

Vorlage – zur Kenntnisnahme –

Beschluss der Open Data-Strategie des Landes Berlin

Der Senat von Berlin
RBM-Skzl – VI A 3 Öz
9(0)223 - 1623

An das
Abgeordnetenhaus von Berlin
über
Senatskanzlei – G Sen –

Vorlage

- zur Kenntnisnahme -
des Senats von Berlin

über den Beschluss der Open Data-Strategie des Landes Berlin

Wir bitten, zur Kenntnis zu nehmen,
dass der Senat die Open Data-Strategie des Landes Berlin beschlossen hat:

A. Begründung:

a) Allgemeines:

- In Berlin wurde frühzeitig die Notwendigkeit erkannt, Open Data in die Verwaltung zu integrieren. Bereits 2012 wurde die erste Open Data-Strategie veröffentlicht, um diese Ziele zu verfolgen. In den vergangenen Jahren konnten die rechtlichen Rahmenbedingungen für Open Data in Berlin mit dem § 13 EGovG Bln und der Open Data-Verordnung als auch dem Betrieb des Open Data-Portals des Landes Berlin daten.berlin.de sukzessive verbessert werden.
- Rund 10 Jahre nach der ersten Strategie war es an der Zeit, einen Zwischenstand zu dem bisher Erreichten zu ziehen und neue Impulse für die zukünftige Ausrichtung von Open Data im Land Berlin zu sammeln.
- Für die Beteiligung des gesamten Open Data-Ökosystems wurde in einer ersten Phase eine Online-Umfrage auf mein.berlin.de durchgeführt. In der zweiten Phase fanden Stakeholder-Workshops gemeinsam mit der OKFN e.V. und unterstützt durch die Open Data Informationsstelle Berlin statt.

- Während des gesamten Prozesses wurden die Zwischenergebnisse auf mein.berlin.de zur Diskussion und Online-Kommentierung veröffentlicht. Auch die Endergebnisse der Online-Bürgerinnen und -Bürger-Beteiligung wurden auf der Informationsseite für die breite Stadtgesellschaft zum Nachverfolgen veröffentlicht.
- Die Handlungsfelder und Maßnahmen dieser Open Data-Strategie wurden aus den Erkenntnissen aus dem Beteiligungsprozess ausgearbeitet und für die Umsetzung in der aktuell laufenden Legislaturperiode in dieser Strategie mit Maßnahmen konkretisiert.
- Mit dieser Strategie wird erstmals im Land Berlin das Ziel formuliert und verankert, Open Data als wesentlichen Bestandteil einer modernen digitalen Verwaltung zu begreifen und entsprechend Maßnahmen zu ergreifen. Die Open Data-Strategie verfolgt damit die Vision eines damit verbundenen Verwaltungskulturwandels für eine verbesserte interne Integration von Open Data in die Verwaltungsprozesse, um insbesondere die eigene Verwaltungsarbeit mit Open Data zu erleichtern und zu verbessern. Mit dem Kulturwandel ist auch die Notwendigkeit verbunden, bestehende Prozesse zu hinterfragen und Open Data in der Praxis zum Beispiel bei der Datensammlung oder der Vergabe an externe Dienstleisterinnen und Dienstleister von Anfang an mitzudenken. Der Zuständigkeitswechsel von Open Data in der Berliner Verwaltung zur Senatskanzlei ist damit gleichzeitig eine große Chance, Open Data stärker in der IT-Ausrichtung des Landes tief zu verankern und IT-Prozesse in Richtung Open Data zu entwickeln.
- Dabei hat die Strategie den Anspruch, mit konkreten Maßnahmen die Ziele für Open Data in den nächsten fünf Jahren zu erreichen. Statt um die theoretische Formulierung von Zielen, geht es um die konkreten Umsetzungsschritte mit Zeitplänen, in denen diese Maßnahmen umgesetzt werden sollen.
- Im Land Berlin wurde die Strategie „Gemeinsam Digital:Berlin“ als sogenannte Dachstrategie aller Digitalisierungsthemen der Berliner Verwaltung verabschiedet, die in ihren Handlungsfeldern auf unterschiedliche Fachstrategien des Landes Berlins für die konkrete Umsetzung mit Projekten verweist. Die neue Open Data-Strategie wird folglich als eine Fachstrategie der Gemeinsam Digital:Berlin-Strategie verstanden. So sollen die wesentlichen Handlungsfelder mit Bezug auf die Themenbereiche „Datenkompetenz“, „Datennutzung“, „Data Governance“, „internes Datenmanagement“/öffentliche Infrastruktur das Prinzip „Open-by-default“ und die dafür notwendige technische Infrastruktur in der Behördenlandschaft innerhalb der Strategie „Gemeinsam Digital:Berlin“ in der Open Data-Strategie mit Maßnahmen konkretisiert werden. Weiterhin wurde darauf geachtet, dass mit der neuen Open Data- Strategie ebenfalls die Ziele der Bundesregierung für Open Data und der Open Data-Verordnung auf Landesebene mitverfolgt werden.

B. Rechtsgrundlage:

§ 13 Absatz 2 des Gesetzes zur Förderung des E-Government (E-Government-Gesetz Berlin - EGovG Bln), verkündet als Artikel 1 des Berliner E-Government-Gesetzes vom 30. Mai 2016 (GVBl. 2016, 282) und die dazugehörige Open Data Verordnung des Landes Berlin, die am 1.1.2021 in Kraft getreten ist.

C. Gesamtkosten:

Die Kosten in Zusammenhang mit der Open Data-Strategie sind derzeit nicht quantifizierbar.

D. Auswirkungen auf die Gleichstellung der Geschlechter:

Die Open Data-Strategie hat keine Auswirkungen auf die Gleichstellung der Geschlechter.

E. Kostenauswirkungen auf Privathaushalte und/oder Wirtschaftsunternehmen:

Die Open Data-Strategie hat keine Auswirkungen auf die Privathaushalte und Wirtschaftsunternehmen.

F. Auswirkungen auf die Zusammenarbeit mit dem Land Brandenburg:

Die Open Data-Strategie hat keine unmittelbaren Auswirkungen auf die Zusammenarbeit mit dem Land Brandenburg.

G. Auswirkungen auf den Klimaschutz und die Umwelt:

Die Open Data-Strategie hat keine Auswirkungen auf den Klimaschutz und die Umwelt.

H. Auswirkungen auf das elektronische Verwaltungshandeln:

Die Open Data-Strategie hat folgende Auswirkungen auf das elektronische Verwaltungshandeln: Die Open Data-Strategie legt den Rahmen für die Freigabe von Daten fest, wodurch ein breites Spektrum von Informationen für die Öffentlichkeit zugänglich gemacht wird. Dadurch werden mehr Daten elektronisch verfügbar gemacht, um den Bedürfnissen der Bürgerinnen und Bürger, Unternehmen und anderen Interessengruppen gerecht zu werden.

Durch die Veröffentlichung von Daten gemäß der Open Data-Strategie wird die Transparenz in der Verwaltung erhöht. Bürgerinnen und Bürger können Einblicke in staatliche Aktivitäten, Finanzen, politische Entscheidungen und andere relevante Informationen erhalten. Dies fördert

die Rechenschaftspflicht der Regierung und ermöglicht den Bürgerinnen und Bürgern eine kritische Überwachung des Verwaltungshandelns.

Eine Open Data-Strategie ermutigt zur aktiven Beteiligung der Bürgerinnen und Bürger am Verwaltungshandeln. Durch den Zugang zu offenen Daten können sie diese analysieren, nutzen und neue Anwendungen oder Dienstleistungen entwickeln. Dies fördert die Zusammenarbeit zwischen der Verwaltung und der Öffentlichkeit, was zu besseren Entscheidungen, innovativen Lösungen und einem verbesserten elektronischen Verwaltungshandeln führen kann. Die Open Data-Strategie kann die Effizienz der Verwaltungsprozesse verbessern. Wenn Daten frei verfügbar sind, können sie von verschiedenen Akteuren genutzt werden, um neue Erkenntnisse zu gewinnen, bessere Entscheidungen zu treffen und die Planung zu verbessern. Durch den Zugang zu umfangreichen Datenbeständen können Verwaltungen ihre Ressourcen effizienter einsetzen und ihre Dienstleistungen optimieren.

I. Flächenmäßige Auswirkungen:

Die Open Data-Strategie hat keine flächenmäßigen Auswirkungen.

J. Auswirkungen auf den Haushaltsplan und die Finanzplanung:

a) Auswirkungen auf Einnahmen und Ausgaben:

Ausgaben der Open Data-Strategie sind im Einzelplan 25 im Rahmen der zur Verfügung stehenden Mittel zu finanzieren. Die Umsetzung der Maßnahmen durch die jeweiligen Fachverwaltungen erfolgt im Rahmen der im jeweiligen Einzelplan zur Verfügung stehenden Mittel. Durch die Umsetzung der Open Data-Strategie potentiell entstehende Mehreinnahmen sind derzeit nicht quantifizierbar.

b) Personalwirtschaftliche Auswirkungen:

Die zu erfüllenden Aufgaben im Rahmen der Umsetzung der Maßnahmen der Open Data-Strategie, können im Rahmen bereits bestehender Personalressourcen für die Umsetzung des Gesetzes zur Förderung des E-Government realisiert werden. Nach § 8 der Open Data-Verordnung wurden in den Bezirks- und Senatsverwaltungen größtenteils Open Data-Beauftragte für die Koordinierung und Unterstützung der Maßnahmen zu Open Data sowie für die Förderung der Bereitstellung von offenen Daten innerhalb ihrer Behörden ernannt. Die Tätigkeit der Open Data-Beauftragten spiegelt sich mit einem unterschiedlichen Prozentanteil der Gesamtarbeitszeit in den Anforderungsprofilen der ernannten Beauftragten wieder und kann entsprechend der tatsächlich benötigten zeitlichen Ressource von den Verwaltungen selbst angepasst werden.

Berlin, den 7. November 2023

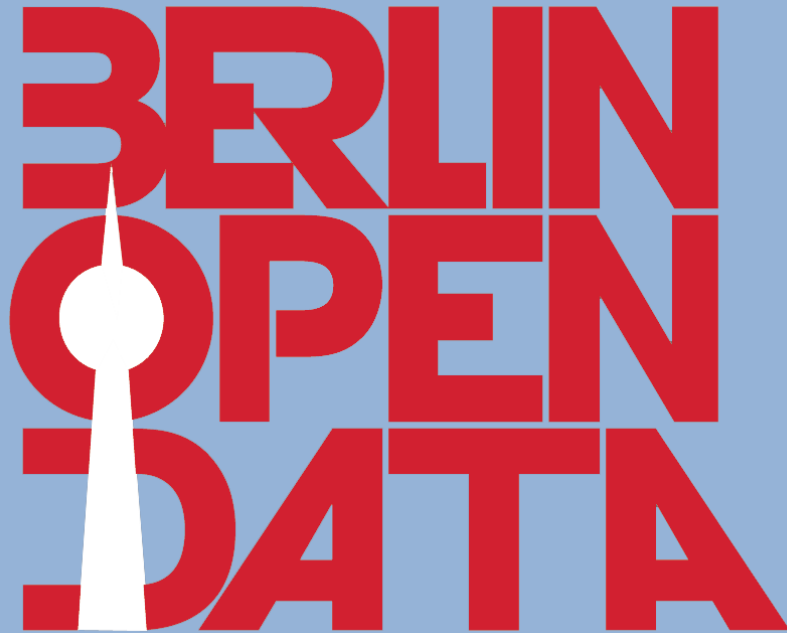
Der Senat von Berlin

Regierender Bürgermeister



OPEN DATA-STRATEGIE

Stand: 25.09.2023



Inhaltsverzeichnis

Inhaltsverzeichnis	2
1.Vorbemerkung	4
2. Mit Open Data als Booster zur Verwaltungsmodernisierung	5
3. Zielstellung der Open Data-Strategie.....	6
3.1. Eine Vision für Open Data als zentraler Baustein der Verwaltungsmodernisierung	6
3.2. Open Data Strategie als Fachstrategie.....	6
3.3. Die Open Data-Strategie als eine lernende und resiliente Strategie	7
4. Der Beteiligungsprozess zur Open Data-Strategie	7
4.1 Erkenntnisse aus den Beteiligungsworkshops	8
4.1.1 Aufbau einer Data Governance	9
4.1.2 Open-by-default zur Effizienzsteigerung	10
4.1.3 Linked Open Data.....	10
4.1.4 Berliner Data Hub	11
5. Kompass zu Open Data Berlin	12
5.1 Wo stehen wir mit Open Data im Land Berlin?	12
5.2 Handlungsfelder und Maßnahmen	14
5.2.1 Data Literacy und Governance	15
5.2.2 Datenkompetenz.....	16
5.2.3 Open Data Mindset stärken	16
5.2.4 Datenbereitstellung verbessern	17
5.2.4.1 Berlins Kerndatensätze auf dem Open Data Portal veröffentlichen	18
5.2.4.2 Dateninventuren in den Bezirks- und Senatsverwaltungen	18
5.2.5 Musterdatenkatalog	19
5.2.6 Qualität & Technische Interoperabilität mit Linked Open Data.....	20
5.2.7 Internes Datenmanagement etablieren.....	22
5.2.7.1 Leitprinzip „Open-by-default“-Strategie	23
5.2.7.2 Open Data in Verträgen für Datensouveränität	24
5.2.7.3 Data Hub mit „Cloud Speicher“/Datenplattform	26
5.2.8 Intelligente Datennutzung	27
5.2.8.1 Open Data-Anwendungen für die Verwaltung	27

5.2.8.2 Open Data-Anwendungen in Kollaboration mit dem Open Data Ökosystem.....	27
6. Ausblick.....	28
7. Anhang.....	28
8. Maßnahmenplanung	29

1.Vorbemerkung

Der Strategiefindungsprozess für die Berliner Open Data-Strategie wurde von der Senatsverwaltung für Wirtschaft, Energie und Betriebe (SenWiEnBe), der damaligen federführenden Senatsverwaltung für das Thema Open Data im Land Berlin, mit einem breiten Partizipationsprozess im letzten Jahr begonnen.

Für die Organisation und Durchführung der Workshops und die anschließende Evaluation der Ergebnisse wurde die Open Knowledge Foundation Deutschland e.V. beauftragt. Die Open Data-Informationsstelle Berlin unterstützte den Partizipationsprozess mit Ihrer Expertise im Bereich Open Data im Land Berlin¹.

