständigen Stellen verantwortet werden. Diesen Stellen obliegt es auch, ihre Daten nach der Veröffentlichung im Datenportal zu registrieren.



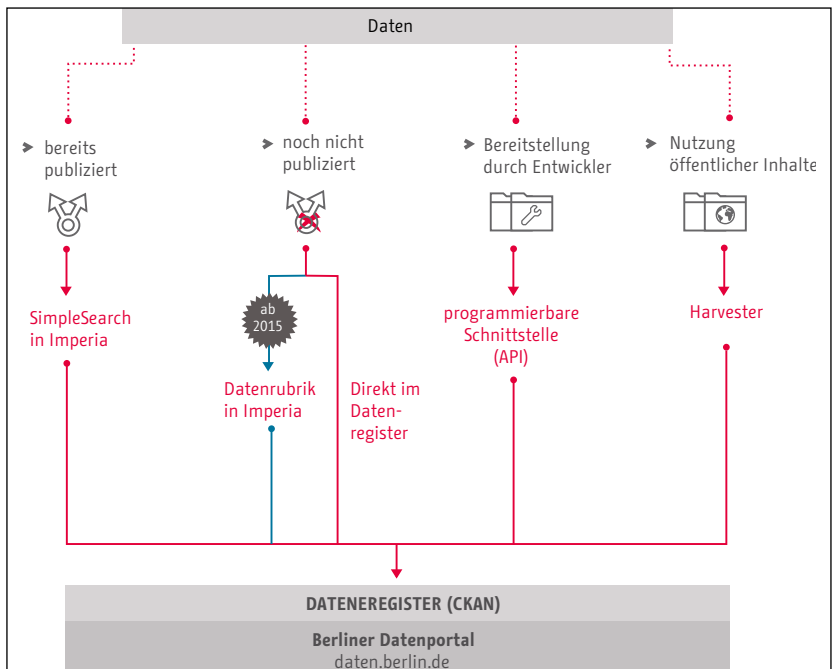
Berliner Datenportal

Dem Berliner Datenportal liegt das Prinzip der Föderation zu Grunde. Das bedeutet, dass einerseits die Datenbestände untergeordneter Instanzen (etwa von Bezirken oder den Portalen anderer Senatsverwaltungen) in einem automatisierten Prozess in das Portal übernommen werden können („Harvesting“). Andererseits bedient das Berliner Portal übergeordnete Instanzen, etwa die

Datenportale des Bundes (govdata.de) und der Europäischen Union (european-dataportal.eu). Im Dezember 2017 sind rund 1.300 Datensätze im Berliner Datenportal registriert. Als größte Datenbereitsteller treten gegenwärtig jene Häuser in Erscheinung, die bereits über eigene Datenportale verfügen: Die Senatsverwaltungen für Stadtentwicklung und Wohnen, respektive Umwelt und Verkehr (FIS-Broker) die Senatsverwaltung für Gleichstellung, Gesundheit und Pflege (GSI), sowie das Amt für Statistik (Statist BBB).

Veröffentlichungsvorgang

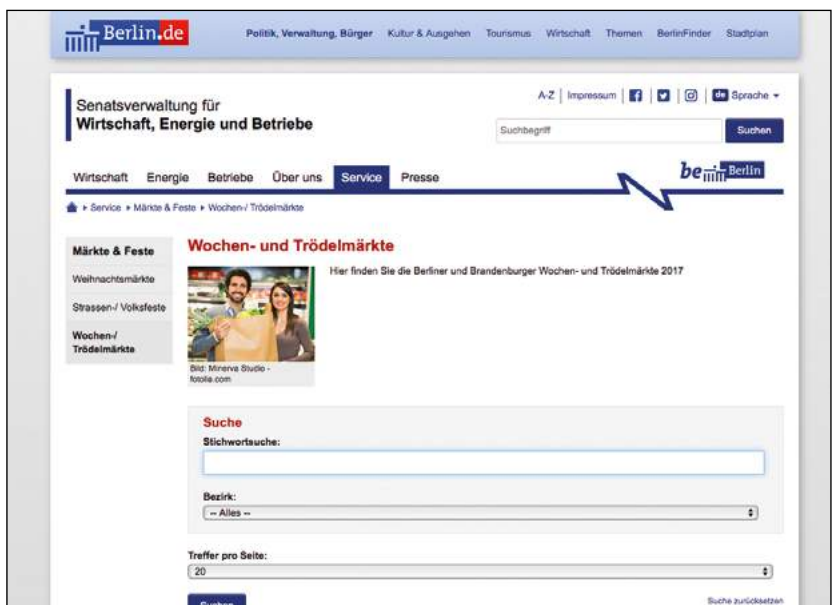
Neben dem automatisierten „Harvesting“ anderer Portale bestehen Möglichkeiten zur manuellen Registrierung von Datensätzen im Berliner Datenportal.



Veröffentlichungsmöglichkeiten im Datenportal im Überblick

Der häufigste Weg führt über das einheitliche Content Management System „Imperia“ der Berliner Verwaltungen. Imperia ist Grundlage des Web-Angebotes „Berlin.de“ und damit auch der Online-Auftritte sämtlicher Haupt- und Bezirksverwaltungen. Als Teil der Basisfunktionalität bietet Imperia über das Tool „SimpleSearch“ die Möglichkeit, strukturierte Tabellen im Backend zu hinterlegen, die dann auf der Webseite als Suchformulare, Listen oder Karten dargestellt werden. So lassen sich aus tabellarischen Verwaltungsdaten (etwa aus einer Liste mit Standorten) einfach zu bedienende Informationsangebote für Bürger*innen generieren.

Durch eine zusätzlich geschaffene Schnittstelle lassen sich die Tabellen, die hinter SimpleSearch-Formularen liegen, in wenigen Schritten auch im Berliner Datenportal veröffentlichen. Dazu müssen lediglich in einem separaten Reiter die zugehörigen Metadatenfelder ausgefüllt und eine Freigabe erteilt werden. Vor allem für die zahlreichen Web-Redakteur*innen des Landes Berlin liegt damit eine praktikable Lösung vor, um neben gängigen PDF-Dokumenten auch maschinenlesbare Rohdaten bereitstellen zu können. SimpleSearch akzeptiert ausschließlich Tabellen im Dateiformat CSV (Comma Separated Values). Die Einschränkung auf dieses weit verbreitete offene Dateiformat hat den Vorteil, dass bereits beim Upload eine erste formale Vereinheitlichung der Datenbestände erzwungen wird. Über RESTful-Webservices ist es anschließend möglich, die einmal bereitgestellten Daten auch in anderen offenen Formaten (XML, JSON etc.) abzufragen.

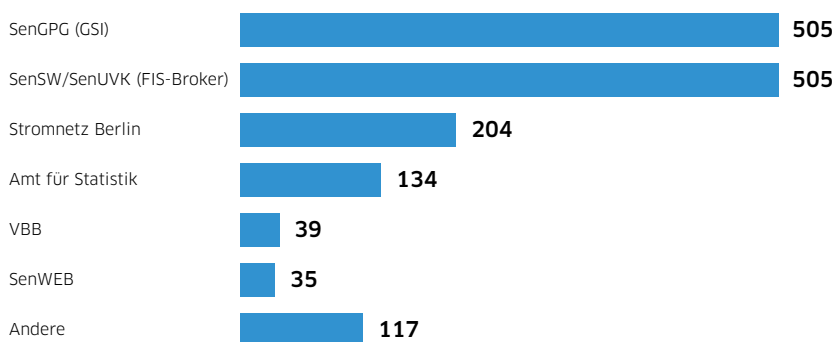


SimpleSearch-Formular

Wenn Daten nicht im CSV-Format vorliegen oder aus anderen Gründen der Umweg über SimpleSearch nicht zielführend ist, können mit einem entsprechenden Zugang Metadaten auch direkt im Berliner Datenregister eingetragen werden. Nicht zuletzt besteht die Möglichkeit, Fachanwendungen und andere Software via Schnittstelle (API) direkt in das Datenportal publizieren zu lassen. Von dieser technisch avanciertesten Möglichkeit wird bislang noch so gut wie kein Gebrauch gemacht, es ist aber absehbar, dass sie in Zukunft eine wichtigere Rolle spielen wird.¹⁹ Dies würde jedoch voraussetzen, dass auch auf Seiten der in der Verwaltung genutzten Fachverfahren und Datenbanken eine entsprechende Schnittstelle implementiert ist.

19 Zur „API-fizierung“ offener Daten als Best Practice vgl. bitkom, S. 59f.

Ein Blick auf die bisherigen Aktivitäten im Berliner Datenportal macht deutlich, dass ein Großteil der dort publizierten Rohdaten aus anderen Portalangeboten stammt. Mit 1214 von insgesamt 1547 im Backend registrierten Datensätzen²⁰ kommen gegenwärtig fast 80 Prozent des Bestandes aus den Portalen GSI (Senatsverwaltung für Gesundheit, Pflege und Gleichstellung, 505 Datensätze), FIS-Broker (Senatsverwaltung für Stadtentwicklung und Wohnen/Senatsverwaltung für Umwelt, Verkehr und Klimaschutz, 505 Datensätze) und Netzdaten-Berlin.de (Stromnetz Berlin, 204 Datensätze). Weitere 134 Datensätze steuert das Amt für Statistik bei. Auch wenn die genannten Portalangebote meist umfangreicher und folglich nur in Auszügen im registriert sind, so scheint das Föderationsprinzip doch im Grundsatz zu funktionieren: Das Berliner Datenportal konzentriert andere Angebote und bietet damit eine zentrale Anlaufstelle für offene Verwaltungsdaten aus Berlin.



Datenbereitsteller von daten.berlin.de (eigene Darstellung)

Gleichzeitig verdeutlicht diese Übersicht, dass die Möglichkeit zur Bereitstellung offener Daten im Berliner Datenportal von einem beträchtlichen Teil der Verwaltungen noch nicht oder nur zögerlich genutzt wird. So haben bislang nur sechs der zehn Senatsverwaltungen und sieben der zwölf Bezirksämter Daten im Portal registriert (siehe Übersicht im Anhang). Auch in qualitativer Hinsicht gibt es Verbesserungspotenzial: Die Verwendung einheitlicher Standards bei den Datenstrukturen und Metadaten ist noch nicht durchgängig etabliert. Gerade bei den bezirklichen Datenbeständen ist das Angebot bislang fragmentiert und kaum aufeinander abgestimmt. Einige Bezirke veröffentlichen Haushalts- oder Bebauungspläne, andere Bezirksamtsbeschlüsse, wieder andere einzelne Geodaten mit Standorten öffentlicher Einrichtungen.


²⁰ Für unsere Analyse haben wir am 28.11.2017 einen JSON-Dump direkt aus dem Datenregister erhalten.

Fazit

Seit der Veröffentlichung der Berliner Open Data Strategie und dem Start des Berliner Datenportals im Jahr 2012 ist das Angebot an offenen Verwaltungsdaten in Berlin kontinuierlich gewachsen. Viele Verwaltungen und deren Beschäftigte haben sich in den letzten Jahren erstmals mit den Potenzialen offener Daten auseinandergesetzt, einige auch mit der Veröffentlichung begonnen. Mit der Verabschiedung des E-Government-Gesetzes ist das Thema Open Data fest auf Landesebene verankert. Neu geschaffene Stellen für Datenkoordination in den Bezirken sind Ausdruck der gewachsenen Bedeutung des Datenmanagements in der öffentlichen Verwaltung. Wichtige Grundlagen für den Aufbau einer offenen Dateninfrastruktur sind damit geschaffen.