Der zweistufige Partizipationsprozess erfolgte vorab mit einer Umfrage auf der Beteiligungsplattform des Landes Berlin <https://mein.berlin.de/>. Die Berliner Öffentlichkeit wurde im Rahmen dieser Umfrage zu der zukünftigen Ausrichtung der Open Data-Strategie und der Schwerpunktsetzung der Workshops befragt. Das Berliner Open Data-Ökosystem, bestehend aus Verwaltung, Wirtschaft, Wissenschaft und Zivilgesellschaft, konnte in den Stakeholder Workshops seine Perspektiven zum Status Quo Open Data in Berlin sowie ihre Anregungen und Impulse im Hinblick auf die Weiterentwicklung der neuen Strategie einbringen. Die sogenannten Schlaglichter aus der Online-Beteiligung und aus den Stakeholder Workshops wurden in dieser Open Data-Strategie berücksichtigt und in konkret umzusetzende Maßnahmen formuliert.

Seit Mai 2023 ist die Senatskanzlei die federführende Senatsverwaltung für das Thema Open Data im Land Berlin und hat die Strategieausarbeitung und die senatsinterne Abstimmung/Befassung weiterverfolgt.

¹ Auf der Website <https://strategie.odis-berlin.de/> wird der Strategieprozess offen und transparent begleitet. Dort ist alles Wissenswertes zur Open Data Landschaft Berlins und rund um den Prozess, die Dokumentationen und Ergebnisse der Workshops und die beteiligten Akteur:innen zu finden.

2. Mit Open Data als Booster zur Verwaltungsmodernisierung

Die Umsetzung von Open Data in Berlin stellt eine wichtige Komponente bei der Entwicklung der Stadt zu einer datengetriebenen, digitalen Metropole dar. Die Verfügbarkeit qualitativ hochwertiger Daten aus Bereichen wie Stadtplanung, Umwelt, Demografie und Finanzen trägt dazu bei, die Wertschöpfung und Innovationskraft der Stadt zu steigern. Der freie Zugang zu diesen Datenbeständen ermöglicht neue Formen der digitalen Teilhabe und Wertschöpfung für Bürgerinnen und Bürger, Unternehmen, Wissenschaftlerinnen und Wissenschaftler. Die Berliner Verwaltung verfügt über eine Vielzahl dieser Daten, die für die Entwicklung der Stadt von großem Nutzen sein können. Durch die Veröffentlichung dieser Daten kann die Transparenz der Politik gesteigert, die Verständlichkeit der Stadt erhöht und neue Dienste für die Bürgerinnen und Bürger entwickelt werden.

In Berlin wurde frühzeitig die Notwendigkeit erkannt, Open Data in die Verwaltung zu integrieren. Bereits 2012 wurde die erste Open Data-Strategie² veröffentlicht, um diese Ziele zu verfolgen. In den vergangenen Jahren konnten die Rahmenbedingungen für Open Data in Berlin sukzessive verbessert werden. So wurde mit der Open Data-Verordnung oder der Benennung behördlicher Open Data-Beauftragten weitere Meilensteine gesetzt. Mit der Open Data Informationsstelle des Landes Berlin verfügt das Land zudem über ein externes Unterstützungsangebot für die Berliner Verwaltung. Rund 10 Jahre nach der ersten Strategie war es an der Zeit, einen Zwischenstand zu dem bisher Erreichten zu ziehen und neue Impulse für die zukünftige Ausrichtung von Open Data im Land zu sammeln. Denn trotz der Erfolge der letzten Jahre ist in den Beteiligungsworkshops zur Strategie deutlich geworden, dass die Umsetzung von Open Data an vielen Stellen hakt. Diese Strategie setzt aufbauend auf den Erkenntnissen des Beteiligungsprozesses auf ein neues Narrativ. Die bisherige Verengung des Open Data Potenzials auf Wirtschaft und Wertschöpfung bzw. der Nutzung von Open Data durch Dritte wird erweitert um den Bedarf und Nutzen an Open Data für die Verwaltung selbst. Im Zuge der steigenden Anforderungen an die Verwaltung und dem steigenden Anspruch an die Verwaltung als eine digitale, moderne und datenaffine Verwaltung kann Open Data ein entscheidender Hebel sein. Gelingt es, hochwertiges Open Data in der Verwaltung in der Fläche zu etablieren, können Prozesse verbessert, Verwaltungshandeln effizienter ablaufen, Datenhoheiten geklärt, Ressourcen gespart und die Zusammenarbeit innerhalb der Verwaltung kollaborativer gestaltet werden. Der Clou dabei: Nutzt die Verwaltung ihre (offenen) Daten selbst, steigt damit die Qualität der Daten und damit das Nutzungspotenzial für Dritte von selbst.

Die hier vorliegende Open Data-Strategie fasst die Ergebnisse des Strategie-Beteiligungsprozesses zusammen und leitet aus den wesentlichen Erkenntnissen konkrete Handlungsfelder und Maßnahmen ab, die auf die überarbeitete Zielstellung für Open Data in Berlin einzahlen.

² Both/Schieferdecker

3. Zielstellung der Open Data-Strategie

3.1. Eine Vision für Open Data als zentraler Baustein der Verwaltungsmodernisierung

Mit dieser Strategie wird erstmals im Land Berlin das Ziel formuliert und verankert, Open Data als wesentlichen Bestandteil einer modernen digitalen Verwaltung zu begreifen und entsprechend Maßnahmen zu ergreifen. Die Open Data-Strategie verfolgt damit die Vision eines damit verbundenen Verwaltungskulturwandels für eine verbesserte interne Integration von Open Data in die Verwaltungsprozesse, um insbesondere die eigene Verwaltungsarbeit mit Open Data zu erleichtern und zu verbessern. Mit dem Kulturwandel ist auch die Notwendigkeit verbunden, bestehende Prozesse zu hinterfragen und Open Data in der Praxis (zum Beispiel bei der Datensammlung oder der Vergabe an externe Dienstleisterinnen und Dienstleister) von Anfang an mitzudenken. Der Zuständigkeitswechsel von Open Data in der Berliner Verwaltung zur Senatskanzlei ist damit gleichzeitig eine große Chance Open Data stärker in der IT-Ausrichtung des Landes tief zu verankern und IT-Prozesse in Richtung Open Data zu entwickeln.

Dabei hat die Strategie den Anspruch, mit konkreten Maßnahmen die Ziele für Open Data in den nächsten fünf Jahren zu erreichen. Statt um die theoretische Formulierung von Zielen, geht es um die konkreten Umsetzungsschritte mit Zeitplänen, in denen diese Maßnahmen umgesetzt werden sollen.

3.2. Open Data Strategie als Fachstrategie

Im Land Berlin wurde die Strategie „Gemeinsam Digital:Berlin als sogenannte Dachstrategie aller Digitalisierungsthemen der Berliner Verwaltung verabschiedet, die in ihren Handlungsfeldern auf unterschiedliche Fachstrategien des Landes Berlins für die konkrete Umsetzung mit Projekten verweist. Die neue Open Data-Strategie wird folglich als eine Fachstrategie der Gemeinsam Digital:Berlin-Strategie verstanden. So sollen die wesentlichen Handlungsfelder mit Bezug auf die Themenbereiche Datenkompetenz, Datennutzung, Data Governance, internes Datenmanagement/öffentliche Infrastruktur³ das Prinzip „Open-by-default“⁴ und die dafür notwendige technische Infrastruktur in der Behördenlandschaft innerhalb der Strategie „Gemeinsam Digital:Berlin“ in der Open Data-Strategie mit Maßnahmen konkretisiert werden. Weiterhin wurde darauf geachtet, dass mit

³ Gemeinsam Digital: Berlin Strategie, Seite 13 „Handlungsfeld: Offenes IT-Ökosystem städtischer Anwendungen entwickeln“

⁴ Gemeinsam Digital: Berlin Strategie, Seite 14 „Handlungsfeld: Offenes IT-Ökosystem städtischer Anwendungen entwickeln“

der neuen Open Data-Strategie ebenfalls die Ziele der Bundesregierung für Open Data⁵ und der Open Data-Verordnung auf Landesebene mitverfolgt werden.

3.3. Die Open Data-Strategie als eine lernende und resiliente Strategie

Wie die Strategie „Gemeinsam Digital:Berlin“ versteht sich die Open Data-Strategie als lernende Strategie. Die Open Data-Strategie beansprucht damit für sich, nicht von Beginn an alle Maßnahmen im Bereich Open Data abzudecken. Vielmehr ist sie offen für Entwicklungen und notwendige Anpassungen, die in den nächsten Jahren zum Beispiel aufgrund sich ändernder Rahmenbedingungen für Open Data auf EU- oder Bundesebene auch auf Berlin auswirken können.

Nicht nur die einzelnen Maßnahmen, sondern auch die Strategie selbst wird in regelmäßigen Zyklen evaluiert und gegebenenfalls angepasst. Dafür sollen mindestens einmal im Jahr ab der Veröffentlichung der Strategie die im Beteiligungsprozess involvierten Akteursgruppen mit einbezogen werden.

4. Der Beteiligungsprozess zur Open Data-Strategie

Die Open Data-Strategie aus 2011 vom Fraunhofer FOKUS⁶ und die bereits umgesetzten Open Data-Projekte im Land Berlin, insbesondere für die Umsetzung der Open Data-Verordnung, waren eine optimale Vorlage und ein guter Ausgangspunkt für den Open Data-Partizipationsprozess, der in den Monaten März, April und Mai im letzten Jahr mit den Stakeholder Gruppen Wirtschaft, Wissenschaft, Verwaltung und Zivilgesellschaft durchgeführt wurde. Für die Beteiligung des gesamten Open Data-Ökosystems wurde in einer ersten Phase eine Online-Umfrage auf mein.berlin.de⁷ durchgeführt. In der zweiten Phase fanden Stakeholder-Workshops gemeinsam mit der Open Knowledge Foundation e.V. (OKFN e.V.) und unterstützt durch die Open Data Informationsstelle Berlin statt.

Während des gesamten Prozesses wurden die Zwischenergebnisse auf mein.berlin.de zur Diskussion und Online-Komentierung veröffentlicht. Auch die Endergebnisse der Online-Bürgerinnen- und -Bürger-Beteiligung wurden auf der Informationsseite⁸ für die breite Stadtgesellschaft zum Nachverfolgen veröffentlicht. Die Handlungsfelder und Maßnahmen wurden aus den Erkenntnissen des Beteiligungsprozesses entwickelt und für die Umsetzung in der laufenden Legislaturperiode in der vorliegenden Strategie mit Maßnahmen konkretisiert.

⁵ <https://www.bundesregierung.de/re-source/blob/976074/1940386/1d269a2ad1b6346fcf60663bdea9c9f8/2021-07-07-open-data-strategie-data.pdf?download=1>

⁶ Fraunhofer FOKUS https://www.berlin.de/sen/wirtschaft/digitalisierung/assets/Open_Data-strategie.pdf

⁷ https://mein.berlin.de/projekte/Open_Data-strategie-berlin/

⁸ https://odis-berlin.de/assets/file-download/OD_Strategie_online_umfrage.pdf

4.1 Erkenntnisse aus den Beteiligungsworkshops

Im Folgenden werden die Kernergebnisse des Beteiligungsprozesses auf mein.berlin.de und den Stakeholder-Workshops zusammengefasst.

Schlaglichter der Online-Bürgerinnen- und -Bürger-Beteiligung

Durch die Online-Bürgerinnen und -Bürger-Beteiligung auf mein.berlin.de⁹ wurden folgende Themen identifiziert, die in den Beteiligungsworkshops mit Schwerpunktsetzung diskutiert werden sollen:

- personelle/finanzielle/technische Ressourcen,
- Kulturwandel,
- Know-How,
- politischer Wille,
- Open Data-Schnittstellen in Fachverfahren,
- Datenmanagement in den Verwaltungen,
- automatisierte Datenbereitstellung und
- das Verständnis, dass nur die Gesamtheit der veröffentlichten Daten einen Mehrwert bringen.

Die Auswertung verdeutlicht, dass Open Data mehr als nur die automatisierte Bereitstellung der Daten über Open Data-Schnittstellen bedeutet. Open Data setzt ebenfalls einen Kulturwandel in der Berliner Verwaltung voraus. Dafür muss der Umgang mit Daten erlernt werden. Der politische Wille für Open Data muss vorhanden sein, so dass auch die Führung der Abteilungen und Referate in ihrer Arbeit an die Umsetzung der Open Data-Verordnung stets von Beginn an mitdenken.

Die aktuelle Datenqualität wurde in der Online Umfrage wie folgt bewertet:

- „Daten auf dem Open Data Portal sind in der Regel“
- so veröffentlicht, dass sie einfach zugänglich für mich sind: 20%
 - vollständig und qualitativ hochwertig: 15%
 - in einem gut nutzbaren Format veröffentlicht: 20%
 - automatisiert (z.B. über Schnittstellen) verlässlich abrufbar: 10%
 - Aufbereitung von Open Data für die Nachnutzung ist sehr aufwendig: 55%

Die Auswertung zeigt, dass die Datenqualität gesteigert werden muss, damit die Nachnutzung der Daten besser gewährleistet wird. Es braucht ebenfalls mehr verlässlich abrufbare Schnittstellen für automatisierte Veröffentlichungen aus den bestehenden Systemen und Fachverfahren.

⁹ https://odis-berlin.de/assets/file-download/OD_Strategie_online_umfrage.pdf

In den vier Stakeholder-Workshops wurden Wardley-Maps¹⁰ von den Teilnehmerinnen und Teilnehmern erstellt, die mehrheitlich folgende Entwicklungsmöglichkeiten für Open Data in Berlin zum Ausdruck brachten.

Schlaglichter der Stakeholder- Workshops:

1. Open Data braucht eine neue Erzählung, die sich auf die interne Nutzung in der Verwaltung und die damit verbundene Nützlichkeit für die einzelnen Verwaltungsbeschäftigten fokussiert.
2. Teil dieser Erzählung muss die Verknüpfbarkeit von Datensätzen über viele Verwaltungsquellen und darüber hinaus sein, um verteilte Informationen zu einem gemeinsamen, abfragbaren Wissensschatz zusammentragen zu können.
3. Der geplante Data Hub als zentrale Datenplattform ist wichtig, braucht klare Vorgaben und sollte möglichst modular in seinen Funktionalitäten wachsen, während gleichzeitig die Grundlagen für Linked Open Data als Standard gelegt werden müssen.¹¹

Aus diesen drei Erkenntnissen ließen sich folgende konkrete Forderungen der Workshop-Teilnehmenden für die Maßnahmen der Open Data-Strategie ableiten:

4.1.1 Aufbau einer Data Governance

Über die zehn Jahre Open Data Praxis wurde den Stakeholdern deutlich, dass die Berliner Verwaltung eine Data Governance benötigt, um die Zuständigkeiten und die notwendigen Prozesse für die Öffnung der Verwaltungsdaten transparent zu machen. Konkret sind Zuweisungen von Zuständigkeit und Verantwortlichkeit zwischen allen Ebenen der behördlichen Verwaltungsstruktur, also in den Bezirks- und in den Senatsverwaltungen, für die Öffnung von Daten wichtig. Auch die Erkundung vorhandener Datenbestände in der Verwaltung kann hierfür sinnvoll sein. In den Workshops wurde vielfach berichtet, dass der Informationsfluss besonders dann stockt, wenn Personen auf Daten von anderen Stellen in der Berliner Verwaltung angewiesen sind. Dort zeigt sich immer wieder, dass angefragte Daten nicht strukturiert, sondern als Aktenvermerke vorliegen.

Die Verwaltungsmodernisierung darf nicht darauf beschränkt sein, analog vorliegende Informationen und Prozesse in eine elektronische Form zu überführen. Vielmehr muss es auch darum gehen, ein umfassendes internes Datenmanagement in der Verwaltung aufzubauen, das es ermöglicht, Prozesse neu zu definieren und Aufgaben effizienter zu erledigen.

Dabei sollte auch die Verwaltungsarbeit mit strukturierten Datensätzen im Vordergrund stehen. Diese Daten sollten auch anderen Verwaltungen zur Verfügung gestellt oder die Verknüpfbarkeit dieser Daten mit anderen Daten sichergestellt werden können. Wenn in der Verwaltung mit Daten gearbeitet wird, müssen diese von Anfang an für eine mögliche Veröffentlichung vorbereitet werden („Open-by-default“).

¹⁰ <https://learnwardleymapping.com/> Eine Wardley-Karte ist eine Karte für die Geschäftsstrategie

¹¹ Erkenntnisse aus dem Beteiligungsprozess zur Erarbeitung einer neuen Open Data-Strategie Berlin, OKFN e.V.

Die Datenqualität wird steigen, wenn die Berliner Verwaltung die Arbeit mit den Daten in ihre vorhandenen Prozesse integriert. In diesem Fall ist auch die Nutzung der Daten für Dritte - wie z.B. für die Wissenschaft, Wirtschaft und Zivilgesellschaft - besser gewährleistet, wenn die Verwaltung die Daten so qualitativ gut bearbeitet, dass sie auch selber damit arbeiten kann.

4.1.2 Open-by-default zur Effizienzsteigerung

Obwohl auf dem Berliner Open Data-Portal immer mehr Datensätze aus der Berliner Verwaltung veröffentlicht werden und Fachverfahren an das Open Data-Portal angeschlossen werden, setzt sich das Prinzip Open-by-default in den Berliner Behörden nur langsam durch. Dabei ist das Prinzip Open-by-default fest in der Berliner Open Data-Verordnung verankert, die am 1.1.2021 in Kraft getreten ist.

Damit Open-by-default sich durchsetzt, muss die Bereitstellung von Open Data in der Weise in den eigenen Verwaltungsalltag integriert werden, dass sie *nicht* dauerhaft zu einem Mehraufwand für die Verwaltungsbeschäftigten führt. Das Ziel muss es sein, durch API-Schnittstellen die Prozesse so zu automatisieren, dass man sie nicht wieder anfassen muss. Open Data muss durch das Prinzip Open-by-default so empfunden werden, dass es die eigene Verwaltungsarbeit effizienter und einfacher macht. Um dieses Ziel zu erreichen, müssen die über 300 Fachverfahren geöffnet und mit API-Schnittstellen an das Open Data-Portal angebunden werden können.

Rahmenverträge mit Dienstleisterinnen und Dienstleistern, die IT-Lösungen für Fachverfahren anbieten, müssen daraufhin überprüft werden, ob sie die Bereitstellung von Open Data in gebotener Qualität und Granularität ermöglichen (Open-by-Design) und ob sie die notwendigen Voraussetzungen für die nächste Qualitätsstufe von Open Data schaffen: „Linked Open Data“.

4.1.3 Linked Open Data

Open Data ist nicht gleich Open Data. Bislang wurde der Erfolg von Open Data meist an der Anzahl bereitgestellter Datensätze auf dem Open Data-Portal <https://daten.berlin.de> gemessen. Ob die veröffentlichten Datensätze PDF-Dokumente, veraltete Excel-Listen oder CSV-Dateien sind, wurde kaum unterschieden. Der wirkliche Erfolg von Open Data hängt aber maßgeblich auch vom Reifegrad der Offenheit ab.

Ein Maßstab für die Offenheit von Daten ist die sog. 5-Sterne-Open Data-Einteilung. In dieser Einteilung ist der fünfte und höchste Grad der Offenheit nur dann erreichbar, wenn die veröffentlichten Datensätze miteinander „sprechen lernen“: durch Linked Open Data¹².

¹² Vgl. Definition Linked Open Data bei Wikipedia: https://de.wikipedia.org/wiki/Linked_Open_Data. Tim Berners-Lee, Erfinder des World Wide Web, stellte 2009 in einem TED-Talk seine Zukunftsvision für ein künftiges WWW der Daten vor. Darauf aufbauend wurde das Modell der 5 Sterne für Open Data entwickelt: In den ersten Stufen sind Daten überhaupt im Web verfügbar, weitere Sterne gibt es für Daten in strukturierter Form (Tabellen). Die eigentlich visionäre Leistung Berners-Lees liegt jedoch in den Stufen 4 und 5 dieses Modells. Vier Sterne bedeutet, dass nicht nur eine bloße Tabelle vorliegt, sondern dass ein Datensatz sich selbst beschreibt: „Dieser Datenpunkt bildet Datum und Uhrzeit ab, hierbefinden sich Adressen in einem

Für Linked Open Data müssen die Datensätze so angelegt werden, dass sie möglichst standardisiert, gut beschrieben und die einzelnen Datenpunkte eindeutig identifizierbar sind. Nur so können Verknüpfungen automatisch generiert werden. Dieses Linked-Open Data-System, das vollständig beschriebene Daten enthält, erlaubt die Zusammenfassung und Verknüpfung von Datensätzen, die aus unterschiedlichen Quellen stammen.

4.1.4 Berliner Data Hub

Das Nutzungsversprechen von Open Data für die Berliner Verwaltung kann nur dann eingelöst werden, wenn in der Zukunft mit einem modernen Datenmanagementsystem gearbeitet wird.

Der Berliner Data Hub soll diese Funktion eines zentralen Datenmanagementsystems erfüllen. Der Data Hub wird nur dann gut funktionieren, wenn das gesamte Daten-Ökosystem, das sowohl aus der IT-Infrastruktur, den bereits etablierten Workflows, und neuen Zuständigkeiten mit personellen Ressourcen, gut aufgestellt ist.

Der Data Hub sollte von Anfang an darauf ausgerichtet sein, auch verknüpfte Datensätze bereitzustellen - der Aufbau von Speicher- und Verbreitungssystemen für Linked Open Data sollte in sogenannten Triple Stores¹³ strategisch im Vordergrund stehen und in die IKT-Architektur aufgenommen werden.

Der geplante Data Hub soll kein monolithisches System sein, das alle erdenklichen Bedürfnisse und Anforderungen von Anfang an erfüllen kann. Vielmehr muss es sich um einen Verbund von Diensten und Systemen handeln, der verschiedene Aufgaben übernimmt. Weitere Funktionalitäten und Anforderungen können in einem modularen System nach und nach aufgebaut werden. So wächst der Data Hub in seinen Funktionalitäten und wird immer weiter verbessert. Folgende Aufgaben wurden von den Stakeholdern als wichtige Anforderungen gesehen, die in einem Data Hub aufgenommen werden sollen:

- Auffindbarkeit und Katalogisierung von Daten
- Triple Store für Linked Data
- Speicherort für nicht-Linked Data
- Zeitreihendatenbanken z.B. für Sensorikdaten
- Extraktions-, Transformations- und Ladeverfahren (ETL) für den automatisierten Import aus Datenbanken von Fachverfahren¹⁴

definierten Format.“ In der 5-Sterne-Ausbaustufe sind verschiedene Datensätze dann miteinander in Bezug gebracht – verlinkt – worden.

¹³ Ein Triple Store organisiert Informationen in einer Semantik mit den drei Bestandteilen „Subjekt – Prädikat – Objekt“ – deswegen Tripel. Als Beispiel kann das Wikidata-Datenobjekt für das Land Berlin herangezogen werden: „Berlin – ist Instanz von – Bundesland (Deutschland)“. So können beispielsweise Abfragen in der Form „zeige alle Bundesländer Deutschlands mit der zugehörigen Bevölkerungszahl in aufsteigender Reihenfolge an“ getätigt werden.

¹⁴ Erkenntnisse aus dem Beteiligungsprozess für Open Data, OKFN e.V.

5. Kompass zu Open Data Berlin

Mit dem Strategiebeteiligungsprozess ist sowohl eine umfassende Evaluation als auch eine Aufnahme der wichtigsten Bedarfe und Hinweise für die zukünftige Ausrichtung von Open Data in Berlin erfolgreich abgeschlossen. Es gilt nun, auf Basis dieser Ergebnisse und im weiteren Austausch mit alten und neuen Akteuren für Open Data in Berlin neue Impulse zu setzen. Dazu bedarf es von zentraler Stelle im Land konkrete Maßnahmen, um Open Data strategisch, politisch, personell und prozessual zu flankieren. Mit dem Zuständigkeitswechsel von Open Data in der Berliner Verwaltung von der SenWiEnBe zur Senatskanzlei ist die Grundlage dafür geschaffen, an zentraler Stelle unter dem/der Chief Digital Officer (CDO) des Landes das Thema Open Data gemeinsam mit der zentralen Stelle für Open Data in Berlin als Grundlage für die Verwaltungsmodernisierung auszurichten. Dazu gehören die Formulierung und die Konkretisierung von messbaren Maßnahmen für die nächsten Jahre.

In diesem Kapitel werden erste Maßnahmen formuliert, die auf das Ziel der Strategie einzahlen, Open Data in die flächendeckende Umsetzung zu bringen. Dabei kann das Land Berlin auf die Erfolge der letzten Jahre, wie zum Beispiel die rechtlichen Rahmenbedingungen durch die Open Data-Verordnung oder den Aufbau der Open Data-Informationsstelle aufbauen.

5.1 Wo stehen wir mit Open Data im Land Berlin?

Open Data ist kein neues Thema in der Berliner Verwaltung. Schon im Jahr 2011 hat das Land Berlin die Wichtigkeit der Formulierung konkreter Ziele und Maßnahmen für die technischen, rechtlichen und organisatorischen Rahmenbedingungen zur Umsetzung von Open Data in der Berliner Verwaltung erkannt und gemeinsam mit Fraunhofer FOKUS die erste Open Data-Strategie des Landes Berlin herausgebracht. Basierend auf der Open Data-Strategie 2011 von Fraunhofer FOKUS¹⁵ wurden aktuell folgende Ziele erreicht, die eine gute Basis für die neue Open Data-Strategie darstellen:

Technische Rahmenbedingungen:

Die erste Phase für Open Data Berlin begann 2011/2012 mit der Formulierung einer Open Data-Strategie und dem Testbetrieb des Open Data-Portals. Das Portal befindet sich seitdem beim landeseigenen Dienstleister BerlinOnline im Regelbetrieb und ist beständig weiterentwickelt worden, wobei die Anzahl an verfügbaren Datensätzen kontinuierlich gewachsen ist.

¹⁵ <http://publica.fraunhofer.de/documents/N-195507.html>

Rechtliche Rahmenbedingungen:

Mit der Verabschiedung des Berliner E-Government-Gesetzes 2016¹⁶ und insbesondere mit dem §13 EGovG zur Bereitstellung allgemein zugänglicher Datenbestände wurde eine gesetzliche Grundlage für den Aufbau einer offenen behördlichen Dateninfrastruktur geschaffen.¹⁷ Die dazugehörige Open Data-Verordnung¹⁸ ist am 1.1.2021 in Kraft getreten.

Organisatorische Rahmenbedingungen:

In den meisten Bezirksämtern und Senatsverwaltungen wurden Open Data-Beauftragte ernannt, die gemäß § 8 der Open Data-Verordnung für die Koordinierung und Unterstützung der Maßnahmen zu Open Data sowie Förderung der Bereitstellung von offenen Daten innerhalb ihrer Behörden zuständig sind. Eine Umfrage unter den behördlichen Open Data-Beauftragten ergab, dass sich die Tätigkeit mit einem unterschiedlichen Prozentanteil der Gesamtarbeitszeit in den Anforderungsprofilen der ernannten Beauftragten widerspiegelt. Die Umfrage bestätigte ebenfalls, dass die tatsächliche zeitliche Ressource für die Einarbeitung und Umsetzung von Open Data in den eigenen Behörden kaum ausreicht. Im Grundsatz ist das Implementieren der behördlichen Open Data-Beauftragten ein großer Erfolg und ein wichtiger Schritt, allerdings muss auch ausreichend Arbeitszeit für die Erfüllung der Aufgabe eingeplant und zur Verfügung gestellt werden.

Open Data-Informationsstelle Berlin:

Mit der Open Data-Informationsstelle (ODIS), angesiedelt bei der Technologiestiftung Berlin, steht den Berliner Behörden und ihren Mitarbeiterinnen und Mitarbeitern seit 2018 ein Unterstützungsangebot rund um alle Fragen zum Thema Open Data zur Verfügung. Das Angebot der ODIS umfasst u.a. Beratungsangebote zur Datenaufbereitung und -bereitstellung, Unterstützung bei Dateninventuren, Kompetenzaufbau und Wissenstransfer sowie die Entwicklung von prototypischen Anwendungen und Tools. Insbesondere Leuchtturmprojekte der ODIS mit den Berliner Behörden zeigen als Best-Practice-Beispiele das Potenzial offener Daten und helfen das Thema Open Data im Open Data-Ökosystem zu etablieren.

Kompetenzaufbau Open Data

Als Schulung mit Zertifikat wird der eintägige Crashkurs Open Data des Landes Berlin angeboten, um sowohl die Open Data-Beauftragten als auch die Datenkoordinatorinnen und Datenkoordinatoren im Umgang mit Daten und der Datenveröffentlichung zu unterrichten.