Gleichzeitig bestehen vielfältige Möglichkeiten zur Weiterentwicklung einer kontinuierlichen und qualitativ hochwertigen Datenbereitstellung. In der Fragmentiertheit des bestehenden Angebotes spiegelt sich die Komplexität der Berliner Verwaltungsstruktur, aus einer nutzenorientierten Perspektive wäre jedoch eine stärkere Vereinheitlichung wünschenswert. Durch eine strategische Koordination und Qualitätssicherung der Veröffentlichung könnte die Weiterverwertbarkeit von Verwaltungsdaten enorm gesteigert werden. Ging es in der ersten Phase nach dem Start des Portals darum, die Anzahl der verfügbaren Datensätze zu erhöhen, so müssen in einem nächsten Schritt vor allem qualitative Aspekte berücksichtigt werden. Denn erst durch verlässliche Qualitätsstandards und regelmäßige Pflege werden offene Daten produktiv nutzbar.

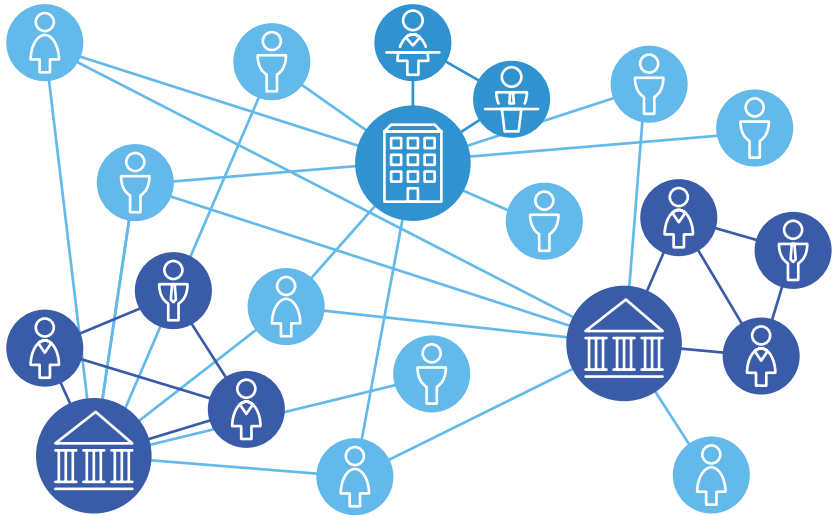
Im Rahmen zahlreicher Interviews mit Beschäftigten von Bezirks- und Hauptverwaltungen haben wir versucht, die zentralen Herausforderungen auf dem Weg zu einer systematischeren Datenbereitstellung zu identifizieren. Drei Handlungsfelder, die sich während unserer Gespräche mit den Mitarbeiter*innen als wesentlich für die Themen Datenmanagement und -veröffentlichung herausstellten, haben wir im Folgenden unter den Begriffen „Koordination“, „Orientierung“ und „Unterstützung“ kategorisiert und näher ausgeführt. Aus ihnen lässt sich abschließend ein Vorgehensmodell zum Aufbau einer offenen Dateninfrastruktur ableiten.



3 Bedarfe der Verwaltung bei der Bereitstellung von Daten

3.1

Koordination: Wer ist eigentlich zuständig?



Zuständigkeit innerhalb einzelner Verwaltungen

Die effektive Gestaltung von Wissensmanagement und Informationsaustausch ist für viele große Organisationen eine Herausforderung. Die öffentliche Verwaltung bildet hier keine Ausnahme. Bei Digitalthemen wird dieses Problem besonders deutlich, da in vielen Berliner Verwaltungen bislang keine klaren Zuständigkeiten für damit zusammenhängende Aufgabenfelder bestehen. Schon die Identifikation geeigneter Ansprechpartner*innen für das Aufgabenfeld „Datenmanagement und -bereitstellung“ erwies sich im Rahmen unserer Untersuchung mitunter als schwierig. In einigen Fällen wurden wir an die IT-Administration verwiesen, in anderen an die Öffentlichkeitsarbeit, an die Prozesssteuerung oder an einzelne Beschäftigte der Fachämter, die innerhalb der jeweiligen Häuser den Ruf hatten, besonders „datenaffin“ zu sein.

„Leider ist es [...] noch häufig so, dass eine Hand nicht weiß, was die andere macht, gerade mit Blick auf Digitalisierung und Daten. Da hat jede Abteilung so über die Jahre eigene Verfahren entwickelt und dann ist so ein Amt hier ja schon so groß, dass man vieles nicht mitbekommt.“

In diesem Gewirr aus (vermeintlichen) Zuständigkeiten zeigt sich eine grundlegende Schwierigkeit bei der Umsetzung des Querschnittsthemas Open Data: Einerseits sind nahezu alle Beschäftigten und Abteilungen in irgendeiner Weise davon betroffen, andererseits fehlt bislang in den meisten Verwaltungen eine klare institutionelle Verortung. Ob eine solche Verortung zielführend sein kann, ist umstritten: Häufig wird argumentiert, dass sich die Digitalisie-

rung nicht an eine Fachabteilung delegieren lässt, sondern einen Kulturwandel erfordert, der letztlich die Organisation in Gänze betrifft. In der gegenwärtigen Situation lässt sich jedoch vielerorts eine Art Lähmungserscheinung beobachten, die nach Ansicht vieler Beschäftigten einer mangelnden strategischen Koordination und Steuerung geschuldet ist: Zwar besteht durchaus Interesse an der Veröffentlichung von Daten, aber angesichts hoher Auslastung und unklarer Verantwortlichkeit erhält das Thema gegenwärtig geringe Priorität.

Die fehlende Zuständigkeit resultiert auch darin, dass in den meisten Verwaltungen kein wirklicher Überblick über die eigenen Daten besteht. Innerhalb der Fachabteilungen laufen alltägliche Prozesse sowohl in technischer als auch in organisatorischer Hinsicht weitgehend isoliert voneinander ab, so dass jeweils eigene, von anderen Bereichen abgeschottete „Daten-Silos“ entstehen. Nicht selten entstehen durch dieses Vorgehen auch Redundanzen: Ähnliche, aber nicht deckungsgleiche Datensätze existieren in verschiedenen Bereichen, ohne dass darüber Kenntnis besteht. Ein rudimentärer Datenaustausch findet zwar zwischen bestimmten Abteilungen statt (etwa via E-Mail oder über geteilte Verzeichnisse), erfolgt aber eher einzelfallabhängig und „auf Zuruf“.

Durch die Anforderung, Datenbestände unabhängig von der fachlichen Verortung in einheitlicher Form bereitzustellen, sehen sich Verwaltungen folglich mit einer eher ungewohnten Perspektive konfrontiert. Die übergreifende Abstimmung ist in erster Linie eine koordinatorische Herausforderung. Der Großteil der Befragten war dann auch der Ansicht, dass die Bereitstellung offener Verwaltungsdaten sie nicht in erster Linie vor technische, sondern vor organisatorische Probleme stellt. Zum einen gibt es Bedarf nach Koordinationsstellen innerhalb einzelner Ämter, zum anderen aber auch nach einem übergreifenden Austausch, insbesondere zwischen Bezirken und Senatsverwaltung.

Gute Erfahrungen wurden, wie bereits erwähnt, auf Bezirksebene mit der Einrichtung einer dezidierten Datenkoordination als Teil der OE SPK gemacht. Insbesondere der Aufbau von bezirklichen Datenbibliotheken als Teil des Aufgabenspektrums der Datenkoordination könnte eine wichtige Grundlage für eine systematische und regelmäßige Veröffentlichungspraxis werden. Dafür ist es jedoch notwendig, dass Liefervereinbarungen mit den datenliefernden Stellen der Bezirke geschlossen und diese auch eingehalten werden. Eine Unterstützung der Leitungsebene ist für dieses Vorhaben unumgänglich.

Eine zentrale, bereichsübergreifende Anlaufstelle für die Archivierung, Aufbereitung und Weiterverarbeitung verwaltungseigener Daten hat weitere Vorteile: In den Fachabteilungen fehlen oft die Ressourcen für weiterführende datenbasierte Analysen. Sowohl für die Verwaltung selbst, als auch für Politik und Bürger*innen werden derartige Auswertungen aber zukünftig an Bedeutung gewinnen. Als Querschnittsposition fungiert die Datenkoordination als Bindeglied, kann übergreifende Zusammenarbeit fördern und Mehrwerte sicht-

bar machen, die aus der Verschneidung der Daten verschiedener Abteilungen entstehen. Nicht zuletzt ist sie Kontaktstelle für datenbezogene Anfragen anderer Behörden, Medien oder Zivilgesellschaft.

In vielen Verwaltungen fehlt es bislang an einer klaren Zuständigkeit für das Thema Datenkoordination und -veröffentlichung. Empfehlenswert wäre, in jedem Haus eine dafür verantwortliche Stelle einzurichten. Diese müsste mit der nötigen Befugnis ausgestattet werden, Daten einzelner Fachämter und Referate zu sichten und für eine weitere Nutzung aufzubereiten. Durch die lokale Zusammenführung verteilter Bestände kann ersichtlich werden, welche Daten innerhalb der Verwaltung in welcher Form vorliegen. Für eine systematische Datenveröffentlichung ist dies eine wichtige Grundlage.

Austausch zwischen Verwaltungen und mit der Open Data-Community

Das Fehlen von klaren Zuständigkeiten macht sich nicht nur innerhalb einzelner Verwaltungen negativ bemerkbar, sondern auch und gerade in der übergreifenden Koordination. Ein beträchtlicher Teil der Verwaltungsdaten verbleibt nicht über die gesamte Dauer des „Data Lifecycle“ an einem Ort, sondern zirkuliert zwischen verschiedenen Einrichtungen. Ein Beispiel: Ein bezirkliches Bürgeramt erfasst Daten und leitet diese an das zuständige Landesamt weiter. Dort werden die Daten mit denen anderer Bürgerämter zusammengeführt. Das Amt für Statistik greift auf diesen aggregierten Datensatz zu, führt weitere Verarbeitungsschritte durch (etwa eine datenschutzkonforme Anonymisierung) und liefert ihn im Rahmen entsprechender Vereinbarungen zurück an Bezirke und Senatsverwaltungen.

„Wir brauchen jetzt Verbindlichkeit: Klare Absprachen und vielleicht auch Verpflichtungen. Wenn wir von der ‚Smart City Berlin‘ reden dann geht es ja auch um eine vernetzte, um eine integrierte Stadt. Und dann geht es nicht, dass bei den Daten jeder sein eigenes Süppchen kocht.“

Nun aber stellt sich die Frage, wer in diesem Fall als der für die Veröffentlichung zuständige Datenhalter gilt, zumal ein Datensatz über die verschiedenen Stationen diverse Modifikationen und Überarbeitungen erfährt. Das Grundprinzip, Daten so nah wie möglich an der Quelle abzugreifen, lässt sich nicht uneingeschränkt anwenden, da der Datensatz durch die diversen Aufarbeitungsschritte durchaus an Wert gewinnt, eine genuine Quelle also gar nicht ohne Weiteres auszumachen ist. Letztlich müssen hier pragmatische Entscheidungen hinsichtlich der Zuständigkeit getroffen werden, die dann auch mit der nötigen Autorität durchgesetzt werden können. Im Rahmen unserer Gespräche wurde in diesem Zusammenhang wiederholt die Einrichtung einer „Landesdatenkoordination“ angeregt, die als Letztinstanz eines übergreifenden Datenmanagements über entsprechende Entscheidungsbefugnisse verfügt.