¹⁹ Über die bei Open Data federführende Senatskanzlei werden das Open Data-Handbuch (als „lebendes Dokument“ von Berlin Online verfasst und kontinuierlich erneuert) sowie Best

¹⁶ <https://gesetze.berlin.de/bsbe/document/jlr-EGovGBErahmen>

¹⁷ Vgl. Gesetz zur Förderung des E-Government vom 30. Mai 2016.

¹⁸ <https://www.berlin.de/sen/wirtschaft/digitalisierung/assets/rvo.pdf>

¹⁹ Die Anmeldung läuft zentral über die Webseite der VAK für alle Verwaltungsbeschäftigten:

<https://www.berlin.de/vak/>

Practices und Publikationen für den Start als Open Data-Beauftragte/r zur Verfügung gestellt.²⁰ Auch auf der Webseite der ODIS gibt es ein stetig wachsendes Angebot an Unterstützungsmaterialien und Selbstlernformaten für Verwaltungsmitarbeiterinnen und Verwaltungsmitarbeiter zu Open Data relevanten Themen.²¹

Vernetzungsaktivitäten:

Der Austausch innerhalb der Berliner Behörden, als auch mit den anderen Ländern und dem Bund wird weitergeführt. Ebenfalls ist es wichtig, den Austausch herzustellen zwischen den Verwaltungsbeschäftigten, die mit Daten arbeiten, und den Open Data-Beauftragten, die Daten veröffentlichen. Für den regelmäßigen Austausch werden seit über zehn Jahren Vernetzungsveranstaltungen wie der Berlin Open Data Day, die AG Open Data oder die Open Data Lunches durchgeführt.

Datenangebot im Land Berlin:

Zu Anfang der Legislaturperiode Ende 2016/Anfang 2017 umfasste das Open Data Portal rund 1100 veröffentlichte Datensätze von 56 Veröffentlichenden. Stand Dezember 2022 sind es knapp 3200 Datensätze von mehr als 150 Veröffentlichenden. Damit sind in der vergangenen Legislaturperiode rund 2000 Datensätze veröffentlicht worden. Davon entfallen auf das Jahr 2022 rund 430 Datensätze. Hervorzuheben, ist das umfangreiche Datenangebot der Senatsverwaltung für Stadtentwicklung, Bauen und Wohnen im Geoportal des Landes Berlin, das über eine Schnittstelle an das Open Data Portal angebunden ist und somit rund 70% des Open Data Angebots ausmacht.

5.2 Handlungsfelder und Maßnahmen

In diesem Kapitel werden die in den Stakeholder-Workshops genannten Anforderungen und Bedarfe an Open Data in Berlin in konkrete Maßnahmen überführt. Ebenfalls werden die Handlungsempfehlungen der „Gemeinsam Digital:Berlin“-Strategie aus den Themenbereichen internes Datenmanagement, Open-by-default, Datenkompetenz und Data Governance in den Handlungsfeldern dieser Strategie aufgegriffen.

In den Handlungsfeldern wurden sowohl die Aspekte der technischen Weiterentwicklung des Open Data Portals, als auch die Erhöhung der Qualität der Datenveröffentlichung, wie auch die Datennutzung und die internen Prozesse und Zuständigkeiten innerhalb der Berliner Verwaltung für Open Data als Dimensionen für das Erreichen der nächsten Phase für Open Data Berlin in Betracht gezogen.

20 Diese können direkt über die E-Mail Adresse: opendata@seninnds.berlin.de angefragt werden.

21 <https://odis-berlin.de/ressourcen/>

Die Senatskanzlei ist nunmehr die federführende Senatsverwaltung für Open Data im Land Berlin und wird gemeinsam mit der ODIS und den Open Data-Beauftragten der Bezirks- und Senatsverwaltungen die Maßnahmen umsetzen. Die Umsetzung der Maßnahmen erfolgt im Rahmen der im jeweiligen Haushalt zur Verfügung stehenden Mittel.

5.2.1 Data Literacy und Governance ²²

Für die intrabehördliche Datenarbeit innerhalb der Berliner Verwaltungslandschaft ist die Kenntnis wichtig, welche Daten in welcher Behörde vorliegen und von wem sie bearbeitet werden. Data Governance betont stärker die organisatorischen Aspekte wie zum Beispiel die Zuständigkeiten, Rollen, Abstimmungsprozesse des Datenmanagements in den Behörden.

Im Umgang mit Daten und Datenarten in einer Organisation können verschiedene Rollen im internen Datenmanagement identifiziert werden. Dazu gehören Dateneigentümerinnen und Dateneigentümer, Datenbereitstellende und Datennutzende. Beim Blick auf eine bestehende Organisation und die Rollen, die sich mit (Stamm-)Daten befassen, können unterschiedliche Tätigkeits- und Aufgabengebiete abgegrenzt werden.

Das Land Berlin verfügt derzeit noch über keine übergreifende Strategie, die relevante Daten identifiziert, eine entsprechende IT-Architektur für das Datenmanagement definiert und Governancefragen, wie z.B. „Wer darf wann welche Daten nutzen?“, „Wer muss wann welche Daten mit wem teilen?“, „Wie garantieren wir Datensouveränität?“, klärt. Noch viel zu oft werden Daten nicht konsistent erhoben, sondern in „Daten-Silos“ gefangen. Es gibt keine übergreifende Dateninfrastruktur und keine allgemein gültigen Standards, um den Austausch zwischen bestehenden „Insellösungen“ in den einzelnen Senatsverwaltungen und städtischen Unternehmen zu ermöglichen sowie die Auswertung für eine ressortübergreifende Steuerung zu nutzen. In diesem Zusammenhang werden Aktivitäten zur Umsetzung der Open Data Strategie am tatsächlichen Bedarf der Datennutzenden ausgerichtet und dieser Bedarf wird insbesondere bei den Verwaltungen erhoben, die Daten bei anderen Verwaltungen nachfragen. Die ODIS unterstützt diese Bedarfsanalyse und hilft bei der Realisierung einer Nachnutzung benötigter Daten. Darüber hinaus soll im Rahmen einer Studie, die in Auftrag gegeben wird, ein Data-Governance-Modell als Vorschlag zur Umsetzung für die Berliner Behördenlandschaft, die aus Bezirks- und Senatsverwaltungen und den ihnen nachgeordneten Behörden besteht, erarbeitet werden.

Es sollen unterschiedliche Rollen in Zusammenhang mit der Arbeit mit Daten sowohl innerhalb der Behörde als auch zwischen den Behörden aufgezeigt werden, die mit ihren Aufgaben in einer Hierarchie dargestellt werden. Dabei sollen die klaren personellen Verantwortlichkeiten als auch die dafür notwendigen zeitlichen und finanziellen Ressourcen in Betracht gezogen werden.

²² <https://gemeinsamdigital.berlin.de/en/grünbuch/handlungsebene-3-sektorübergreifende-strategische-fragen/daten-governance/>

Federführung: Senatskanzlei

Beteiligte: Auftragsvergabe an Dienstleister oder Dienstleisterin für die Studie

Zeitraum: ab 2024 bis 2026

5.2.2 Datenkompetenz

Auch die Datenkompetenzen, also die Fähigkeiten, der für Daten-verantwortlichen Personen, die Daten auf kritische Art und Weise zu sammeln, zu managen sowie im richtigen Kontext zu bewerten und anzuwenden, müssen stets mit Weiterbildungsmaßnahmen ausgebaut werden. Denn es hat sich in den letzten Jahren gezeigt, dass zwar die Quantität der Daten kontinuierlich zunimmt, die Qualität der Datensätze aber häufig auf der Strecke bleibt. Dies zeigt sich auch an den vielen Datensätzen, die lediglich als PDF- oder Excel-Dateien bereitgestellt werden. Wir wollen das Wissen im Umgang mit Daten und auch im Sinne von Open Data unterstützen. Dabei muss das Datenverständnis in der Breite der Verwaltungslandschaft vergrößert werden, so dass die Verantwortung für das Datenmanagement von vielen Abteilungen und Referaten getragen wird.

- Ein vom Land Berlin zertifizierter Kurs für Datenmanagement und Open Data wird den datenverantwortlichen Verwaltungsbeschäftigten wie z.B. den Fachverfahrensverantwortlichen, Datenkoordinatorinnen und Datenkoordinatoren, Open Data-Beauftragten und den datenhaltenden Fachreferentinnen und Fachreferenten Fachreferenten in den Senats- und Bezirksverwaltungen sowie den nachgeordneten Behörden und den Berliner Einrichtungen angeboten.

Federführung: Senatskanzlei

Beteiligte: ODIS, weitere Referentinnen und Referenten und externe Seminarleiterinnen und Seminarleiter zur Vermittlung von Datenkompetenz, VAK Berlin

Zeitraum: ab 2023 fortlaufend

- Die ODIS erweitert ihr Angebot an Unterstützungsmaterialien für die Verwaltungsmitarbeiterinnen und Verwaltungsmitarbeiter mit je nach Zielgruppe und Thema geeigneten Formaten. Das können Checklisten, Handouts, Videoformate oder Schulungsformate sein.

Federführung: Senatskanzlei

Beteiligte: ODIS

Zeitraum: ab 2023 fortlaufend

5.2.3 Open Data Mindset stärken

Schließlich ist Open Data ein Kulturwandel in der Verwaltung vom Aktengeheimnis zu mehr Transparenz. Die Identifizierung, Bereitstellung und Nutzung von Open Data stellt für Teile der Verwaltung noch eine organisatorische und vor allem eine kulturelle Hürde dar. Bislang wird Open Data von einer „Koalition der Willigen“ in der Berliner Verwaltung vorangetrieben. Damit der Open Data-

Gedanke und die damit zusammenhängenden Vorteile von Open Data sich nachhaltig in der Verwaltungsarbeit verankern, gilt es Überzeugungsarbeit zu leisten und die Vorteile von Open Data herauszustellen. Neben den Maßnahmen unter Handlungsfeld „Intelligente Datennutzung fördern“ sind folgende Maßnahmen vorgesehen:

- Um die Verwaltungsbeschäftigten auf Ihre neuen Herausforderungen im Umgang mit den technischen Systemen vorzubereiten, soll das Open Data-Handbuch²³ erweitert werden.

Federführung: Senatskanzlei

Beteiligte: Berlin Online

Zeitraum: seit 2020 fortlaufend

- Die regelmäßigen AG Open Data-Sitzungen werden fortgeführt und bilden das zentrale Vernetzungsformat der behördlichen Open Data-Beauftragten für den interbehördlichen Erfahrungsaustausch. Die federführende Senatsverwaltung für Open Data, Senatskanzlei wird das Format nutzen, um von den Teilnehmerinnen und Teilnehmern Erfahrungswerte und Herausforderungen zur Umsetzung von Open Data in den einzelnen Senats- und Bezirksverwaltungen in der Praxis einzuholen und in die strategische Umsetzung einfließen zu lassen.

Federführung: Senatskanzlei

Beteiligte: ODIS zur Unterstützung der Arbeitssitzungen

Zeitraum: seit 2020 fortlaufend

- Open Data Berlin, angesiedelt bei der Senatskanzlei wird verstärkt in bestehende Verwaltungsrunden und Gremien der Verwaltungsmodernisierung reingehen, um für Open Data und die Anpassung bestehender Prozesse und Praktiken an den Open Data-Grundgedanken zu werben und die Erkenntnisse des Open Data-Strategieprozesses vorzustellen.

Federführung: Senatskanzlei

Beteiligte: ODIS als externe Referentin

Zeitraum: ab 2023 fortlaufend

5.2.4 Datenbereitstellung verbessern

Wir wollen die Zahl der offenen Datensätze der Berliner Verwaltung deutlich steigern und orientieren uns dabei an den „Kerndatensätzen“,²⁴ deren Veröffentlichung einen Mehrwert für die Stadtge-

²³ <https://github.com/berlinonline/open-data-handbuch>

²⁴ <https://odis-berlin.de/projekte/kerndatensaetze/>

sellschaft darstellen. Dabei ist nicht allein die Quantität, sondern ebenso die Qualität der bereitgestellten Daten wichtig. Denn nur qualitativ hochwertige Daten können optimal durch Dritte genutzt werden. Das Open Data Angebot wollen wir nicht nur aus Anbietersicht erweitern, sondern auch im Rahmen der Identifikation der „Kerndatensätze“ einen konkreten an die Verwaltung als Anbieterin herangetragenen Bedarf ermitteln.

5.2.4.1 Berlins Kerndatensätze ²⁵ auf dem Open Data Portal veröffentlichen

Verschiedentlich kritisierten Verwaltungsbeschäftigte, dass es an klaren Vorgaben fehlt, welche Daten vorrangig und nachfrageorientiert veröffentlicht werden sollen. Eine solche Konkretisierung ist dringend notwendig, um die abstrakte Forderung nach einer allgemeinen nachfrageorientierten Datenbereitstellung in die Tat umzusetzen. Aber welche Daten sind für eine Stadt und auch für die Berliner Verwaltungsbeschäftigten eigentlich relevant? Um diese Frage zu beantworten hat die ODIS in Anlehnung an die „High Value Datasets“ die Berliner Kerndatensätze als Orientierung für die Berliner Behörden herausgebracht, die Daten veröffentlichen wollen. Kerndatensätze sind Datensätze, die als besonders wertvoll für Berlin bewertet worden sind, weil sie einen großen Mehrwert für ein breites Publikum liefern können. Für diese Liste wurden die Bedarfe und Interessen der Wirtschaft, der Wissenschaft, der Zivilgesellschaft und der Verwaltung selbst betrachtet.

Ziel der Berliner Open Data-Strategie ist es, die Kerndatensätze in eine konkrete Veröffentlichungspraxis zu überführen und die Mehrwerte der Öffnung der Daten greifbarer zu machen.

- In den AG Open Data-Sitzungen werden unter Federführung von Open Data Berlin und der ODIS die Kerndatensätze den Open Data Beauftragten vorgestellt und eine sukzessive Veröffentlichung dieser Kerndatensätze mit ihnen geplant. Neben den Senats- und Bezirksverwaltungen werden auch landeseigene Betriebe oder das Amt-für-Statistik-Berlin-Brandenburg mit den „Kerndatensätzen“ adressiert, die über eine Vielzahl an wichtigen Daten verfügen.

Federführung: Senatskanzlei & ODIS

Beteiligte: ODIS unterstützend für die Veröffentlichung der Kerndatensätze mit den Senats- und Bezirksverwaltungen

Zeitraum: ab 2023 bis 2025

5.2.4.2 Dateninventuren in den Bezirks- und Senatsverwaltungen

Um sich einem umfassenden Datenangebot in der Berliner Behördenlandschaft zu nähern, bieten sich Dateninventuren in den Behörden an. Im Rahmen von Dateninventuren lassen sich zudem Open Data relevante Daten identifizieren, die dann in einem weiteren Schritt bereitgestellt werden können.