Neben der Festlegung von Zuständigkeiten und Veröffentlichungszyklen besteht auch durch die Anforderung einer behördenübergreifenden Standardisierung von Daten Bedarf nach stärkerer Koordination. Beschäftigte von Hauptverwaltungen berichten, dass Daten, die von Bezirken abgefragt werden, mitunter in unterschiedlichen, zueinander inkompatiblen Formaten und Strukturen geliefert werden. Zum Teil liegt die Ursache für diese Fragmentierung in der heterogenen Systemlandschaft der Verwaltung, in der unterschiedliche Softwarelösungen für gleiche Aufgaben verwendet werden. Häufig fehlt es aber auch schlicht an Vorgaben hinsichtlich zu verwendender Standards.

Auch beim Aufbau bezirklicher Datenbibliotheken, der in einigen Fällen bereits begonnen hat, sollte frühzeitig auf intensive Abstimmung und Wissensaustausch geachtet werden, damit nicht erneut untereinander inkompatible Inselösungen entstehen. Zudem ist das Föderationsprinzip des Berliner Datenportals zu berücksichtigen: Eigene Plattformen sollten so gestaltet werden, dass ihr Bestand automatisiert in übergeordnete Portale übernommen werden kann. Viele der Befragten äußerten darüber hinaus den Wunsch nach einer regelmäßigen Abstimmungsrunde, die sich mit der Sichtung und Qualitätssicherung interner Datenbestände befasst und dazu geeignet wäre, gemeinschaftliche Vorhaben im Bereich Open Data voranzubringen.

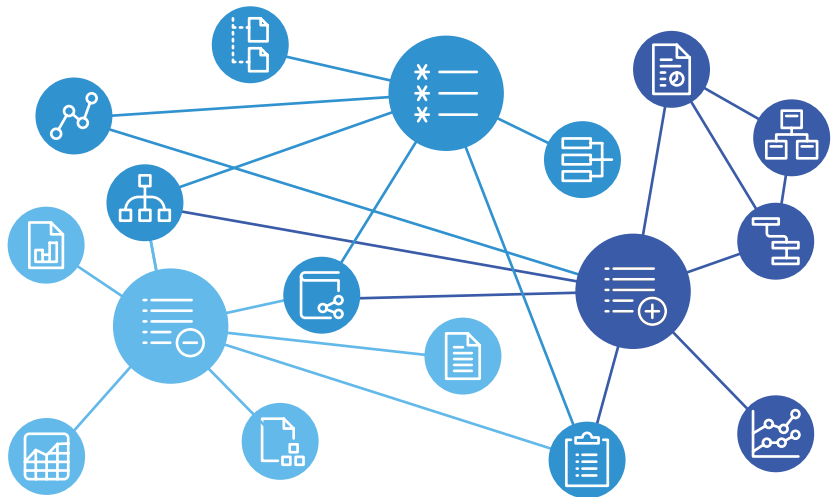
Bei der qualitativen Verbesserung der Datenveröffentlichung spielt auch der Austausch zwischen Verwaltung und der Open Data-Community eine wichtige Rolle. Um eine Weiternutzung von offenen Daten zu fördern, sollten Nutzergruppen möglichst früh in den Veröffentlichungsprozess einbezogen werden. Bei der regelmäßigen Erfassung von Bedarfen, aber auch bei der Orientierung an etablierten Open Data-Standards (z.B. Open311²¹) kann die Verwaltung von den Perspektiven und der Expertise der Community profitieren. Diese Ansicht teilten viele unserer Interviewpartner*innen, wenngleich auch vereinzelt Skepsis bezüglich kultureller Unterschiede zwischen Verwaltung und „Digital Natives“ zum Vorschein kam. Gerade für den Abbau wechselseitiger Vorbehalte könnte aber ein intensiverer Austausch, bis hin zur Durchführung konkreter Kooperationsprojekte zielführend sein.

21 Open311 ist ein ursprünglich aus den USA stammender, offener Standard für standortbasiertes Nachverfolgen von Bürgeranliegen und -beschwerden (civic issue tracking), vgl. www.open311.org. Inzwischen wird der Standard auch von einigen deutschen Städten genutzt, etwa in Bonn (<http://anliegen.bonn.de/seiten/open-data-mit-open311>). Berlin verfügt zwar mit „Ordnungsamt Online“ ebenfalls über ein digitales Anliegenmanagement, unterstützt den Standard jedoch bislang nicht.

Wo Daten im Rahmen der Weiterverarbeitung und Aufbereitung durch mehrere Verwaltungen zirkulieren, ist die Zuständigkeit der Veröffentlichung oft ungeklärt. Zudem ist, bedingt durch die vergleichsweise große Autonomie einzelner Stellen beim Datenmanagement, die übergreifende Harmonisierung von Datenbeständen in Berlin eine besondere Herausforderung. Um Abhilfe zu schaffen, sollte auch auf Landesebene eine Koordinations- und Steuerungsinstanz mit entsprechenden Befugnissen geschaffen werden. Auch der Austausch zwischen Behörden sowie mit weiteren Stakeholdern sollte durch Netzwerkangebote intensiviert und verstetigt werden.

3.2

Orientierung: Was soll überhaupt veröffentlicht werden?



Angebots- oder Nachfrageorientierung?

Beim Thema Open Data sehen sich Verwaltungsbeschäftigte gegenwärtig mit einer recht unscharfen Anforderung konfrontiert: Sie sollen „ihre“ Datenbestände in maschinenlesbarer Form aufbereiten und für Dritte zur Verfügung stellen. Die Festlegung klarer Zuständigkeiten ist dabei nicht die einzige Herausforderung. Vielmehr ist häufig bereits unklar, was genau einen eigenständigen Datensatz eigentlich ausmacht und anhand welcher Kriterien beurteilt werden kann, ob Daten zur Veröffentlichung geeignet sind.

Auch wenn in der Verwaltung zweifellos große Mengen an Daten erhoben und verarbeitet werden, so zeigt sich, dass die pauschale Forderung nach der Bereitstellung „sämtlicher“ Daten die hybride Verwaltungsrealität verfehlt. Viele Daten liegen nicht einfach in publizierbarer Form vor, sondern werden nach Bedarf beschafft und aufbereitet, etwa zur Beantwortung von Anfragen oder der Erstellung von Berichten. Anlassbezogene Recherchen, bei denen

Daten aus bestehenden Systemen extrahiert oder von externen Quellen angefragt werden, gehören für viele Beschäftigte zum Alltag. Die Berliner Open Data-Strategie macht hingegen bislang keine spezifischen Vorgaben zu Art, Umfang oder Regelmäßigkeit der zu veröffentlichenden Daten.

„Wenn bestimmte Informationen benötigt werden, mache ich mich auf die Suche. Ich weiß schon, wo ich was beschaffen kann. Aber wenn mir niemand sagt, wonach ich suchen soll, kann ich auch nichts finden. ‚Veröffentlichen Sie eben einfach ihre Daten‘ – Das ist mir zu pauschal, mit so einer Aufforderung kann ich nichts anfangen.“

Viele Beschäftigte äußerten, dass sie gerne einen stärkeren Beitrag zum Berliner Open Data-Angebot leisten würde, dazu jedoch klarere Informationen nötig seien, welche Daten in welcher Form gewünscht werden. Genau diese Spezifizierungen sind jedoch schwierig, solange eine wirkliche Übersicht über die vorhandenen Daten fehlt. Dieses „Henne-Ei“-Problem ist im Zusammenhang mit Open Data bekannt und viel diskutiert. Einseitig auflösen lässt es sich nicht. Stattdessen wird eine stärkere Systematisierung der Veröffentlichung nur iterativ und in einem fortlaufenden Austausch zwischen Anbieter- und Nachfrageseite erfolgen können.

Erforderlich ist dazu einerseits eine größere Transparenz über die in der Verwaltung verwendeten Verfahren, Datenbanken und Systeme. Dies wäre auch aus Sicht der Beschäftigten wünschenswert, da die mangelnde Übersicht über vorhandene Datenbestände bislang zu Verzögerungen und Mehrarbeit führt. Andererseits lassen sich auch von Nachfrageseite strategische Zielsetzungen formulieren, welche Daten verwaltungsseitig erhoben und bereitgestellt werden sollten. Solche Ziele müssen sich nicht ausschließlich am bereits Bestehenden orientieren: Mittelfristig wäre denkbar, dass Daten von großem öffentlichem oder wirtschaftlichem Interesse von Seiten der Verwaltung gezielt erhoben werden, um sie für eine Weiternutzung zugänglich zu machen.

Verschiedentlich kritisierten Verwaltungsbeschäftigte, dass es an klaren Vorgaben fehlt, welche Daten überhaupt veröffentlicht werden sollen. Eine solche Konkretisierung ist dringend notwendig, um die abstrakte Forderung nach einer allgemeinen Datenbereitstellung in die Tat umzusetzen. Entsprechend gilt es jetzt Prozesse zu erarbeiten, die für die Praxis der Datenveröffentlichung die nötige Orientierung geben.

Systematische Inventur von Datenbeständen auf Verwaltungsseite

Um einen Überblick über das mögliche Datenangebot zu gewinnen, wären zunächst die vorhandenen Prozesse, Fachverfahren und Datenbanken daraufhin zu analysieren, welche Daten dort gegenwärtig bezogen, gespeichert oder aufbereitet werden. Zu Beginn gibt es hier kaum eine Alternative zu einem kleinteiligen Vorgehen auf Ebene einzelner Referate und Fachabteilungen.

Jedoch verspricht eine Daten-Inventur, die unter einheitlichen Maßgaben durchgeführt und transparent dokumentiert wird, für sich genommen schon einen großen Erkenntnisgewinn: Das Nachvollziehen von Datenflüssen in der Berliner Systemlandschaft ermöglicht es, Quellen und Knotenpunkte zu identifizieren, schafft Transparenz und hilft so dabei, Zuständigkeiten zu definieren und Mehrfachaufwand bei der späteren Veröffentlichung zu vermeiden.

Auf eine solche Inventur kann in einem nächsten Schritt eine Priorisierung der identifizierten Datenbestände folgen. In unseren Gesprächen wurde deutlich, dass ein Anspruch auf Vollständigkeit bei der Datenveröffentlichung gegenwärtig (noch) unrealistisch ist. Auch andere Städte und Kommunen folgen hier dem „Pareto-Prinzip“, nach dem sich schon mit 20% der Daten 80% des Mehrwerts realisieren lässt.²² Statt nur auf Quantität zu setzen, sollte die Bereitstellung qualitativ hochwertiger Basisdaten stärker im Fokus stehen. Welche Datensätze in den jeweiligen Arbeitsfeldern am relevantesten sind, können die Beschäftigten selbst in der Regel gut einschätzen. Darüber hinaus bestehen hilfreiche Werkzeuge zur Priorisierung von Verwaltungsdaten, etwa der Leitfaden des britischen Open Data Institute²³ oder die von der Stadt Wien entwickelte Bewertungsmatrix.²⁴ Auch der mit einer Veröffentlichung verbundene Aufwand sollte in die Bewertung einfließen.

„Die Frage, welche Daten wir eigentlich haben, ist überhaupt nicht trivial. Es klingt zwar so, als müssten wir das wissen, aber das wird ja nirgends zentral erfasst und folglich gibt es hier eigentlich niemanden, der Ihnen das beantworten kann.“

Inventur und Priorisierung können schließlich zur Grundlage eines Veröffentlichungsplans werden, der Termine und Anforderungen an die zu liefernden Daten festlegt. Mehr Verbindlichkeit bei der Datenbereitstellung wäre in der jetzigen Situation für viele Beschäftigte eine Erleichterung, weil sie bestehende Unsicherheiten reduziert und den mit der Veröffentlichung verbundenen Aufwand planbarer macht. Daten, die sich häufiger ändern, müssen natürlich auch in regelmäßigen Abständen aktualisiert werden, denn eine aktuelle Datenbasis ist für viele Anwendungen unerlässlich. Hier kann ein Veröffentlichungsplan für die Anwenderseite die nötige Planungssicherheit schaffen.

Um das Datenangebot der Berliner Verwaltung zu verbessern, ist zunächst eine umfassende Inventur und anschließende Priorisierung von Daten auf Ebene einzelner Referate und Abteilungen nötig. Ziel sollte die Erstellung eines Veröffentlichungsplans sein, der Ziele und Fristen spezifiziert.

22 Vgl Krabina/Lutz 2016, S. 19.

23 <https://theodi.org/guides/prioritise-open-data-to-drive-global-development>

24 Vgl Krabina/Lutz 2016, S. 25.

Gezielte Ergänzung der Datenveröffentlichung aus Nachfragesicht

Eine systematische Inventur der verwaltungsinternen Bestände ist ein notwendiger Schritt auf dem Weg zu einer umfassenderen Praxis der Datenveröffentlichung. Um die Wertschöpfung aus offenen Daten zu fördern, reicht eine bloße Orientierung am bereits verfügbaren Angebot aber nicht aus. Auch viele der Befragten äußerten Zweifel, ob die ursprünglich zur rein internen Verarbeitung vorgesehenen Datensätze in dieser Form für Außenstehende gewinnbringend nutzbar sein könnten. So enthalten manche Datensätze interne Verweise oder Kennziffern, die eine Vertrautheit mit den teils sehr komplexen Organisationsstrukturen und Abläufen einer Verwaltung voraussetzen.

Gleichzeitig gehört es für Beschäftigte der Verwaltung durchaus zum Alltag, externe Anfragen datengestützt zu beantworten und die eigenen Datenbestände in diesem Zuge für die Öffentlichkeit verständlich aufzubereiten. Dies ist etwa bei regelmäßigen Zuarbeiten zu parlamentarischen Anfragen aus dem Abgeordnetenhaus der Fall. Auch für die Erstellung von Berichten oder die Beantwortung von Presseanfragen bestehen oft eingespielte Arbeitsabläufe, um Daten gezielt zu beschaffen. Diese vorhandenen Routinen können auch genutzt werden, um die Datenveröffentlichung stärker an bestehender Nachfrage auszurichten.

„Wenn wir das Thema Datenbereitstellung ernst nehmen und uns auch einen Nutzen davon erhoffen, sollten wir schon den Anspruch haben, Relevantes zu liefern und uns nicht einfach auf das beschränken, was sowieso bei der täglichen Arbeit abfällt.“

Während eine interne Inventur Datenbestände zu Tage fördern kann, die sich mit vergleichsweise geringem Aufwand veröffentlichen lassen, kann andererseits die Formulierung konkreter Bedarfe dazu dienen, Lücken zu schließen und bestimmte Informationen mit hoher Relevanz für die Bevölkerung gezielt zu beschaffen. Eine fortschrittliche Dateninfrastruktur mit entsprechendem Nutzen für Gesellschaft und Wirtschaft sollte perspektivisch an den Bedarfen der verschiedenen Nutzergruppen ausgerichtet sein. Entsprechend wäre genauer zu prüfen, welche Datenbestände aus relevanten Bereichen wie Verkehr, Energie, Wohnen, Demographie, Gesundheit Klimaschutz, Ver- und Entsorgung etc. einen Mehrwert bieten könnten, um sie im Zweifel gezielt zu beschaffen und für die Veröffentlichung aufzubereiten.

Zur Spezifikation solcher Bedarfe empfiehlt sich ein Blick auf bestehende datenbasierte Monitoring-Projekte. Genannt seien hier etwa das EU-Vorhaben „CityKEYS“, in dem eine Liste von 73 Indikatoren zur vergleichenden Entwicklung von Stadtentwicklungsprozessen ausgearbeitet wurde²⁵, sowie die internationale Städtestatistik „Urban Audit“.²⁶ Auch ein regelmäßiger Abgleich mit

25 <http://nws.eurocities.eu/MediaShell/media/CITYkeystheindicators.pdf>

26 www.urbanaudit.de

einschlägigen Open Data-Rankings wie dem „Global Open Data Index“²⁷ bietet sich an, um das Datenangebot zielgerichtet auszubauen. Nicht zuletzt sollte der regelmäßige Austausch mit der Nutzercommunity dazu genutzt werden, Open Data-Bedarfe abzufragen. Eine Feedback-Funktionalität, die im Berliner Datenportal bereits existiert, setzt natürlich voraus, dass Anfragen auch zeitnah bearbeitet werden können.

Eine an Bedarfen orientierte Datenveröffentlichung macht nicht nur die Nutzung des Open Data-Angebotes attraktiver, sondern fördert auch die Motivation auf Seiten der Bereitsteller. Häufig wurde in unseren Gesprächen darauf hingewiesen, dass konkrete Zielsetzungen und Ergebnisorientierung der Motivation zur Datenbereitstellung förderlich sind. Positiv hervorgehoben wurden Web-Projekte wie der „Kiezatlas“²⁸ oder „Breitband Berlin“²⁹ die offene Verwaltungsdaten in interaktiver Form visualisieren und auswerten. Hier wurde die Bereitschaft, die zugrundeliegenden Datenbestände regelmäßig zu pflegen und zu aktualisieren als sehr viel höher eingeschätzt als bei einer nicht-zweckgebundenen Veröffentlichung.

Um die Potenziale einer offenen Dateninfrastruktur zu realisieren, sollte sich die Bereitstellung von Verwaltungsdaten auch an den Bedarfen der Nutzergruppen orientieren. Ein Abgleich mit geläufigen Indikatorensets kann dabei helfen, Datenlücken zu identifizieren und zu schließen. Bei Daten von hoher Relevanz, die bislang nicht vorliegen, sollte geprüft werden, ob diese gezielt erhoben und verfügbar gemacht werden können.

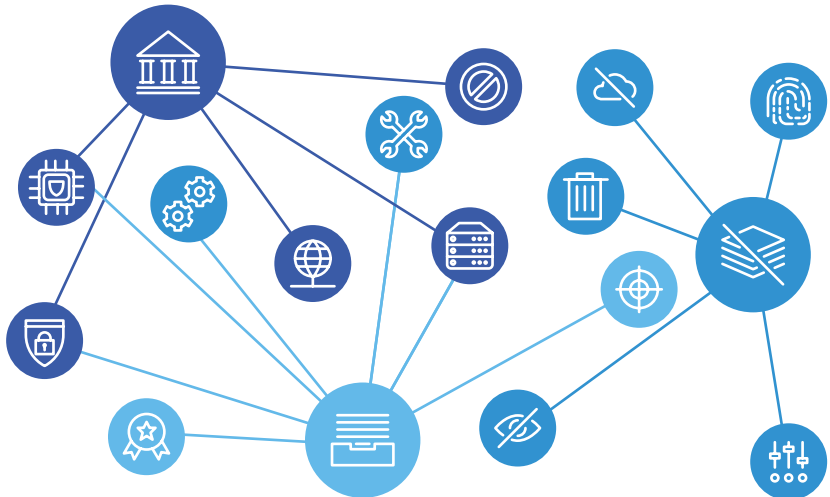
27 <https://index.okfn.org>

28 www.kiezatlas.de

29 www.breitband-berlin.de

3.3

Unterstützung: Wie kann die Veröffentlichung erleichtert werden?



Unterstützung durch eine zentrale Servicestelle

Die Veröffentlichung von Daten bedeutet für Beschäftigte eine zusätzliche Verantwortung: Was bisher nur für den internen Gebrauch bestimmt war, wird nun öffentlich einsehbar. Hinzu kommt, dass die meisten Mitarbeiter*innen der Verwaltung keinen IT-Hintergrund haben, wohingegen die Open Data-Zielgruppen (Data Scientists, Softwareentwickler*innen etc.) häufig ein hohes Maß an Digitalexpertise aufweisen. Diese Asymmetrie kann zu Verunsicherung führen und die Initiative vor allem jener Beschäftigten hemmen, die bislang keine Erfahrung mit der Datenbereitstellung haben.

„Wir haben hier noch zu wenig wirkliche Experten für dieses Thema. Oft sind es ja nur Kleinigkeiten, dass jemand nochmal draufschaute, ob das Dateiformat passt oder ob eine Tabelle in dieser Form in Ordnung ist. Aber selbst als jemand, der dem Thema offen gegenübersteht, ist das schwer zu überblicken.“

Im Rahmen unserer Gespräche wurde deutlich, dass bei der Erfassung von Datenbeständen und bei der Planung einer initialen Datenveröffentlichung nach wie vor individueller Unterstützungsbedarf besteht. Bei der Vielzahl von Aufgabengebieten und begleitenden IT-Lösungen bietet jeder Datensatz eigene Herausforderungen, die sich nur schwer verallgemeinern lassen. Trotz bestehender Informationsangebote tauchen in der konkreten Arbeit mit Datenbeständen häufig Fragen auf, für die bislang kein geeigneter Ansprechpartner verfügbar ist.³⁰

³⁰ Die Senatsverwaltung bietet unter <http://daten.berlin.de/datenbereitsteller> unter anderem einen Schritt-für-Schritt-Leitfaden zur Datenveröffentlichung an. Dieser war jedoch nicht allen Beschäftigten bekannt.

Unsicherheiten bestanden einerseits hinsichtlich rechtlicher Aspekte: Neben der Sorge, gegen Datenschutz- oder Sicherheitsbestimmungen zu verstoßen, wurde die Notwendigkeit zur Vergabe einer individuellen Lizenz für jeden veröffentlichten Datensatz als überfordernd beschrieben. Auch bezüglich der technischen Aufbereitung bestand Informationsbedarf: Woran erkenne ich ob ein Datensatz den Ansprüchen an Maschinenlesbarkeit genügt? Wie kann ein effizienter Workflow vom Datenbezug bis zur Bereitstellung aussehen? Wie können verlässliche Qualitätsprüfungen durchgeführt werden? Hier wurde der Ruf nach einer verbindlichen Servicestelle laut, die schnell und unbürokratisch entsprechende Expertise anbieten kann.