²⁵ <https://odis-berlin.de/projekte/kerndatensaetze/>

Mit einer Dateninventur soll auch das Datenmanagement innerhalb und zwischen den Behörden unterstützt werden. Für die Durchführung einer Dateninventur liegen erste Erfahrungswerte sowie unterstützende Ressourcen bereit. Es gilt nun, vermehrt Dateninventuren in den Behörden anzustoßen. In einigen Fällen bietet es sich für die Bezirks- und Senatsverwaltungen an, Dateninventuren gemeinsam mit dem Geschäftsprozessmanagement durchzuführen.

- Die Hauptverwaltungen und Bezirke stoßen interne Dateninventuren an. Die ODIS steht unterstützend zur Seite. Die Erfahrungen aus den Dateninventuren werden von der ODIS eingeholt und mit der Verwaltung z.B. im Rahmen der AG Open Data geteilt.

Federführung: Senatskanzlei & ODIS

Beteiligte: Haupt- und Bezirksverwaltungen, ODIS zur Unterstützung

Zeitraum: ab 2023 fortlaufend

5.2.5 Musterdatenkatalog

Bereits die geplante Veröffentlichung und dafür nötige Auseinandersetzung mit Daten kann eine organisationsinterne Qualitätssteigerung bewirken. Metadaten sind für die Qualitätssteigerung elementar und erlauben nicht nur die erleichterte Auffindbarkeit, sondern beschreiben die Daten auch für die Datennutzenden.

Damit die Datennutzung für Wirtschaft, Zivilgesellschaft und Wissenschaft optimal ausgeschöpft werden kann, müssen Daten in sehr guter Qualität bzw. Metadatenqualität auf dem Open Data Portal veröffentlicht werden.

Daten, die aufgrund der gut beschriebenen Metadaten auch mit den offenen Daten anderer Kommunen vergleichbar sind, bringen einen gesteigerten Nutzenwert.

Der Musterdatenkatalog, der in Zusammenarbeit zwischen dem Datenportal GovData des Bundes, der Open Knowledge Foundation, dem KDZ – Zentrum für Verwaltungsforschung und der Bertelsmann Stiftung entstanden ist, bietet eine Vergleichbarkeit der offenen Daten unterschiedlicher Datenportale in Deutschland.

Der Musterdatenkatalog definiert gut 300 abstrakte Musterdatensätze und ordnet diese rund 60 verschiedenen kommunalen Themenbereichen zu. Die tatsächlichen Datenkataloge (oder Datenportale) der Länder und Kommunen können dann für jeden ihrer Datensätze sagen, welchem Musterdatensatz dieser entspricht. Dadurch können vergleichbare Datensätze in unterschiedlichen Katalogen schnell gefunden werden. Dies erhöht die Vergleichbarkeit der Daten und kann gleichzeitig Anreize für Kommunen schaffen, weitere offene Daten bereitzustellen.

Damit auch die offenen Daten des Berliner Open Data-Portals mit den offenen Daten der anderen Kommunen verglichen werden können, werden die technischen Voraussetzungen für die Anwendung des Musterdatenkatalogs durch Berlin Online, dem Betreiber des Open Data-Portals, geschaffen.

Beim Anlegen neuer Datensätze können diejenigen, die Daten veröffentlichen bereits im Laufe des ersten Quartals den entsprechenden Musterdatensatz auswählen.

Bei in der Vergangenheit angelegten Datensätzen sollen die jeweiligen Behörden sukzessive die Metadaten mit Verknüpfungen zum Musterdatenkatalog anpassen können.

Im Rahmen der AG Open Data oder auch den Datenkompetenzschulungen wird die Anwendung des Musterdatenkatalogs anhand von Best Practice Beispielen den Open Data-Beauftragten und den Datenveröffentlichenden gezeigt.

Federführung: Senatskanzlei

Beteiligte: Berlin Online/Betreiber des Open Data Portals, Datenveröffentlichende

Zeitraum: ab 2023 bis 2025

5.2.6 Qualität & Technische Interoperabilität mit Linked Open Data

Die Qualitätskriterien offener Daten werden im sogenannten „Fünf-Sterne-Modell“ für offene Daten²⁶ gezeigt. Die Bewertung der Veröffentlichungsqualität von offenen Daten findet zunehmend anhand dieses „Fünf-Sterne-Modells“ statt.

Mit der Veröffentlichung von Daten als Open Data und im besten Fall, auch der Vernetzung von Daten untereinander, wächst die Wissensbasis der Verwaltung kontinuierlich. Um offene Verwaltungsdaten effektiv nutzen zu können, braucht es vor allem Interoperabilität der veröffentlichten Daten. Interoperabilität bezeichnet allgemein die Fähigkeit verschiedener Systeme und Organisationen zusammenzuarbeiten. Technisch sorgen insbesondere unklare Kennzeichnungen der für Datenbestände geltenden Nutzungsbedingungen, die Bereitstellung der Daten in herstellungsgebundenen und unstrukturierten Formaten (PDF) sowie die fehlende Skalierungsmöglichkeit bei händisch exportierten Datensätzen für Probleme.

Die Erfahrung zeigt: Eine automatisierte Verknüpfbarkeit als Linked Open Data (LOD) nach dem FAIR-Prinzip sorgt für optimale Vernetzung zwischen möglichst vielen Datenquellen. Das im Jahr 2016 entwickelte FAIR-Schema fordert, dass Daten auffindbar («findable»), zugänglich («accessible»), verknüpfbar («interoperable») und wiederverwendbar («reusable») sein müssen.²⁷ Die praktische Umsetzung dieser Prinzipien entspricht dem Fünf-Sterne-Ziel auf der aus einem Vortrag von Tim Berners-Lee entwickelten fünfstufigen Bereitstellungsskala für Open Data.²⁸

Eine solche Infrastruktur würde nicht nur die behördeninterne Nutzung vorhandener Informationen drastisch vereinfachen, sondern auch die – wo erforderlich umfassend anonymisierte – Veröffentlichung weitgehend automatisiert ermöglichen.

²⁶ <https://5stardata.info/de>

²⁷ (vgl. Wilkinson et al. 2016).

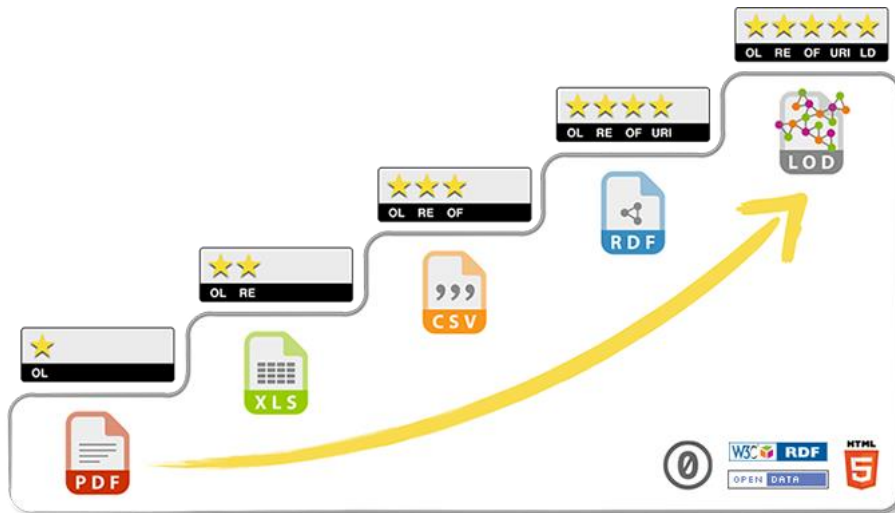
²⁸ (vgl. Berners-Lee 2009, Hasnain/Rebholz-Schuhmann 2018)

Daher sollte das Ziel der Open Data-Strategie sein, Daten in Fünf Sterne Qualität, also Linked Open Data, auf dem Open Data-Portal zur Verfügung zu stellen, die per Uniform Resource Identifier (URI) identifiziert sind und darüber direkt abgerufen werden können und ebenfalls per URI auf andere Daten verweisen.

Die bisherige Strategie der Bereitstellung endet nach dem „5-Sterne-Modell“ für offene Daten regelmäßig am dritten Stern – der manuellen Bereitstellung von Tabellen in offenen Formaten. Die automatisierte Verknüpfbarkeit als Linked Open Data als vierter und fünfter Stern des Modells bleibt in dieser Strategie meist von vornherein ausgeschlossen. Damit werden allein auf der Ebene der technischen Umsetzung große Chancen vergeben.

- Für das Ziel, eine datenbasierte/digitale Verwaltung zukunftssicher aufzustellen, müssen alle Fachverfahren so angepasst werden, dass voll automatisiert und in kurzen Abständen Datenexporte ermöglicht werden und die Ergebnisse direkt im Berliner Open Data-Portal erscheinen. Die Daten müssen so wie bisher auch in einfach für Menschen zu lesenden Tabellen, aber zusätzlich in maschinell lesbaren Formaten erscheinen. Die neuen Formate müssen semantisch korrekt ausgezeichnet werden und alle Objekte der Berliner Verwaltung über einen eindeutigen Identifier (ID bzw. URI) beschrieben werden. Somit gilt über alle Verwaltungs- und Fachverfahrensgrenzen hinweg der gleiche Identifier für jede gleiche Informationseinheit. Dies ist die Grundlage für die geforderte Verlinkung der Daten und wird über den Aufbau eines entsprechenden offenen Registers (Registry) umgesetzt.
- Für die Umsetzung von Linked Open Data müssen zuerst die technischen Voraussetzungen beim Open Data-Portal und dem zukünftigen Data Hub geschaffen werden. Für den Aufbau von Speicher- und Verbreitungssystemen für Linked Open Data werden sog. „Triple Stores“ strategisch implementiert und in die IKT-Architektur aufgenommen. Triple Stores sind (wie oben bereits erwähnt) spezielle Datenbanken, die den Einsatz von Linked Open Data ermöglichen. Es gibt zahlreiche Triple-Store-Produkte am Markt, deren gemeinsames Merkmal die Unterstützung von Linked-Data-Standards wie dem RDF-Modell und der SPARQL-Abfragesprache ist. Weit verbreitete und seit vielen Jahren erprobte Beispiele sind etwa *Open Link Virtuoso*, *Apache Jena* oder *RDF4J*, die alle als Open Source, oder zumindest in einer Open-Source-Variante verfügbar sind. Diese Produkte sollten ebenfalls in die IKT-Architektur aufgenommen werden.

Abbildung: 5-Sterne-Open Data-Modell nach Tim Berners-Lee



Quelle: <https://5stardata.info/de/>

- **Barcamp Linked Open Data:** in diesem Barcamp soll die Verwaltung gemeinsam mit einer Organisation der Open Data-Zivilgesellschaft Prototypen im Bereich Linked Open Data in einem Wettbewerb entwickeln und öffentlichkeitswirksam präsentieren. Das Ziel ist es, Linked Open Data mit den Prototypen der Verwaltung näher zu bringen und in die Anwendung im Open Data-Portal zu bringen. Diese Beispiele für Linked Open Data sollen konkrete Anschauungen über Effizienzgewinne für die Verwaltung, Kostenersparnisse und einfachere Zusammenarbeit zwischen den Verwaltungen bieten.

Federführung: Senatskanzlei

Beteiligte: Senatskanzlei für die Aufnahme Linked Open Data in die IKT-Architektur, Berlin Online/Betreiber des Open Data Portals für die Schaffung der technischen Voraussetzungen für Linked Open Data und die ersten Open Data- Anwendungsbeispiele

Zeitraum: ab 2023 bis 2025

5.2.7 Internes Datenmanagement etablieren

Das interne Datenmanagement beinhaltet technische Aspekte der Regelung von Datennutzung (Datenqualität, Datenpflege, Qualitätsmanagement, technische Standards).

Die technischen Möglichkeiten des Open Data-Portals wollen wir erweitern und auch Speicherkapazitäten mit dem Data Hub anbieten als auch die Visualisierungen von veröffentlichten Daten wie zum Beispiel mit Dashboards mit sogenannten Datenstories vorantreiben.

Da das Open Data-Portal ein Metadatenportal ist und keine Speichermöglichkeiten zur Verfügung stellt, sollte ein Data Hub im Land Berlin aufgebaut werden, das unter anderem die interne Datenspeicherung und auch die Veröffentlichung von Open Data relevanten Datensätzen über eine Schnittstelle vom Data Hub auf dem Open Data-Portal ermöglicht.

Daten und die dazugehörigen Metadaten sollen in Zukunft in hoher Qualität bereitgestellt werden, so dass Daten besser durch Dritte genutzt werden können. Für die Qualität der Metadaten können sowohl die nischenspezifischen Kategorisierungen (aus dem Musterdatenkatalog) als auch die Verlinkbarkeit der Datensätze im Sinne von Linked Open Data herangezogen werden.

5.2.7.1 Leitprinzip „Open-by-default“-Strategie

Eine automatisierte Bereitstellung von Open Data ist ein Indikator für die erfolgreiche Umsetzung von Verwaltungsdigitalisierung und des modernen Datenmanagements in der Verwaltung. „Open-by-default“ bedeutet, das Vorbereiten der Systeme und Abläufe für eine Veröffentlichung der Informationen von Grund auf in alle Systeme und Prozesse zu integrieren, so dass Aufwände für eine Veröffentlichung sinken.

Nach dem Prinzip „Open-by-default“ sollen die Daten aus der Verwaltung oder auch aus externen Plattformen mit so wenigen „Handgriffen“ wie möglich automatisch an das Open Data-Portal des Landes Berlin angebunden werden. Daher wird Open Data nur dann funktionieren, wenn die bestehenden Fachverfahren Öffnung und Bereitstellung von Daten in die Arbeitsprozesse integrieren. Fachverfahren auf allen Ebenen müssen daher auf Ihre Anbindung an das Open Data-Portal überprüft werden.

Wenn Daten „per-default“ bereitgestellt werden sollen, wird im besten Fall eine technische Infrastruktur benötigt, die eine möglichst automatisierte Bereitstellung ermöglicht.