Interessanterweise betrachteten Mitarbeiter*innen, die bereits in der Datenveröffentlichung aktiv sind, die bestehenden rechtlichen und technischen Hürden als vergleichsweise gering und auch den für eine Veröffentlichung nötigen Mehraufwand als überschaubar. Dies deutet darauf hin, dass Unterstützung vor allem im Sinne einer „Starthilfe“ nötig ist. Wiederholt wurde etwa der Wunsch geäußert, Daten vor der Erstveröffentlichung einer externen Prüfung unterziehen zu lassen, oder zumindest mit klareren Vorgaben zu arbeiten, welche Kriterien die Daten zu erfüllen haben. Auch angesichts der angestrebten Harmonisierung des Open Data Angebotes scheint eine solche zentrale Qualitätskontrolle in der Initiationsphase sinnvoll zu sein.

Es besteht Bedarf nach einer „Servicestelle Open Data“, die den Veröffentlichungsprozess mit technischer Expertise begleitet und bei konkreten Fragen als Ansprechpartner fungiert. Eine solche Servicestelle könnte weitere Aufgaben rund um den Veröffentlichungsprozess übernehmen, etwa eine stärkere redaktionelle Betreuung des Datenangebotes und eine zentrale Qualitätskontrolle.

Unterstützung durch geeignete Software-Tools

Eine ganzheitliche Open Data-Strategie muss berücksichtigen, wie die Bereitstellung von maschinenlesbaren Verwaltungsdaten im Feld der Verwaltungsdigitalisierung insgesamt verortet ist und welche Abhängigkeiten und möglichen Synergien sich zwischen der Weiterentwicklung von IT-Systemen einerseits und der Datenveröffentlichung andererseits ergeben. Dies gilt insbesondere für die IT-Infrastruktur der Berliner Verwaltung, die auch bedingt durch die im E-Government-Gesetz formulierten Anforderungen vor einer grundlegenden Neustrukturierung steht. Die angestrebte Konsolidierungen der Verwaltungs-IT bieten auch für eine Systematisierung der Open Data-Veröffentlichungspraxis große Chancen und können, richtig durchgeführt, zugleich für Entlastung bei den Beschäftigten sorgen.

Bislang stellt die fragmentierte IT-Landschaft der Berliner Verwaltung ein Hindernis für eine systematische Bereitstellung offener Daten dar. Ein Großteil der Verwaltungsverfahren durchläuft nach wie vor Medienbrüche, die eine einheitliche Erfassung und ein effizientes Datenmanagement erschweren. Um mit der Vielzahl der erforderlichen manuellen Arbeitsschritte zurecht zu kommen, haben sich viele Beschäftigte individuelle Lösungen geschaffen, häufig in Form rudimentärer Excel-Makros oder kleinerer Access-Datenbanken. Diese Werkzeuge erleichtern den konkreten Arbeitsalltag, erweisen sich jedoch oft als wenig nachhaltig. Im Falle einer Abwanderung von Personal führen individuelle Lösungen zu Wissensverlust und zur Entstehung von „Datengräbern“. Eine stärkere Standardisierung der Prozesse von Datenerfassung und -verarbeitung bis zur letztlichen Veröffentlichung wäre hier geboten.

„Gerade am Anfang ist Datenveröffentlichung mit viel manueller Arbeit verbunden. Und dabei passieren dann natürlich auch Fehler, die eigentlich vermeidbar wären. Ich denke, es wäre eine große Erleichterung, wenn wir Daten einfach auf Knopfdruck publizieren könnten und zum Beispiel eine Qualitätsprüfung automatisch durchgeführt werden würde.“

Viele Beschäftigte äußerten im Gespräch den Wunsch nach technischen Hilfsmitteln, die etwa eine Datenbereitstellung „auf Knopfdruck“ ermöglichen könnten. Nicht nur unter dem Aspekt der Arbeitserleichterung und des Wegfalls repetitiver, fehleranfälliger Tätigkeiten, sondern auch angesichts des angestrebten Ziels einer Harmonisierung von Datenformaten und -strukturen sollten technische (Teil-)Lösungen in Erwägung gezogen werden, die im Prozess der Datenbereitstellung verwendet werden können. Dabei muss es in einem ersten Schritt nicht unbedingt um die Aufsetzung großer Systeme gehen – in vielen Fällen wären schon kleine Werkzeuge, etwa zur Metadatenprüfung oder zum automatischen Datenaupload eine Erleichterung. Entscheidend wäre jedoch, das Angebot solcher Werkzeuge zentral zu koordinieren, um von der daraus resultierenden Standardisierung auch tatsächlich profitieren zu können.

Beispiele für solche Werkzeuge existieren bereits: So bietet etwa das Open Data Institute (ODI) in London ein Web-Tool an, mit dem CSV-Dateien vor dem Upload unkompliziert auf Fehler geprüft werden können.³¹ In Österreich werden im Rahmen der „ADEQUATE“-Initiative (Analytics & Data Enrichment to Improve the Quality of Open Data) eine Reihe freier Werkzeuge entwickelt, die der Verbesserung der Qualität offener Verwaltungsdaten dienen.³² Auch das Erproben von ETL-Prozessen (Extract, Transform, Load), bei denen heterogene Daten in einem automatisierten Vorgang zusammengeführt, konvertiert und veröffentlicht werden, könnte in vielen Fällen Mehrwerte bieten. Neben einschlägigen kommerziellen Lösungen existieren hier leistungsstarke Open

31 <https://csvlint.io>

32 www.adequate.at

Source-Produkte.³³ In Berlin muss also das Rad nicht neu erfunden werden. Vielmehr wäre im Einzelfall zu prüfen, welche bereits vorhandenen Werkzeuge sich in bestehende Arbeitsabläufe integrieren lassen.

Mittelfristig ist denkbar, dass große Teile des Veröffentlichungsvorgangs automatisiert ablaufen. Bereits jetzt nutzen viele Beschäftigte individuelle technische Lösungen, die ihre Arbeit mit Daten erleichtern. Die Verwendung solcher Werkzeuge erfordert perspektivisch aber stärkere Koordination. Die Möglichkeiten zur Automatisierung einzelner Arbeitsschritte (Datenbezug, Konvertierung, Qualitätsprüfung, Upload) wäre zu prüfen und schrittweise auszubauen. Wo immer möglich sollte dabei auf bereits bestehende Werkzeuge zurückgegriffen werden.

Auf dem Weg zu einem integrierten Datenmanagement

Perspektivisch sollten Schritte in Richtung einer Automatisierung der Datenaufbereitung und -bereitstellung direkt bei den verwendeten Fachverfahren und Datenbanken ansetzen. Bisher ist dies nur in Ausnahmefällen möglich, weil viele Fachverfahren keine unkomplizierte Weiterverarbeitung der Daten erlauben. Eben ein solcher „Open Data Ready“-Export in geeignete Formate wäre für viele Beschäftigte der einfachste Weg, ohne großen Mehraufwand Datensätze für eine Veröffentlichung direkt aus Fachanwendungen zu generieren. Auch eine direkte Publikation aus Anwendungen ins Datenportal ist theoretisch denkbar. Die manuell nötigen Arbeitsschritte könnten damit auf ein Minimum reduziert werden, während gleichzeitig höhere Konsistenz und Regelmäßigkeit bei der Bereitstellung gewährleistet wären.

„Gefühlt ist es gerade so, dass zehn verschiedene Leute dieselben Daten an zwölf verschiedenen Orten speichern. Das ist natürlich nicht optimal. Jedes Mal wenn ich eine Kollegin anrufe und sie frage, ob sie mir eine Excel-Tabelle mailen kann, wünsche ich mir ein vernünftiges ‚Data Warehouse‘, wie es ja große Unternehmen längst haben.“

Angesichts der großen Potenziale, die in einer stärkeren Automatisierung der Datenbereitstellung liegen, wäre es jetzt umso wichtiger, den Themen Datenmanagement und Open Data im Rahmen anstehender IT-Konsolidierungen einen höheren Stellenwert einzuräumen. Bei der Überarbeitung und Neuentwicklung von Verwaltungsanwendungen sollte die Möglichkeit zum Datenexport in offenen Formaten unbedingt berücksichtigt werden. Idealerweise sollten Fachverfahren zukünftig quelloffen (Open Source) entwickelt werden, um die Transparenz der Datenverarbeitung weiter zu fördern und eine leichtere Implementierung offener Schnittstellen auch nachträglich noch zu gewährleis-

33 <https://frictionlessdata.io/software>

ten.³⁴ Auch Datenlieferverträge mit externen Dienstleistern sollten zukünftig so geschlossen werden, dass eine Open Data-konforme Veröffentlichung unkompliziert möglich ist.

Aufgrund der wachsenden Bedeutung von Datenmanagement und -analyse für öffentliche Verwaltungen sollten zudem Grundlagen für ein integriertes Datenbanksystem geschaffen werden, das den übergreifenden Zugriff und die Verwaltung auch größerer, teils zeitkritischer Datenmengen (bis hin zu sensorgenerierten Echtzeitdaten) erlaubt. Mit einem entsprechenden Rechtemanagement ausgestattet, könnte eine umfassende Plattformarchitektur einen Großteil der bisherigen Insellösungen ersetzen, ein höheres Maß an Datenschutz und -sicherheit gewährleisten und die Grundlagen für einen effizienten Datenaustausch schaffen.

Mit Blick auf die in diesem Bericht geschilderten Herausforderungen wäre der Aufbau eines solchen Systems ein komplexes und zweifellos ambitioniertes Vorhaben. Mit der Einrichtung von Strukturen zur Standardisierung und Qualitätssicherung offener Verwaltungsdaten können jedoch schon heute wichtige Grundlagen geschaffen werden um das Datenmanagement der Berliner Verwaltung insgesamt zu systematisieren. Umgekehrt kann natürlich die Berliner Open Data-Strategie ungemein von technischen Weiterentwicklungen der IT-Architektur profitieren: Die Implementierung offener Programmierschnittstellen (APIs), die externen Nutzer*innen einen direkten Zugriff auf ausgewählte Datenbestände geben, könnte die Beschäftigten der Verwaltung perspektivisch vom Aufwand einer manuellen Bereitstellung entlasten.

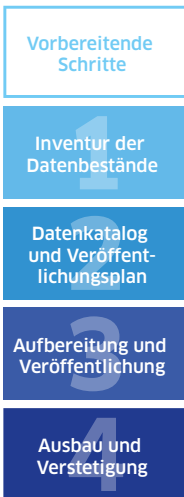
Der Aufbau einer offenen und qualitativ hochwertigen Dateninfrastruktur wird letztlich nur auf Grundlage eines fortschrittlichen Datenmanagements möglich sein. Bei der Weiterentwicklung der IT-Architektur kann Open Data durchaus Motivation und Treiber sein. Mit den zur Veröffentlichung notwendigen Schritten werden zugleich die Voraussetzungen für eine übergreifende Plattform zum Management von Verwaltungsdaten geschaffen. Mit der Realisierung einer solchen Plattform könnte in Zukunft der manuelle Aufwand bei der Datenveröffentlichung reduziert und die Datenqualität weiter verbessert werden.

34 Stand Dezember 2017 sind erst 74 der insgesamt 316 verwendeten Fachverfahren quelloffen, vgl. <http://bit.ly/2DiaTq6>

The background features a blurred financial report with line graphs and tables. A large, semi-transparent geometric shape, composed of a light blue triangle and a dark blue triangle meeting at a point, is overlaid on the image. A red triangle also points towards the same central point from the right side. The overall aesthetic is professional and data-oriented.