- In der IKT-Architektur ist bereits verankert, dass neue Fachverfahren eine Schnittstelle beinhalten sollen, die einen Export von Informationen in maschinenlesbaren Dateiformaten und Metadaten weitestgehend über eine automatisierte Bereitstellung der Informationen auf dem Open Data-Portal gewährleisten.
- Im Rahmen der Open-by-default-Strategie sollen bei den älteren Bestandsfachverfahren die Schnittstellen für eine automatisierte Bereitstellung nachträglich idealerweise bei wesentlichen Weiterentwicklungen nach und nach implementiert werden. Die Anbindung der Open-Data-relevanten-Fachverfahren aus der IT-Bestands- und Planungsübersicht (IT-BePla) soll fortgeführt werden. Über einen längeren Zeitraum wird eine Inventur aller Fachverfahren aus der IT-Bestands- und Planungsübersicht (IT-BePla) durchgeführt, um mittelfristig alle Open Data-relevanten bestehenden und zukünftigen Fachverfahren an das Open Data-Portal anzubinden.
- 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 gewährleisten. Durch den Einsatz von Open-Source-Software in der Verwaltung können langfristig Abhängigkeiten von einzelnen Anbieterinnen und Anbietern verhindert werden.

- In den Beschaffungsverträgen bei denen Systeme, Plattformen, Visualisierungen oder Fachverfahren durch die Vertragspartnerin oder den Vertragspartner betreut oder erstellt werden, soll es einen Hinweis geben, dass eine API bzw. Programmierschnittstelle von Anfang an auch für eine „Open-by-default“- Veröffentlichung auf dem Open Data-Portal mitbedacht werden muss.

Federführung: Senatskanzlei

Beteiligte: Senatskanzlei IT-Bestands- und Planungsübersicht der Fachverfahren des Landes Berlin (IT Be-Pla) und Fachverfahrensverantwortliche der Senats- und Bezirksverwaltungen

Zeitraum: ab 2024 bis 2026

5.2.7.2 Open Data in Verträgen für Datensouveränität

Die Datensouveränität ist die „Verfügungsgewalt“ über die eigene Datennutzung. Diese muss bei der Verwaltung selbst liegen, damit die Daten umfangreich von der Öffentlichkeit, von Unternehmen, der Zivilgesellschaft und der Wissenschaft weiterverwendet werden können. Verwaltungen müssen sich die Hoheit über ihre Daten sichern. Sie müssen souverän über ihre eigene Datennutzung entscheiden und Einschränkungen und Abhängigkeiten von Dritten reduzieren. Nicht nur in den Fachbereichen oder bei den Open Data-Beauftragten, sondern auch in den Vergabestellen sind und werden Fragen des Umfangs der Datennutzung relevant. Verwaltungen sind von Anfang an auf der sicheren Seite, wenn sie die Kenntnisse zu Datensouveränität und Datenhoheit bereits in den Vergabeverfahren und bei der Beschaffung berücksichtigen. Das bedeutet, wenn Dienstleisterinnen und Dienstleister involviert sind, sollte bereits bei der Erhebung und in der Datenbeschaffung entschieden werden, wem welche Rechte an den Daten zustehen sollen und das jeweilige Ergebnis mit der beauftragten Dienstleisterin oder dem beauftragten Dienstleister vereinbart werden.

Damit wird die Basis für die künftige Weiterverwendung von Daten und die Veröffentlichung als Open Data gelegt.

- Bei Abschluss oder Anpassung von vertraglichen Regelungen mit Dritten, welche die Erhebung, Erstellung, Verarbeitung oder Nutzung von Informationen betreffen, soll darauf hingewirkt werden, dass das Recht der betroffenen Behörde auf die uneingeschränkte öffentliche Bereitstellung dieser Informationen zur freien Weiterverwendung im Sinne des §9 Absatz 1 der Open Data-Verordnung aufgenommen wird.²⁹ Das Ziel ist, Situationen zu vermeiden, in denen die Verwaltung nicht mehr die komplette Kontrolle über ihre eigenen Daten hat, obwohl deren Erhebung und/oder Verarbeitung aus Behördenmitteln finanziert wur-

²⁹ Absatz 2 aus §7 der Open Data-Rechtsverordnung

den. Daten können von den Auftragnehmenden erhoben, verarbeitet oder gespeichert/gehostet werden. Die Verwaltung kann auch Daten Dritter einkaufen. In allen Fällen geht es darum, dass in der Beschaffung und auch in der Vertragsgestaltung, sämtliche Rechteeinräumungen von der Auftragnehmerin oder dem Auftragnehmer an die Auftraggeberin oder dem Auftraggeber und auch umgekehrt sowie die Pflicht zur Sicherstellung des technischen Zugriffs auf die im Vertrag erhobenen/verarbeiteten Daten geregelt werden.

Folgende Open Data-Muster-Klausel wurde von Open Data Berlin und Berlin Online, dem Betreiber des Open Data-Portals formuliert. Es wird empfohlen, die Klausel als beispielhafte Orientierung für Verträge zu verwenden, bei denen Systeme, Plattformen, Visualisierungen oder Fachverfahren durch Vertragspartner betreut oder erstellt werden, und eine API bzw. Programmierschnittstelle von Anfang an auch für eine „Open-by-default“-Veröffentlichung auf dem Open Data-Portal mitbedacht werden muss. Die Muster-Klausel entbindet die bereitstellungsverpflichteten bzw. vertragsschließenden Stellen nicht von ihrer Verantwortung und Pflicht zur Prüfung und etwaigen Anpassung der Klausel im jeweiligen Einzelfall.

„Die Auftragnehmerin/der Auftragnehmer wird im Rahmen des Beschaffungsvertrages von Informations- und Kommunikationstechnik (Hardware/Software) verpflichtet, die Veröffentlichung von Informationen auf dem Open Data-Portal daten.berlin.de zu ermöglichen. Dies betrifft sowohl Daten als auch Metadaten.

Daten, die gemäß §3 Absatz 2 der Open Data-Verordnung des Landes Berlin veröffentlicht werden sollen, müssen in maschinenlesbaren Formaten über einen öffentlich verfügbaren Download-Endpunkt oder eine Programmierschnittstelle (API) veröffentlicht werden.

Metadaten müssen gemäß dem Metadatenschema des Landes Berlin über die Programmierschnittstelle des Open Data-Portals veröffentlicht werden.

Folgende Open Data-Muster-Klausel wurde von Open Data Berlin und dem Justitiariat formuliert. Es wird empfohlen, die Klausel als beispielhafte Orientierung für Verträge zu verwenden, bei denen Dienstleisterinnen und Dienstleister vom Land Berlin mit der Erhebung, Speicherung oder Verarbeitung von Daten beauftragt werden. Die Muster-Klausel entbindet die bereitstellungsverpflichteten bzw. vertragsschließenden Stellen nicht von ihrer Verantwortung und Pflicht zur Prüfung und etwaigen Anpassung der Klausel im jeweiligen Einzelfall.

"Die Auftragnehmerin/der Auftragnehmer wird verpflichtet, im Rahmen dieses Auftrages durch ihn/sie erhobene und erstellte Informationen unverzüglich in einem offenen, maschinenlesbaren Format im Sinne des §3 Absatz 2 der Open Data-Verordnung des Landes Berlin dem Auftraggeber/der Auftraggeberin bereitzustellen, sofern nicht personenbezogene Daten betroffen sind. Weiterhin vereinbaren die Vertragsparteien, dass diese Informationen von dem Auftraggeber/der Auftraggeberin unter einer offenen Lizenz gemäß §9 Absatz 1 und 2 der Open Data-Verordnung des Landes Berlin auf dem Open Data-Portal <https://daten.berlin.de> zur uneingeschränkten privaten wie kommerziellen Nutzung durch Dritte bereitgestellt werden können, soweit nicht eine der in §5 der Open Data-Verordnung des Landes Berlin genannten Ausnahmen vorliegt."

Federführung: Senats- und Bezirksverwaltungen, die Verträge für Fachverfahren
oder

Aufträge für die Datenverarbeitung vergeben

Beteiligte: Fachverfahrensverantwortliche, Fachbereiche, die Aufträge vergeben;
Vergabestellen

Zeitraum: ab 2023 bis 2025

5.2.7.3 Data Hub mit „Cloud Speicher“/Datenplattform

Das Nutzungsversprechen von Open Data kann nur eingelöst werden, wenn mit einem modernen Datenmanagementsystem gearbeitet wird. Aufgrund der wachsenden Bedeutung von Datenmanagement und -analyse für öffentliche Verwaltungen sollten zudem Grundlagen für eine integrierte Datenmanagementplattform geschaffen werden, die den übergreifenden Zugriff für die Verwaltung auf größere Datenmengen erlaubt.

Die Datenmanagementplattform soll sich in die bestehende IKT-Architektur einfügen und diese um eine übergreifende Daten-Architektur ergänzen. 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.

Verschiedene Datenplattformen des Landes Berlin sollten über Schnittstellen an die gemeinsame Datenplattform angebunden werden können. Upload- und Downloadprozesse in der Datenplattform sollten möglichst niedrigschwellig funktionieren.

- Für die Weiterentwicklung des Konzepts einer gemeinsamen Datenmanagementplattform, das damals im Auftrag der SenInnDS und der SenWiEnBe von der KPMG entwickelt wurde, soll eine „Governance für den Data Hub“ für eine gemeinsame Infrastruktur, die die entsprechenden Ressourcen, eindeutige Verantwortlichkeiten und die erforderlichen Kompetenzen im Umgang mit dem Data Hub regelt, auf den Weg gebracht werden. Schließlich soll das Land Berlin einen Auftrag an einen Dienstleister für die Umsetzung des Data Hubs vergeben, der die Anforderungen der damals entwickelten KPMG Studie „Konzept eines Data Hubs zur Vernetzung städtischer Datenquellen“ erfüllt.

Federführung: Senatskanzlei

Beteiligte: Dienstleisterin oder Dienstleister/Auftragnehmerin oder Auftragnehmer für die
Umsetzung des DataHubs für das Land Berlin

Zeitraum: ab 2023 bis 2025

5.2.8 Intelligente Datennutzung

Im Rahmen der Open Data-Strategie soll die Datennutzung weiter sichtbar gemacht werden. Der Nutzen von Open Data wurde bislang auf die Entstehung neuer Geschäftsmodelle und wertschöpfender Ziele verengt. Dabei verspricht Open Data für alle Stakeholder enorme Nutzungspotenziale. Insbesondere in der Verwaltung können Tools und Open Data-basierte Anwendungen einen großen Mehrwert schaffen.

5.2.8.1 Open Data-Anwendungen für die Verwaltung

Einige prototypische Anwendungen, die von der ODIS entwickelt wurden, deuten das Potenzial offener oder maschinenlesbarer Daten für die Verwaltung an. Als Beispiele können hier das Projekt zu offenen Haushaltsdaten oder das Organigramm-Tool der ODIS genannt werden. Um den Mehrwert von Open Data zu zeigen, werden bestehende Anwendungen weiterentwickelt bzw. weitere Anwendungen entstehen. Zudem gilt es, prototypische Anwendungen in einen Regelbetrieb zu überführen und für eine Verstetigung zu sorgen, um eine nachhaltige Nutzung zu gewährleisten.

- Die ODIS entwickelt regelmäßig weitere prototypische Anwendungen oder Tools, die als Demonstratoren das Potenzial offener Daten aufzeigen. Ziel in der Legislaturperiode ist es, mindestens eine prototypische Anwendung in den Regelbetrieb zu überführen.

Federführung: ODIS

Beteiligte: Senats- und Bezirksverwaltungen

Zeitraum: ab 2023 fortlaufend

- Open Data Berlin wirkt darauf ein, dass die Rahmenbedingungen geschaffen werden, dass prototypische Anwendungen aus dem Status eines Prototyps herauskommen und verstetigt werden. Dafür wird Open Data Berlin das Gespräch mit relevanten Akteuren in den Behörden suchen.

Federführung: Senatskanzlei

Beteiligte: Senatskanzlei, beteiligte Behörden

Zeitraum: ab 2023

5.2.8.2 Open Data-Anwendungen in Kollaboration mit dem Open Data Ökosystem

Neben der Verwaltung können Open Data-Anwendungen für und in Kollaboration mit Wirtschaft, Zivilgesellschaft und Wissenschaft dabei helfen, den Nutzen von Open Data zu kommunizieren. Dazu wird die Senatskanzlei mit Unterstützung der ODIS die Entwicklung digitaler Produkte anregen.

Federführung: Senatskanzlei

Beteiligte: Open Data-Ökosystem

Zeitraum: ab 2023

6. Ausblick

Die Open Data-Strategie 2023 versteht sich als eine lernende Strategie, die kontinuierlich mit den Stakeholdern der Stadtgesellschaft weiterentwickelt werden soll. Dafür sind die Vernetzungsformate vorgesehen, die regelmäßige Evaluationen mit Netzwerk- und Lernformaten für alle beteiligten Stakeholder ermöglichen. Dort soll zum einen auch über die Fortschritte der Maßnahmen berichtet, zum anderen aber auch die Möglichkeit geboten werden, sich in der Weiterentwicklung der Strategie einzubringen.