4

Vorgehensmodell



Das folgende Vorgehensmodell beschreibt einen Prozess zur qualitativen und quantitativen Verbesserung der Datenbereitstellung in der Berliner Verwaltung. Es orientiert sich in Teilen an vergleichbaren Modellen anderer Städte und Kommunen, wurde jedoch an die in den Interviews identifizierten Bedarfe der Berliner Verwaltung angepasst.³⁵ Zudem wurde auf eine erneute Erörterung einschlägiger Open Data-Grundlagen verzichtet. Stattdessen konzentriert sich das Modell auf pragmatische Schritte, um ausgehend vom gegenwärtigen Ist-Zustand ein höheres Maß an Systematisierung zu erreichen.

Angesichts knapper personeller und finanzieller Ressourcen in den Verwaltungen muss darauf geachtet werden, den Veröffentlichungsvorgang nicht durch Überformalisierung zu verkomplizieren. Da sich Rahmenbedingungen schnell ändern, sind bei der Umsetzung zudem Flexibilität und Augenmaß gefragt. Das Vorgehensmodell ist folglich nicht als starrer Rahmen zu verstehen, sondern soll lediglich Anregungen zur Weiterentwicklung der Berliner Open Data-Infrastruktur bieten.

Vorbereitende Schritte

Klare Verantwortlichkeiten schaffen

Eine regelmäßige und standardisierte Datenweitergabe erfordert klare Zuständigkeiten und Kommunikationswege zwischen einzelnen Abteilungen. Dafür sollten in jeder Berliner Verwaltung „Open Data-Beauftragte“ ernannt werden, die einerseits für die interne Koordination von Datenveröffentlichungen zuständig sind, andererseits in der verwaltungsübergreifenden Kommunikation als Kontaktpersonen fungieren.³⁶ Diese Stellen sollten zudem befugt sein, Liefervereinbarungen mit internen Abteilungen zu schließen. Auf Bezirksebene bietet es sich an, diese Position mit der bezirklichen Datenkoordination zu verbinden. Auch Online-Redaktionen oder E-Government-Verantwortliche sind für diese Aufgabe geeignet.

- Ernennung einer/eines „Open Data-Beauftragten“ in jeder Behörde
- Einrichtung einer „Landesdatenkoordination“ auf Ebene der Hauptverwaltungen

³⁵ Ein Überblick über verschiedene Reifegrad- und Vorgehensmodelle aus dem Bereich E-Government findet sich bei Fath Allah et al (2014). Einen aktuelleren Literaturüberblick zur Implementation von Open Data-Strategien bietet Thomas (2017). Aus dem deutschsprachigen Raum sind insbesondere die „Leitlinien Open Government Data“ der Stadtverwaltung Bonn und das „Open Government-Vorgehensmodell“ der Stadt Wien zu nennen.

³⁶ Wo die Einrichtung zusätzlicher Verantwortlichkeiten aus formalen Gründen schwierig ist, kann eine solche Rolle zunächst auch informell von engagierten Beschäftigten übernommen werden. Wichtig ist in einem ersten Schritt, möglichst viele Verwaltungen an Bord zu bekommen, um bestehende Informationsdefizite zu beheben.

Servicestelle Open Data einrichten

Eine zentrale Beratungs- und Servicestelle Open Data kann fachliche Expertise hinsichtlich rechtlicher und technischer Fragen bündeln und als verlässlicher Ansprechpartner für konkrete Fragen im Veröffentlichungsprozess dienen. Eine solche Stelle sollte die Katalogisierung und Aufbereitung der Datenbestände begleiten sowie bei der laufenden Qualitätssicherung unterstützen.³⁷ Auch kann sie die strategische Weiterentwicklung einer Dateninfrastruktur vorantreiben und als Schnittstelle zur Open Data Community fungieren. Konkret können die Aufgaben umfassen:

- Monitoring von Bedarfen auf Anbieter- und Nutzerseite
- Unterstützung bei der Katalogisierung und Qualitätssicherung von Daten
- Bereitstellen von Bildungsmaterialien, Durchführung von Trainings
- Erfolgsmessung der Open Data-Strategie anhand quantitativer und qualitativer Indikatoren (Anwendungen, Nutzerzahlen, Datenqualität und -aktualität, Anzahl Datensätze)
- Veröffentlichung von Vorlagen und Referenzdatensätzen
- Bereitstellung von Werkzeugen zur Datenveröffentlichung, -aufbereitung und -nutzung
- Durchführung von Pilotprojekten, die den Nutzen offener Daten demonstrieren

Vernetzung fördern

Offene Daten leben vom Austausch, nicht nur in technischer Hinsicht. Um in der komplexen Verwaltungslandschaft zu einheitlichen Lösungen zu kommen, ist eine laufende Abstimmung und Koordination der am Prozess Beteiligten nötig. Regelmäßige Treffen der Open Data-Beauftragten können neben dem verwaltungsübergreifenden Austausch auch dazu genutzt werden, externe Perspektiven einzubinden, etwa durch die Einbeziehung von Open Data-Nutzer*innen, Wissenschaft, Unternehmen oder Verantwortlichen aus anderen Kommunen. Begleitet werden sollte die Vernetzung durch den Aufbau einer gemeinsamen Wissensbasis, die kollaborativ erweitert und transparent dokumentiert wird.³⁸

- Einrichtung einer regelmäßigen Abstimmungsrunde der Open Data-Beauftragten (AG Open Data)
- Auftakt zu einem kontinuierlichen Dialogformat mit externen Stakeholdern
- Aufbau einer kollaborativen Wissensplattform (Open Data-Wiki o.ä.)

³⁷ Eine solche Servicestelle kann sowohl innerhalb als auch außerhalb der Verwaltung angesiedelt sein. Beispiele für verwaltungs-externe Beratungsstellen wären etwa das Open Data Institute (ODI) in London oder das Zentrum für Verwaltungsforschung (KDZ) in Wien.

³⁸ vgl. auch Lucke 2017, S. 4f.

Inventur der Datenbestände

Eine systematische und regelmäßige Datenveröffentlichung erfordert einen transparenten Überblick über die Datenbestände der Verwaltung. Entsprechend sollte ein Katalog der wichtigsten Berliner Verwaltungsdaten erstellt werden. Ein solches Vorgehen gilt inzwischen als Best Practice und wurde bereits in verschiedenen Städten erfolgreich umgesetzt.³⁹

Der mit einer Dateninventur verbundene Aufwand hängt stark von dem angestrebten Detailgrad ab. Während die vollständige Erfassung sämtlicher Datenbestände ein langwieriges und komplexes Unterfangen sein kann, lässt sich ein erster Überblick über maßgebliche Systeme relativ leicht erstellen und veröffentlichen. Um ein gewisses Maß an Einheitlichkeit zu erreichen, empfiehlt es sich, die Dateninventur der einzelnen Referate und Abteilungen mit zentraler Koordination durchzuführen. Folgende Ziele verfolgt die Dateninventur:

- Katalogisierung der wichtigsten Informationssysteme und Datenbestände des Landes Berlin
- Erfassung der wichtigsten Attribute zu den gespeicherten Daten
- Überblick über einzelne Verarbeitungsschritte und Datenflüsse
- Identifikation zentraler Basisdaten als Grundlage für eine Priorisierung

Die Dateninventur soll im ersten Schritt keinen Anspruch auf Vollständigkeit erheben. Vielmehr soll ein Katalog mit den relevantesten Verwaltungsdaten aufgebaut und schrittweise erweitert werden. Zur Identifikation geeigneter Datensätze sollten verwaltungsinterne Prozesse und Systeme, Datenflüsse und bisherige Veröffentlichungen berücksichtigt werden.

Analyse von Prozessen

In vielen Routinevorgängen der Verwaltung werden Daten erhoben und verarbeitet. Dazu zählen:

- Bearbeiten von Anträgen und Formularen
- Verwaltung von Adressen und Standorten
- Evaluation von Projekten
- Umgang mit öffentlichen Geldern
- Erstellung von Berichten

³⁹ Kürzlich etwa in Helsinki, wo ein [Katalog mit 843 Informationssystemen und Datenbanken](#) veröffentlicht wurde. Zur Dateninventur als Best Practice im Zusammenhang mit Open Data vgl. www.w3.org/2013/share-psi/bp/pomd. In Deutschland erfasst das Verbundprojekt „MetaVer“ Informationen zu Datenbeständen und Fachverfahren von bislang fünf Bundesländern (Brandenburg, Bremen, Hamburg, Mecklenburg-Vorpommern und Sachsen-Anhalt), vgl. www.metaver.de

Analyse von Datenquellen

Datenquellen sind digitale Systeme und Verfahren, in denen Daten erhoben, verarbeitet oder gespeichert werden. Dazu zählen unter anderem:

- Datenbanksysteme
- Fachverfahren
- Online-Anwendungen
- Excel-basierte Verfahren

Zur Identifikation der verwendeten Softwarelösungen können vorhandene IT-Bestandspläne zur Hilfe genommen werden.

Analyse von Datenflüssen

Einzelne Verwaltungen fungieren häufig als Knotenpunkte in einem Netz aus Datenflüssen. Entsprechend lässt sich untersuchen:

- Welche Daten werden von externen Quellen bezogen? (Bundesbehörden, nachgeordnete Ämter, Unternehmen, Dienstleister, freie Träger etc.)
- Welche Daten werden eigenständig erhoben oder aufbereitet? (Messungen, Evaluationen, Abfragen etc.)
- Welche Daten werden an externe Stellen weitergegeben? (andere Landesverwaltungen, Bundesbehörden, Medien & Öffentlichkeit etc.)

Analyse bisheriger Veröffentlichungspraxis

Viele Verwaltungen und Behörden veröffentlichen heute schon Daten, teils in maschinenlesbarer Form, teils in Form von Berichten (PDF- oder Druckerzeugnisse):

- Welche Daten werden bereits veröffentlicht?
- Welche Daten liegen Berichten zugrunde, die bereits veröffentlicht werden?

Identifikation konkreter Datensätze

In der Folge lässt sich detaillierter bestimmen, welche maschinenlesbaren Dateien in Form von Tabellen, Geodaten, Hierarchien, etc. aus den bestehenden Verfahren und Systemen extrahiert werden können. In vielen Fällen sind Datenquelle und Datei identisch (z.B. Excel-Tabellen), aus komplexeren Systemen hingegen lassen sich durch unterschiedliche Abfragen nahezu beliebig viele Dateien erzeugen. Hier ist fallweise und mit Augenmaß zu entscheiden. Grundsätzlich sind Basisdaten wichtiger als weiterführende Aggregationen.



Datenkatalog und Veröffentlichungsplan

2

Datenkatalog

Im Anschluss an die Inventur sollte ein Datenkatalog erstellt und veröffentlicht werden. Hier sollten die identifizierten Datensätze aufgeführt und mit weiteren Zusatzinformationen versehen werden, z. B. zur zuständigen Stelle und Aktualität. Die abgefragten Informationen sollten kongruent zum Metadaten-schema DCAT-AP.de sein, so dass diese Daten im Falle einer Veröffentlichung nicht noch einmal erhoben werden müssen. Mindestens sollte erfasst werden:

- Titel des Datensatzes
- Kurzbeschreibung
- Aktualität/Aktualisierungszyklus
- Primärquelle
- Lizenz
- Raumbezug
- Zeitbezug
- Dateiformat

Priorisierung

Die katalogisierten Daten sollten hinsichtlich einer Veröffentlichung priorisiert werden. Dabei sind erwarteter Nutzen, Aufwand und vorhandene personelle Ressourcen zu berücksichtigen, aber auch technische und rechtliche Voraussetzungen.⁴⁰ Falls Ausschlusskriterien erfüllt sind, die eine Veröffentlichung grundsätzlich verhindern, ist dies entsprechend zu vermerken. Konkret sollten folgende Punkte berücksichtigt werden:

- Ist eine Weiternutzung der Daten möglich?
- Ist die Qualität der Daten vertretbar oder lässt sich die Qualität mit überschaubarem Aufwand auf ein vertretbares Niveau bringen?
- Können aus der Verschneidung der Daten mit weiteren Datensätzen Synergieeffekte entstehen?
- Liegen die Daten in Open Data-konformen Formaten vor oder lassen sie sich leicht in solche Formate exportieren?
- Gibt es grundlegende Einwände, die gegen eine Veröffentlichung sprechen? (Personenbezug, öffentliche Sicherheit, fehlende Rechte)

Ergänzung aus Nutzersicht

Ergänzend zur Übersicht des Datenkatalogs lassen sich Zielsetzungen für weitere Datenerhebungen und -veröffentlichungen formulieren. Als Gegenpol zur Angebotsorientierung sollten hier explizit Fragen aus Nutzersicht gestellt werden, die mit offenen Daten beantwortet werden könnten, z.B.:

40 Ein Toolkit zur Priorisierung findet sich unter <http://bit.ly/2CSh1IB>

- Wo finde ich eine E-Ladesäule?
- Wie ist die Luftqualität in meinem Kiez?
- Welche kostenlosen Freizeitangebote gibt es in meinem Kiez?
- Wie steht es um die medizinische Versorgung in meiner Nähe?
- Wofür werden meine Steuergelder ausgegeben?
- Wann wird mein Altpapier abgeholt?

In User-Workshops oder durch Online-Befragungen können Fragestellungen gesammelt werden. Auch wenn die entsprechenden Daten noch nicht vorliegen, können diese zusätzlichen Perspektiven zum Anlass genommen werden, sie zu beschaffen und der Stadtgesellschaft zugänglich zu machen.

Veröffentlichungsplan




OPEN-GOVERNMENT-VORGEHENSMODELL

Tabelle 6: Muster eines Phasenplanes

	Phase 1	Phase 2	Phase 3	Phase 4
Einreichschluss	01.03.2016	31.05.2016	02.09.2016	15.11.2016
Redaktionsschluss	15.03.2016	14.06.2016	13.09.2016	29.11.2016
Veröffentlichung	29.03.2016	28.06.2016	27.09.2016	13.12.2016

Quelle: Stadt Wien, 2016.

Abbildung 9: Beispiel für einen Datenkatalog inkl. Bewertung und Phasenplanung

Datensatz	Kategorie	#	Datum	Zuständig	Startplanjahr	Kostengutachten	Bewertung	NEP/BI	Ansprechpartner
Geolayer Kindergärten (altsteife amalien - Hyperlink)	Stadtplan, WNF, WFS	10	Phase noch offen	MA 11	JA	A	38	JA	Alice Meislinger
Geolayer Kulturgut: Historische Karten	WMF	10	Phase noch offen	MA 9			29		Dr. Karl Fischer, MA 9
Geolayer Marktgebiete	Stadtplan, WNF, WFS	10	Phase noch offen	MA 21 B			25		Martin Murrer, MA 21B, WNF
Geolayer Flurgassen, Planklöse, Adressen	Stadtplan, WNF, WFS	10	Phase noch offen	MA 21 B	JA		39		Hanno Martin, MA 21B, WNF
Geolayer Ständige Geschäfte	Stadtplan, WNF, WFS	10	Phase noch offen	WDM		B	31	JA	Fayal Thomas, WDM
Geolayer Wohnungskategorie	Stadtplan, WNF, WFS	10	Phase noch offen	MA 21A, MA 21B		B	31		Martin Murrer
Geolayer Behinderteneinrichtungen - Erweiterung	Stadtplan, WNF, WFS	10	Phase noch offen	MA 48 (MA 48)			40		Wolfgang
Geolayer Denkmal	Stadtplan, WNF, WFS	10	Phase 2016, 2017	MA 45		A	38		Stefan Schmalzberger, Thomas
Geolayer Fahrschulen Wien	Stadtplan, WNF, WFS	10	Phase 2016, 2017	MA 65	Nein		26	Nein	Martin Hoffmann
Geolayer nord-südliche Ratha - Kulturgut	Stadtplan, WNF, WFS	10	Phase 2016, 2017	MA 7			32		Barbara Kirch
Geolayer Kunstwerke im öffentlichen Raum - Kulturgut	Stadtplan, WNF, WFS	10	Phase 2016, 2017	MA 7			32		Barbara Kirch
Geolayer Biotopekartierung	Stadtplan, WNF, WFS	10	Phase MA 8, 2017	MA 22	JA	B	33	JA	Dr. Viktoria Krüger

Datenkatalog inklusive Priorisierung und Phasenplanung der Stadt Wien

Auf Grundlage der Priorisierung sollte ein Veröffentlichungsplan erstellt werden. Der Veröffentlichungsplan sollte durch die für Open Data zuständige Senatsverwaltung unter Mitwirkung der veröffentlichenden Einrichtungen und der Servicestelle Open Data erstellt werden. Veröffentlichungstermine können in regelmäßigen Abständen (z.B. quartalsweise) angestrebt werden. Insbesondere sollte im Rahmen eines Veröffentlichungsplans festgelegt werden:

- Wer ist für die Lieferung der Daten zuständig?
- Bis wann sollen die Daten vorliegen?
- Wie häufig werden die Daten aktualisiert?
- In welchem Format sollen die Daten bereitgestellt werden?

Zur laufenden Pflege und Evaluation von Veröffentlichungsplänen existieren verschiedene Werkzeuge, etwa das in den Städten Köln und Bonn sowie im Land Brandenburg verwendete „OGD Cockpit“.⁴¹



Aufbereitung und Veröffentlichung

Um den steigenden Ansprüchen an Datenqualität und -standardisierung Rechnung zu tragen, sollte gerade bei Erstveröffentlichung von Daten eine redaktionelle Prüfung erwogen werden. Diese könnte von der erwähnten Service-stelle Open Data übernommen werden. Eine Qualitätskontrolle kann insbesondere die Harmonisierung heterogener Datenbestände unterstützen, die Qualität von Metadaten sichern und die Einhaltung von Open Data-Standards gewährleisten.

Datenqualität

Ein Blick auf das gegenwärtige Datenangebot zeigt, dass in technischer Hinsicht häufig Inkonsistenzen zwischen Datensätzen bestehen, die durch eine zentrale Prüfung behoben werden könnten. Dies betrifft insbesondere:

- Verwendung korrekter Zeichenkodierungen (z.B. UTF-8)
- Einheitliche Formatierungen von Zeit- und Ortsangaben
- Einheitliche Verwendung von Identifikatoren (z.B. IDs von Planungsräumen)
- Generelle Überprüfung von Kriterien zur leichteren Maschinenlesbarkeit (Tabellenstrukturen, Dateiformate)
- Gleiche Datenstrukturen für gleiche Daten

Metadatenmanagement

Auf den gewachsenen Standardisierungsbedarf der Metadaten von Open Data-Portalen hat die EU mit der Spezifizierung des Metadatenstandards DCAT-AP reagiert. Die Fachgruppe GovData hat im November 2016 eine direkt kompatible Ableitung dieses Standards für deutsche Datenportale beschlossen. DCAT-AP.de sollte 2018 auch im Berliner Datenportal eingeführt werden. Bereits heute besteht bei den Metadaten Verbesserungspotenzial. Geprüft werden sollte zukünftig unter anderem:

- Einheitliche Bezeichnung der veröffentlichenden Stelle
- Korrekte Vergabe geeigneter Schlagwörter (Tags)
- Verknüpfung mit über-/untergeordneten Datensätzen
- Vergabe passender Lizenzen (im Regelfall „CC-0“ oder „Datenlizenz Deutschland Zero“)

Veröffentlichung und Feedback

Nach Überprüfung und etwaigen Aufbereitungsschritten können die Datensätze im Berliner Datenportal registriert werden. Die Prüfung dient dabei nicht nur der Verbesserung der Datenqualität, sondern zielt vor allem darauf, einen Wissensbestand über Best Practices aufzubauen und diesen mit den datenhaltenden Stellen zu teilen. Einmal aufbereitete Datensätze sollten deshalb kommentiert an die Datenhalter zurückgespielt werden und als Vorlage für zukünftige Veröffentlichungen dienen.

- Feedback über erfolgte Aufbereitungsschritte an zuständige Datenhalter
- Bereitstellung der aufbereiteten Daten durch die zuständigen Datenhalter
- Erstellung von Vorlagen und Referenzdatensätzen



Ausbau und Verstetigung

Die Veröffentlichung von Verwaltungsdaten ist kein einmaliger Vorgang. Sie muss in bestehende Routinen integriert, schrittweise verbessert und an wechselnde Rahmenbedingungen angepasst werden. Der Aufbau einer offenen Dateninfrastruktur erfordert eine kontinuierliche Pflege und Erweiterung des Angebots, aber auch des Netzwerks der beteiligten Akteure. Die Weiterentwicklung sollte nach dem „Open by Default“-Prinzip erfolgen: Neu anfallende Verwaltungsdaten sollten grundsätzlich unter offener Lizenz veröffentlicht werden, sofern keine triftigen Gründe dagegen sprechen. Auch in Verträgen mit externen Datenlieferanten sollte eine entsprechende Klausel verankert werden.

Regelmäßigkeit und Aktualisierung

Zuverlässigkeit und Aktualität sind essentielle Aspekte einer offenen Dateninfrastruktur. Entsprechend wichtig ist eine Verstetigung der Veröffentlichung durch verbindliche Verantwortlichkeiten und Termine. Welche Aktualisierungszyklen sinnvoll sind, hängt von den jeweiligen Beständen ab. Bei dynamischen Daten sollte geprüft werden, wie die Bereitstellung automatisiert (Schnittstellen oder Daten-Pipelines) erfolgen kann.

- Festlegung von Zuständigkeiten für Aktualisierung einzelner Datensätze
- Terminierung regelmäßiger Veröffentlichungs- und Aktualisierungszyklen
- Wo möglich Entwicklung von Schnittstellen (APIs) und Data Pipelines für automatisierte Aktualisierungen
- Datengetriebene Verwaltungsakte werden „open by default“

Feedback- und Fehler-Kultur etablieren

In vielen Fällen wird erst nach einer erfolgten Datenveröffentlichung ersichtlich, wo Möglichkeiten bestehen, die Datenqualität weiter zu verbessern. Feedback der Nutzer*innen und ein regelmäßiges Monitoring der Entwicklung in

anderen Städten sollten kontinuierlich dazu genutzt werden, den laufenden Prozess zu optimieren. Den beteiligten Akteuren muss vermittelt werden, dass im ersten Anlauf keine „perfekte“ Veröffentlichung erwartet wird. Wichtiger ist die Bereitschaft, auf Rückmeldungen zu reagieren und das eigene Angebot schrittweise zu optimieren.