7. Anhang

- Strategieprozess: Open Data-Strategie Berlin: Zielstellung, Prozess und Zeitschiene
- „5-Pager“ Erkenntnisse aus dem Beteiligungsprozess zur Erarbeitung einer neuen Open Data-Strategie für Berlin – Open Knowledge Foundation e.V.
- Ergebnisberichte der Strategieworkshops <https://strategie.odis-berlin.de/>
- Open Data in der Berliner Verwaltung, Status quo, Bedarfe und Perspektiven

8. Maßnahmenplanung

Maßnahme	Maßnahmenplanung (Beschreibung)	Zeitraum
Data Literacy und Data Governance	Im Rahmen einer Studie, die in Auftrag gegeben wird, soll ein Data Governance Modell als Vorschlag zur Umsetzung für die Berliner Behördenlandschaft erarbeitet werden.	2024-2026
Datenkompetenz	Ein vom Land Berlin zertifizierter Kurs für Datenmanagement und Open Data wird den datenverantwortlichen Verwaltungsbeschäftigten angeboten. Die ODIS erweitert das Angebot an Unterstützungsmaterialien (Checklisten, Handouts, Videoformaten, Schulungen)	2023-2026
Open Data Mindset stärken	Changemanagement Formate für den Verwaltungskulturwandel in der AG Open Data. Erweiterung des Open Data Handbuchs mit den Changemanagement Formaten im Umgang mit den neuen technischen Systemen	2023-2026
Berlins Kerndatensätze auf dem Open Data Portal veröffentlichen	Kerndatensätze der ODIS sollen in eine konkrete Veröffentlichungspraxis überführt werden, Senats- und Bezirksverwaltungen müssen diese Kerndatensätze veröffentlichen	2023-2026
Dateninventuren durchführen	Die Senats- und Bezirksverwaltungen führen mit der ODIS in ihren Behörden Dateninventuren zur Identifikation von Daten und datenhaltenden Stellen durch.	2023-2026

Musterdatenkatalog	Das Open Data-Portal wird nach dem Musterdatenkatalog Prinzip die Metadaten auflisten, so dass die Datenveröffentlichenden diese nischenspezifischen Metadaten auswählen können. Hierdurch sind die Daten für die Datennutzenden besser kommunal vergleichbar.	2023 ff.
Qualität & technische Interoperabilität mit Linked Open Data	Im Open Data-Portal und im Data Hub sollen für den Aufbau von Speicher- und Vertriebssystemen für Linked Open Data Triple Stores in die IKT-Architektur aufgenommen werden. Neue Daten müssen semantisch korrekt ausgewiesen werden und alle Objekte der Berliner Verwaltung über einen eindeutigen Identifier (ID bzw. URI) beschrieben werden.	2024-2026
„Open-by-default“-Strategie	Über einen längeren Zeitraum wird eine Inventur aller Fachverfahren aus der IT-Bestands- und Planungsübersicht durchgeführt, um mittelfristig alle Open-Data-relevanten bestehenden und zukünftigen Fachverfahren an das Open Data-Portal anzubinden. Eine Open Data-Klausel wird in die Beschaffungsverträge für IKT-Systeme integriert, die eine Schnittstelle für eine Veröffentlichung von Anfang an berücksichtigt.	2023-2026
Open Data in Verträgen für Datensouveränität	Die Open Data-Klauseln werden Pflicht für Ausschreibungen und Vergaben, bei denen Daten die Grundlage des Auftrages bilden oder IT-Systeme beschafft werden, die Schnittstellen haben müssen. Denn Verwaltungen müssen souverän über ihre eigene Datennutzung entscheiden und Einschränkungen und Abhängigkeiten für Open Data durch Dritte vermeiden.	2023-2026
Data Hub mit „Cloud Speicher“	Zusätzlich zu der bestehenden Studie aus 2019 „Konzeption eines Data Hubs zur Vernetzung städtischer Datenquellen“ für die technischen Anforderungen des Data Hubs, sollte die „Governance für den Data Hub“ für eine gemeinsame Infrastruktur, die entsprechende Ressourcen, eindeutige Verantwortlichkeiten und erforderliche Kompetenzen im Umgang mit dem Data Hub konkretisiert, geregelt werden.	2023-2026

Open Data Anwendungen für die Verwaltung & Wirtschaft	Die ODIS entwickelt prototypische Anwendungen und Tools, die als Demonstratoren das Potenzial offener Daten zeigen. Ziel der Legislaturperiode ist es, mindestens eine Anwendung in den Regelbetrieb zu überführen. (z.B. Organigramm-Tool, Schuleinzugsgebieterechner)	2023-2026
---	---	-----------



BERLIN
OPEN
DATA

Open Data Strategie Berlin 2022

Zielstellung, Prozess und Zeitschiene

Hintergrund zur neuen Open Data Strategie

Im Jahr 2011 wurde die erste [Berliner Open Data Strategie](#) (vom Fraunhofer FOKUS im Auftrag der Senatsverwaltung für Wirtschaft, Technologie und Forschung) veröffentlicht. Im gleichen Jahr ist das [Berliner Datenportal](#) gestartet. Seitdem ist viel passiert: Aktuell sind über 2.900 offene Datensätze auf dem Portal veröffentlicht und das Angebot wächst kontinuierlich. Viele Verwaltungen und deren Beschäftigte haben sich in den letzten Jahren erstmals mit den Potenzialen offener Daten auseinandergesetzt, andere wiederum veröffentlichen bereits seit Jahren regelmäßig Daten auf dem Open Data Portal.

Mit der Verabschiedung des E-Government-Gesetzes (EGovG) Berlin im Juni 2016 sind die rechtlichen Rahmenbedingungen für Open Data in der Verwaltung fest verankert. Am 01.01.2021 ist zusätzlich die [Open Data Verordnung](#) in Kraft getreten. Sie konkretisiert den §13 EGovG, indem sie die Art und Weise der Datenveröffentlichung und auch die Formate und Nutzungsbedingungen festlegt, unter denen die Daten auf dem Open Data Portal bereitgestellt werden sollen. Damit wurden in den letzten Jahren die technischen und rechtlichen Grundlagen für Open Data in der Verwaltung erfolgreich gelegt.

Jetzt gilt es für die nächsten fünf Jahre eine neue Open Data Strategie mit konkreten Vorhaben aufzusetzen, die die Bedarfe des gesamten Berliner Open Data Ökosystems aus Wirtschaft, Verwaltung, Wissenschaft, Community und den Bürger:innen berücksichtigt.

Prozess

Der Prozess zur Berliner Open Data Strategie startete im Februar 2022. Er umfasst eine **Online-Beteiligung**, fünf **Beteiligungs-Workshops** und eine kontinuierliche **Stakeholder-Betreuung**. Ende Juli 2022 soll das Strategiepapier fertiggestellt werden.

Dokumentation

Im Sinne des Open Data Ansatzes von "open by default" werden **Prozessschritte und Ergebnisse dokumentiert** und zur Verfügung gestellt.



Dafür gibt eine [öffentliche Informationsseite](#). Dort finden sich auch wichtige **Hintergründe, Definitionen, Termine und Kontaktdaten**. Dazu erfolgen regelmäßige Updates in denen über Fortschritt sowie inhaltliche Neuigkeiten berichtet werden. Über Blogbeiträge werden außerdem Herausforderungen und Learnings kommuniziert, die aus dem Strategieprozesse resultieren.

Die Open Data Stakeholder:innen

Bei der Entwicklung der Strategie für Berlin stehen die Verwaltung, Wirtschaft, Wissenschaft und Stadt- bzw. Zivilgesellschaft im Mittelpunkt. Bei den anstehenden **Beteiligungphasen** sollen möglichst viele Perspektiven berücksichtigt werden. Bis zum Juni 2022 können alle (Open) Data Interessierten und Berliner:innen ihre Ideen & Vorstellungen für die zukünftige Strategie einbringen.

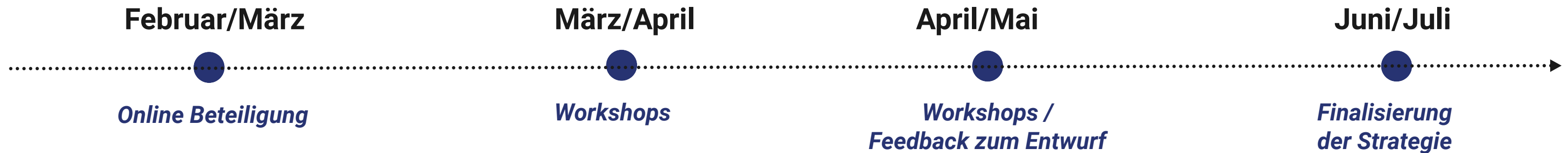
Während des gesamten Prozesses werden die Zwischenergebnisse auf mein.berlin.de zur Diskussion und Online-Kommentierung veröffentlicht.

In den **Stakeholder:innen-Workshops** steht der Austausch mit Expert:innen im Fokus. Die Einladung zu den jeweiligen Workshopterminen erfolgen über das Strategieteam.

Stakeholder-Gruppen

Interessierte Öffentlichkeit	Alle interessierten Einzelpersonen, Gruppen der Stadtgesellschaft, (Open) Data Initiativen und Netzwerke in Berlin sind eingeladen, sich an den Online-Umfragen und Kommentierungsphasen im Prozess zu beteiligen.	Online-Beteiligung auf mein.berlin: • 18.02-18.03.22 • ab Mai 2022 Kommentierung des Strategieentwurfs
Verwaltung	Mitarbeiterinnen und Mitarbeiter aus der gesamten Berliner Verwaltung, die am Prozess mitwirken wollen, können sich unabhängig von ihrer Zuständigkeit und Handlungsebene einbringen.	Online-Beteiligung auf mein.berlin: • 18.02-18.03.22 Stakeholder:innen Workshops: • Verwaltung 29.03.22 - 13:00 - 17:30 • Verwaltung und Zivilgesellschaft 12.05.22 - 17:30-20:00
Zivilgesellschaft	Nutzer von Open Data aus der Zivilgesellschaft können anwendungsorientierte Erwartungen und Bedürfnisse einbringen. Vertreter:innen werden aktiv eingeladen, sich an den Konsultationen und Workshops zu beteiligen.	Online-Beteiligung auf mein.berlin: • 18.02-18.03.22 Stakeholder:innen Workshops: • Zivilgesellschaft 31.03.22, 17:30-20:00 • Verwaltung und Zivilgesellschaft 12.05.22 - 17:30-20:00
Wirtschaft	Offene Daten bergen ein hohes wirtschaftliches Potenzial. Vertreter:innen aus der Wirtschaft sind daher aufgerufen, ihre Perspektiven und Bedürfnisse in den Open Data Strategieprozess einzubringen, in Hinblick auf die Nutzung, wie auch die Bereitstellung eigener Daten.	Online-Beteiligung auf mein.berlin: • 18.02-18.03.22 Stakeholder:innen Workshop: • 28.04.22 - 13:00-17:30
Wissenschaft	Data Science, Big Data und Dateninfrastrukturen sind nur einige von vielen Themenfeldern und Methoden, die helfen Erkenntnisse aus offenen Daten zu generieren. Impulse und Erfahrungen aus der Wissenschaft sind daher wichtig für den Open Data Strategieprozess und erweitern die Perspektive auf das Themenfeld Open Data.	Online-Beteiligung auf mein.berlin: • 18.02-18.03.22 Stakeholder:innen Workshop: • 19.05.22 - 12:00-16:00

Open Data Strategie Berlin 2022 - Zeitschiene



Öffentlichkeit

18.02 - 18.03.2022
**Einreichungen von Ideen,
Wünschen und Anregungen**

**Kommentierung
und Ergänzung der
Zwischenergebnisse aus den
Workshops und der Eckpunkte der
Strategie**

Expert:innen

29.03.2022
Workshop #1
Verwaltung

31.03.2022
Workshop #2
**Zivil-
gesellschaft**

28.04.2022
Workshop #3
Wirtschaft

12.05.2022
Workshop #4
**Verwaltung
+ Zivilgesell.**

19.05.2022
Workshop #5
Wissenschaft

**Feedback zu
den Eckpunkten**

Strategieteam

**Stakeholder:
innen
Identifikation**

**Wissenspeicher
anlegen**

**Aufruf zur
Beteiligung**

Einladung Workshops

**Eckpunkte
der Strategie**

**Integration in das
Umfeld**

**Stakeholder:innenbetreuung und Aufbereitung der
Zwischenergebnisse**

**24.05.2022
Zwischenergebnisse
und Feedback der
Workshops**

**Finalisierung
der Strategie**

**31.07.2022
Veröffent-
lichung der
Strategie**

Weitere Informationen

Öffentliche Informationsseite zum Strategieprozess:

<https://strategie.odis-berlin.de/>

Hier finden Sie den aktuellen Stand im Prozess, Dokumentationen, Ergebnisse und Wissenswertes.

Das Strategieteam

Senatsverwaltung für Wirtschaft, Energie und Betriebe

Betül Özdemir

✉ betuel.oezdemir@senweb.berlin.de

Open Knowledge Foundation

Henriette Litta,
Stefan Kaufmann,
Walter Palmetshofer

✉ opendataberlin@okfn.de

🐦 @okfde

Open Data Informationsstelle

Lisa Stubert,
Klemens Maget

✉ odis@ts.berlin



Erkenntnisse aus dem Beteiligungsprozess zur Erarbeitung einer neuen Open-Data-Strategie Berlin

29.06.2022

Das Wichtigste in Kürze

- ⇒ Die automatisierte Bereitstellung von Open Data ist ein Indikator für die erfolgreiche Umsetzung von Verwaltungsdigitalisierung und des modernen Datenmanagements in der Verwaltung.
- ⇒ Die Bereitstellung von Open Data braucht infrastrukturelle Grundlagen: Neben adäquater Ausstattung (Hard- und Software) ist auch eine Anpassung von Prozessen, Strukturen und Zuständigkeiten unumgänglich.
- ⇒ Eine moderne Verwaltung mit komplexen Aufgaben braucht hochwertige und verknüpfte Datensätze (Linked Open Data).
- ⇒ Der Berlin Data Hub ist ein wichtiger Baustein für ein modernes Datenmanagement und sollte zeitnah auf den Weg gebracht werden. Gleichzeitig muss die Strategie bereits jetzt auf Linked Open Data ausgerichtet sein und parallel mit dem geplanten Data Hub die dafür notwendige Infrastruktur aufgebaut werden.

Einleitung

Mehr als 10 Jahre nach dem ersten Open-Data-Strategiepapier überarbeitet das Land Berlin diese Strategie, um Impulse für die zukünftige Schwerpunktsetzungen zu erarbeiten. Die Federführung für den Strategieprozess liegt bei der Senatsverwaltung für Wirtschaft, Energie und Betriebe; für die Umsetzung von Beteiligungsworkshops wurde die Open Knowledge Foundation Deutschland beauftragt; die Open Data Informationsstelle Berlin (ODIS) unterstützt den Prozess.¹ Unter Beteiligung der Berliner Öffentlichkeit und des Berliner Open Data Ökosystems, bestehend aus Verwaltung, Wirtschaft, Wirtschaft und Zivilgesellschaft konnten in der ersten Jahreshälfte 2022 eine Vielzahl an Perspektiven auf den Status Quo sowie Anregungen und Impulse in Hinblick auf die zukünftige Ausrichtung von Open Data gewonnen werden. Im Vorfeld war die allgemeine Öffentlichkeit zu einer Online-Beteiligung eingeladen worden.

¹ Auf der Website <https://strategie.odis-berlin.de/> wird der Strategieprozess offen und transparent begleitet. Dort ist alles Wissenswertes zur Open Data Landschaft Berlins und rund um den Prozess, die Dokumentationen und Ergebnisse der Workshops und die beteiligten Akteur:innen zu finden.

In diesem Dokument werden die bisherigen Kernergebnisse des Beteiligungsprozesses zusammengefasst, sowie am Ende illustrative Schlaglichter aus den Workshops skizziert, die Interesse an einer weiteren Beschäftigung, Vertiefung und Diskussion mit dem Thema wecken sollen. Workshopübergreifend haben sich drei Erkenntnisse herauskristallisiert: Open Data braucht eine neue Erzählung, die sich auf die interne Nutzung in der Verwaltung und die damit verbundene Nützlichkeit für die einzelnen Verwaltungsbeschäftigten fokussiert. Teil dieser Erzählung muss die Verknüpfbarkeit von Datensätzen über viele Verwaltungsquellen und darüber hinaus sein, um verteilte Informationen zu einem gemeinsamen, abfragbaren Wissensschatz verknüpfen zu können. Der geplante Data Hub ist wichtig, braucht klare Vorgaben und sollte möglichst modular in seinen Funktionalitäten wachsen, während gleichzeitig die Grundlagen für Linked Open Data als Default gelegt werden müssen.