- Open Data als Lernerfahrung für die Verwaltung begreifen
- Aufbau eines effizienten Feedback-Managements
- Regelmäßige Evaluation von Datenkatalog, Veröffentlichungsplan und Vorgehensmodell

Automatisierung vorantreiben

Das Land Berlin sollte den Aufbau einer eigenen Bibliothek von freien Open Data-Werkzeugen anstreben, die Veröffentlichung und Datenaufbereitung erleichtern. Wo immer möglich, sollten dabei bereits bestehende Tools verwendet und auf die im Land Berlin bestehenden Bedarfe angepasst werden. Sofern diese Bibliothek ebenfalls unter freier Lizenz verfügbar gemacht wird, kann sie auch von der Open Data-Community weiterentwickelt oder in anderen Kommunen eingesetzt werden. Die Entwicklung solcher Tools kann nicht nur die Datenbereitstellung unterstützen, sondern auch einen Lernprozess und Kooperationen mit externen Entwickler*innen in Gang setzen.

- Exemplarische Erprobung der Automatisierung einzelner Schritte und Verfahren
- Monitoring internationaler Best Practices und Toolkits
- Aufbau einer offenen und frei nutzbaren Bibliothek von Softwarewerkzeugen zur Datenaufbereitung und -bereitstellung

Open Data-Beauftragte

Servicestelle

Vernetzung

Vorbereitung



1

Prozesse

Systeme

Datenflüsse

Berichte

Datensätze



2

Datenkatalog

Priorisierung

Ergänzung

Veröffentlichungsplan



3

Datenqualität

Metadaten

Veröffentlichung



4

Aktualisierung

Feedback

Automatisierung

Ausbau und Verstetigung



5 Anhang

5.1

Literatur

- Bitkom (Hrsg.) (2017): Open Data. Neue Konzepte erfolgreich umsetzen. www.bitkom.org/noindex/Publikationen/2017/Leitfaden/Open-Data-Leitfaden/171103-Open-Data.pdf
- Both, Wolfgang/Schieferdecker, Ina (Hrsg.): Berliner Open Data-Strategie: Organisatorische, rechtliche und technische Aspekte offener Daten in Berlin. Fraunhofer Verlag. www.berlin.de/sen/inneres/service/publikationen/strategie-studie_zur_umsetzung_des_open-data-gedankens_in_der___ffentlichen_verwaltung.pdf
- Büttner, Christian (2017) Die Sozialraumorientierte Planungskoordination in der bezirklichen Verwaltung, In: Zeitschrift für amtliche Statistik Berlin Brandenburg (2/2017), S. 20-24.
- Erbstöber, Anne-Caroline (2014): Smart City Berlin. Urbane Technologien für Metropolen. Technologiestiftung Berlin. www.technologiestiftung-berlin.de/fileadmin/daten/media/publikationen/140213_Studie_SmartCity.pdf
- Fath-Allah, Abdoullah/Cheikhi, Laila/Al-Qutaish, Rafa E./Idri, Ali (2014): E-Government Maturity Models: A Comparative Study. International Journal of Software Engineering & Applications (5) 3, S. 71-91.
- Lebhart, Gustav (2016): Data Governance – viel Lärm um nichts? Strategisch-analytische Verwaltungssteuerung im Zeitalter der Digitalisierung. Vortrag beim Kongress „Effizienter Staat“, 10./11. Mai 2016, www.digitaler-staat.org/wp-content/uploads/2017/02/Lebhart.pdf
- Manske, Julia/Knobloch, Tobias (2017): Leitfaden für Datenschutz bei Open Data. Ansätze und Instrumente für die verantwortungsvolle Öffnung von Verwaltungsdaten. www.stiftung-nv.de/sites/default/files/policy_brief_leitfaden_open_data_datenschutz.pdf
- Kooperation OGD D-A-CH-LI (2015): Vision eines Daten-Ökosystems im Raum D-A-CH-LI. www.data.gv.at/wp-content/uploads/2015/05/OGD-D-A-CH-LI-Daten-Oekosystem-VISION-1.pdf
- Krabina B, Lutz B (2016) Open-Government-Vorgehensmodell - Umsetzung von Open Government Version 3.0, KDZ, Wien.
- Kuzev, Pencho (Hrsg.) (2016): Open Data. The Benefits. Das volkswirtschaftliche Potential für Deutschland. Konrad Adenauer Stiftung.
- von Lucke, Jörn (2014) Nationale Open Data-Infrastruktur. In: Engel A (Hrsg) IT-Governance in Staat und Kommunen, edition sigma, Berlin, S. 219-236.

- von Lucke, Jörn und Kooperation OGD D-A-CH-LI (2017): Maßnahmen für den Aufbau einer Open Data Kultur. www.egovernment.ch/index.php/download_file/force/1135/3631
- Seibel, Benjamin (2016). Open Data in der Praxis. Bereitsteller und Anwender offener Daten in Berlin. Technologiestiftung Berlin. <http://technologiestiftung-berlin.de/opendatapraxis>
- Senatsverwaltung für Stadtentwicklung (Hrsg.) (2009): Handbuch zur Sozialraumorientierung. Grundlage der integrierten Stadt(teil)entwicklung Berlin. www.stadtentwicklung.berlin.de/soziale_stadt/sozialraumorientierung/download/SFS_Handbuch_RZ_screen.pdf
- Technologiestiftung Berlin (Hrsg.) (2014): Digitales Gold. Nutzen und Wertschöpfung durch Open Data in Berlin.
- Thomas, Jesse (2017): Best Practices in the Implementation of Open Data at a Municipal Government Level. <https://scholarsbank.uoregon.edu/xmlui/bitstream/handle/1794/22563/Thomas2017.pdf?sequence=1&isAllowed=y>
- Wangermann, Tobias (Hrsg.) (2016): Open Data aus internationaler Perspektive. Konrad Adenauer Stiftung.

5.2

Liste der Interviewpartner*innen

- **Björn Unte**
Senatskanzlei
- **Ralf Eberhardt**
Senatsverwaltung für Wirtschaft, Energie und Betriebe
- **Olaf Franke**
Senatsverwaltung für Inneres und Sport
- **Sabine Pöhl**
Bezirksamt Lichtenberg
- **Katrin Becker**
Bezirksamt Friedrichshain-Kreuzberg
- **Frank Petersen**
Bezirksamt Marzahn Hellersdorf
- **Angelika Herda**
Bezirksamt Marzahn Hellersdorf
- **Uta Simdorn**
Bezirksamt Marzahn Hellersdorf
- **Maud Metke**
Bezirksamt Mitte
- **Christian Conrad**
Bezirksamt Neukölln
- **Richard Jäger**
Bezirksamt Pankow
- **Stefen Haberecht**
Bezirksamt Steglitz-Zehlendorf
- **Ulrich Binner**
Bezirksamt Tempelhof-Schöneberg
- **Hr. Becke**
Bezirksamt Treptow-Köpenick
- **Sandy Arndt**
Bezirksamt Treptow-Köpenick
- **Dr. Christian Rentsch**
Bezirksamt Treptow-Köpenick
- **Christoph Effing**
Amt für Statistik Berlin-Brandenburg
- **Jörn Ehlert**
Amt für Statistik Berlin-Brandenburg
- **Frank Gödicke**
Amt für Statistik Berlin-Brandenburg

5.3

Liste der Bereitsteller im Berliner Open Data-Portal

505	Senatsverwaltung für Gesundheit, Pflege und Gleichstellung (GSI)
505	Senatsverwaltung für Stadtentwicklung und Wohnen/Senatsverwaltung für Umwelt, Verkehr und Klimaschutz (FIS-Broker)
204	Stromnetz Berlin
134	Amt für Statistik Berlin Brandenburg
39	Verkehrsverbund Berlin-Brandenburg (VBB)
35	Senatsverwaltung für Wirtschaft, Energie und Betriebe
17	Senatsverwaltung für Finanzen
13	Bezirksamt Friedrichshain-Kreuzberg
12	Bezirksamt Lichtenberg
11	Senatsverwaltung für Kultur und Europa
9	Bezirksamt Marzahn-Hellersdorf
6	Gemeinsame Landeplanungsabteilung Berlin-Brandenburg
5	Landesamt für Bürger- und Ordnungsangelegenheiten
4	Berliner Forsten
4	Bezirksamt Charlottenburg-Wilmersdorf
3	Berliner Stadtreinigung (BSR)
3	Polizei Berlin
3	Senatsverwaltung für Bildung, Jugend und Familie
2	Bezirksamt Steglitz-Zehlendorf
2	Bezirksamt Treptow-Köpenick
2	Landesamt für Gesundheit und Soziales
2	Landesdenkmalamt
2	Verkehrslenkung
1	Berliner Wasserbetriebe
1	Bezirksamt Neukölln

(Stand 1.12.2017)

5.4

Links

- Beispiel: VOIS-Fachverfahren der Bürgerämter
www.vois.org/wp-content/uploads/2016/03/Bildschirmfoto-2016-03-11-um-09.47.23.png
- Auszug aus dem Bezirksregionenprofil Oberschöneweide, Seite 13
www.berlin.de/ba-treptow-koepenick/_assets/aemter-und-se/spk/bzr05-oberschoeneweide/bezirksregionenprofil-oberschoeneweide.pdf
- „Planungsraumbezogene Informationssystem für Monitoring und Analyse (PRISMA)“, Seite 15
www.ioer.de/fileadmin/internet/Oeffentlichkeitsarbeit/Veranstaltungen_2011_pdf/3_DD_FNS_2011/BangertDietrich_PlanungsraumbezogeneInformationsbereitstellungFuerMonitoringUndAnalyse-StadtentwicklungBerlin.pdf
- „Gesundheits- und Sozialinformationssystem Berlin (GSI)“, Seite 16
www.gsi-berlin.info
- „Statistisches Informationssystem Berlin-Brandenburg (StatIS-BBB)“, Seite 16
www.statistik-berlin-brandenburg.de/datenbank/inhalt-datenbank.asp
- „Netzdaten Berlin“, Seite 17
netzdaten-berlin.de
- „Berliner Datenportal“, Seite 18
daten.berlin.de/kategorie/oeffentliche-verwaltung-haushalt-und-steuern
- „Veröffentlichungsmöglichkeiten im Datenportal im Überblick (S. 19)“, Seite 19
ddaten.berlin.de/sites/default/files/Datenbereitletter.pdf
- Beispielfoto SimpleSearch-Formular“, Seite 20
www.berlin.de/sen/web/service/maerkte-feste/wochen-troedelmaerkte
- „Datenkatalog inklusive Priorisierung und Phasenplanung der Stadt Wien“, Seite 44
www.kdz.eu/de/file/18417/download

Die Technologiestiftung Berlin engagiert sich für die Entwicklung Berlins zur Hauptstadt der Digitalisierung. Sie macht die Chancen und Perspektiven deutlich, die mit dem technologischen Fortschritt verbunden sind und formuliert Handlungsempfehlungen. Außerdem unterstützt sie die Open Data-Strategie und setzt sich für eine smarte Infrastruktur ein.

Dr. Benjamin Seibel

Der Autor leitet das Ideation & Prototyping Lab der Technologiestiftung Berlin. Studium der angewandten Kulturwissenschaften an der Universität Lüneburg. Promotion in Darmstadt und Harvard zur Digitalisierung der öffentlichen Verwaltung. Seine Arbeitsschwerpunkte liegen im Bereich Verwaltungsmodernisierung, Open Data und Civic Technologies.