Das Versprechen von Open Data muss neu justiert werden: **Open Data ist Verwaltungsmodernisierung und führt zu effizienterer Arbeit**

Der Mehrwert von Open Data wird vielfach mit extrinsischen Motivationsfaktoren begründet. Diese reichen von Argumentationen rund um Transparenz und Nachvollziehbarkeit von Verwaltungsentscheidungen bis hin zu ungehobenem Potenzial durch die Nutzung der Daten durch Dritte. Forschende könnten die Daten wissenschaftlich nutzen, Unternehmen könnten innovative Geschäftsideen mit den Daten entwickeln und die engagierte Zivilgesellschaft könnte Lösungsvorschläge für gesellschaftliche Probleme anhand der Offenen Daten erarbeiten. So richtig diese Argumente sind, scheinen sie dennoch in den Verwaltungen den Eindruck zu erzeugen, dass sie Open Data in erster Linie für Dritte bereitstellen und dieser erhebliche Aufwand für sie selbst keinen Nutzen bringt.

Über den gesamten Beteiligungsprozess hinaus ist deutlich geworden, dass der Verwaltung die zentrale Rolle im Open-Data-Ökosystem zukommt. So nannten bei der Online-Befragung die Mehrheit der Beteiligten mit 42 Prozent bei der Verwaltung das größte Nutzungspotenzial von Open Data. Unter den Verwaltungsakteur:innen sahen dies mit 55 Prozent sogar noch mehr. Und auch in den Beteiligungsworkshops war übergreifend Konsens, dass die Berliner Verwaltung zwar sowohl ein enormes Potenzial offener Daten besitzt als auch die größte Nutznießerin offener Daten sein kann, dieses Potenzial aber bei weitem noch nicht ausschöpft. Gleichwohl beispielsweise auf dem Berliner Open Data Portal immer mehr Datensätze aus der Berliner Verwaltung veröffentlicht werden oder Fachverfahren an das Open Data Portal angeschlossen werden, setzt sich das Prinzip *open by default*, welches auch in der Berliner Open Data Rechtsverordnung verankert ist, in den Berliner Behörden nur langsam durch. Es bedarf daher eines grundlegenden Verständnis über eine Zielrichtung und einer Herangehensweise für die Berliner Verwaltung, die über personelle, finanzielle und organisatorische Anpassungen hinaus Open Data in der Berliner Verwaltung zukünftig tragen sollen. Dabei ist klar, dass Open Data in der Berliner Verwaltung ein Spiegelbild der generellen Data Governance der Berliner Verwaltung ist.

Die Bereitstellung von Open Data muss in der Weise in den eigenen Verwaltungsalltag integriert werden, dass sie (mittelfristig) *nicht* zu einem Doppel- oder Mehraufwand führt, sondern die eigene Arbeit effizienter und einfacher macht (insbesondere den Austausch zwischen Behörden). Um dies zu erreichen, müssen die bestehenden Fachverfahren aller Berliner Ebenen reformiert und geöffnet werden. Rahmenverträge mit Dienstleister:innen, die IT-Lösungen für Fachverfahren anbieten, müssen daraufhin überprüft werden, ob sie die Bereitstellung von Open Data in gebotener Qualität und Granularität ermöglichen (open by design) und ob sie die notwendigen Voraussetzungen für Linked Open Data schaffen.

In den Workshops wurde vielfach von Verwaltung und landeseigenen Betrieben berichtet, dass der Informationsfluss besonders dann stockt, wenn Personen auf Daten von anderen Stellen in Berlin angewiesen sind. Dort zeigt sich immer wieder, dass die angefragten Daten nicht strukturiert und somit auch nicht durchsuch- und auffindbar vorlagen, sondern beispielsweise als Aktenvermerke oder in gedruckter Form. Für die veraktende Person gibt es häufig weder Anreiz noch die notwendigen Werkzeuge, die in einem Vermerk enthaltenen Faktendaten maschinenles- und -findbar abzulegen; es genügt die Ablage in einer Papierakte. Dritte, die diese Informationen für ihre behördliche Arbeit benötigen, müssen diese Faktendaten deswegen händisch in den Akten finden und übertragen – womöglich über alle Bezirksamter hinweg.

Die Verwaltungsmodernisierung darf nicht darauf beschränkt sein, analoge Prozesse nur zu digitalisieren (z.B. von Akte zu e-Akte). Vielmehr muss es auch darum gehen, ein umfassendes Datenmanagement in der Verwaltung aufzubauen, das es ermöglicht, Prozesse neu zu definieren und Aufgaben effizienter zu erledigen. Es wird dabei immer wichtiger werden, mit strukturierten Datensätzen zu arbeiten und diese auch anderen Verwaltungen zur Verfügung zu stellen oder die Verknüpfbarkeit dieser Daten sicherzustellen.² Wenn in der Verwaltung mit Daten gearbeitet wird, müssen diese schon für eine mögliche Veröffentlichung vorbereitet werden (open by default), ohne dass dies für die Ausführenden als unverhältnismäßige Mehrarbeit empfunden wird. Wenn diese Veröffentlichung ohne weiteres Zutun der erfassenden Person geschehen kann, kann dies ein Indikator für eine gelungene IT-Strategie sein.

Wenn die Bereitstellung von Open Data durch die Berliner Verwaltung wie oben beschrieben in die eigene Verwaltungsarbeit integriert und als Teil des Berliner Datenmanagements verstanden wird, wird die Datenqualität steigen. Denn dort, wo die Verwaltung die Daten auch selbst braucht und regelmäßig mit ihnen arbeitet, erfolgt die Qualitätssicherung quasi automatisch. Dann stehen die Chancen hoch, dass auch Dritte (u.a. Wissenschaft, Wirtschaft, Zivilgesellschaft) diese bereitgestellten Daten nutzen und ihr Potenzial noch erweitern.

² Aufgrund der hohen Relevanz der Verwaltungsmodernisierung und des Entwickelns eines modernen Datenmanagements ist es wichtig, die Ansätze für Open Data auch mit den Überlegungen zu Digitalstrategie und Smart-City-Prozess abzugleichen und auszutauschen.

Mehr PDFs und bessere Excellisten reichen nicht: Eine moderne Verwaltung mit komplexen Aufgaben braucht verknüpfte Datensätze (Linked Open Data)

Open Data ist nicht gleich Open Data. Bislang messen wir den Erfolg von Open Data meist, indem wir die Anzahl von bereitgestellten Datensätzen zählen. Ob diese beispielsweise fehlerhafte PDF-Dokumente, veraltete Excellisten oder mit Metadaten versehene CSV-Dateien sind, unterscheiden wir selten. Der wirkliche Erfolg hängt aber maßgeblich auch am Reifegrad der Offenheit.

Seit vielen Jahren stecken wir bei Open Data bei einem Reifegrad zwischen "PDF" und "Excellisten" fest. Wir könnten noch viele Jahre weiter daran arbeiten, beispielsweise von "PDF" zu "barrierefreies PDF" und von "Excellisten" zu "fehlerfreie Excellisten mit Optimierungen" zu kommen. Dies wird auch einen Großteil der Arbeit in Verwaltungen ausmachen. Der Aufwand dazu ist jedoch immens und wird auch bei jedem neuen Datensatz sehr groß bleiben. Die Klagen über den "großen Mehraufwand ohne spürbaren Nutzen" werden daher auch bleiben. Mit dem demografischen Wandel und der digitalen Transformation müssen die Prozesse in der Verwaltung so umgestaltet werden, dass es zu Effizienzgewinnen durch Automatisierungen kommt. Dies ist nur möglich, wenn Datensätze miteinander sprechen lernen, oder mit anderen Worten: durch Linked Open Data³. Datensätze müssen so angelegt werden, dass sie möglichst offen kompatibel, gut beschrieben und eindeutig identifizierbar sind. Nur so können Verknüpfungen automatisch generiert werden.

Ein Beispiel aus den Workshops: Während der Covid-19-Pandemie nutzten viele Datenjournalist:innen tabellarische Daten des Robert-Koch-Instituts (zum Infektionsgeschehen, zum Impfstatus sowie Geodaten aus anderen Quellen für die Zuordnung auf Bundesländer, Städte und Landkreise), um für journalistische Angebote diese Zahlen automatisiert grafisch aufzubereiten und dadurch erfahrbar zu machen (z.B. tägliches Corona-Dashboard mit Deutschlandkarte in der Süddeutschen Zeitung). Die Auswertungen konnten zwar weitgehend skriptgesteuert vorgenommen werden, den verwendeten Skripten musste jedoch händisch zugeordnet werden, in welcher Spalte der Tabelle Informationen welcher Art zu erwarten waren. Immer wenn das RKI seine Tabellen verändert hat (Spaltenüberschriften, neue Spalten, Spaltenreihenfolge), funktionierten die Skripte der Journalist:innen nicht mehr und mussten händisch erneut angepasst werden. Dies führte oft zu hektischen Telefonaten zwischen Redaktionen und RKI am frühen Mor-

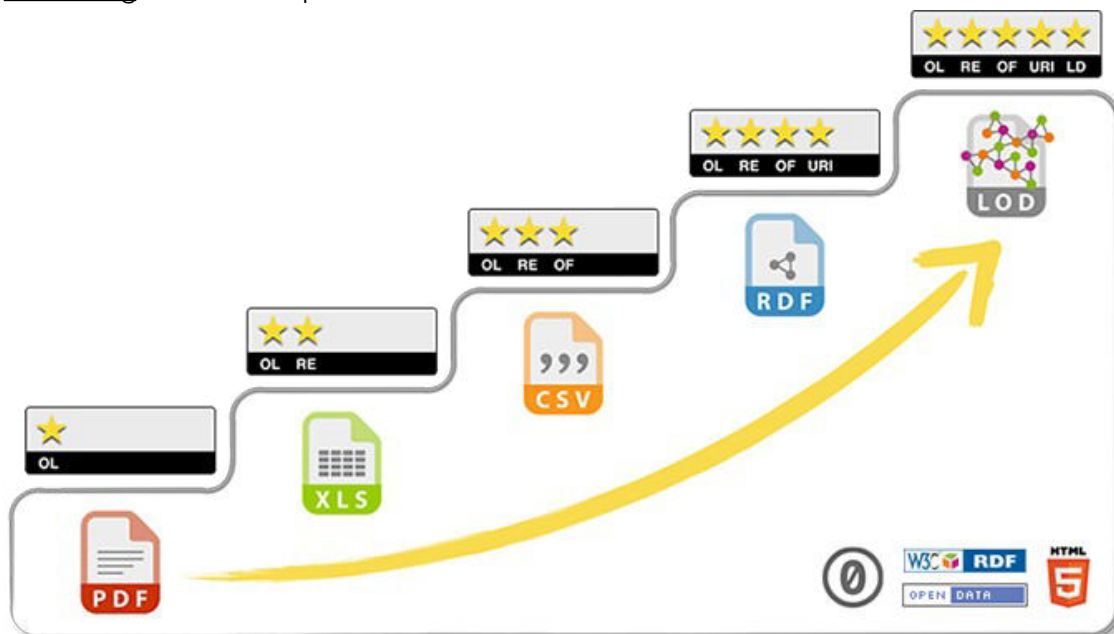
³ Vgl. Definition Linked Open Data bei Wikipedia: https://de.wikipedia.org/wiki/Linked_Open_Data. Tim Berners-Lee, Erfinder des World Wide Web, stellte 2009 in einem TED-Talk seine Zukunftsvision für ein künftiges WWW der Daten vor. Darauf aufbauend wurde das Modell der 5 Sterne für Open Data entwickelt: In den ersten Stufen sind Daten überhaupt im Web verfügbar, weitere Sterne gibt es für Daten in strukturierter Form (Tabellen). Die eigentlich visionäre Leistung Berners-Lees liegt jedoch in den Stufen 4 und 5 dieses Modells. Vier Sterne bedeutet, dass nicht nur eine bloße Tabelle vorliegt, sondern dass ein Datensatz sich selbst beschreibt: „Diese Spalte bildet Datum und Uhrzeit ab, in dieser Spalte befinden sich Adressen in einem definierten Format.“ Erst in der 5-Sterne-Ausbaustufe können verschiedene Datensätze so miteinander in Bezug gebracht – verlinkt – werden.

gen. Wären die Daten in der RKI-Tabelle mit Metadaten selbstbeschreibend kodiert gewesen, hätten Veränderungen in der Tabelle automatisch ausgelesen werden können.

Dieses System mit wohlbeschriebenen Daten erlaubt die Zusammenfassung und Verknüpfung von Datensätzen aus gänzlich unterschiedlichen Quellen. Es ist somit weder die Einrichtung einer autoritativen und zentralen Datenplattform notwendig, noch müssen dafür zwingend die eingesetzten Fachverfahren beispielsweise über alle Berliner Bezirke hinweg vereinheitlicht werden (sie müssen aber reformiert und geöffnet werden). Im Gegenteil erlaubt erst das offene Bekenntnis zu Linked Open Data gleichzeitig eine hohe Zusammenfassbarkeit verschiedener Datenquellen sowie eine weitgehend freie Wahl der eingesetzten IT-Systeme. Diese Herangehensweise dürfte auch die Grundsätze des Föderalismus optimal berücksichtigen. Der geplante Data Hub ist demnach ein wichtiger und willkommener Zwischenschritt, um schon jetzt Daten beispielsweise aus Fachverfahren automatisiert zu exportieren und der Öffentlichkeit bereitzustellen. Wenn gleichzeitig jedoch bereits jetzt die strategischen Entscheidungen für die Bereitstellung von Linked Open Data getroffen werden, können künftig auch Daten aus verschiedensten Quellen berücksichtigt und gemeinsam verwendet werden, die sich aus praktischen Gründen nur mit unverhältnismäßigem Aufwand oder gar überhaupt nicht auf den Nenner eines überall durchgehend genutzten Systems bringen lassen.

Der Fokus sollte daher auf die Ziele Linked Open Data und automatisierte Bereitstellungsprozesse gelegt werden. Je besser das interne Datenmanagement organisiert ist, desto einfacher lassen sich auch komplexere Vorgänge wie Linked Open Data oder datengetriebene automatisierte Entscheidungsmechanismen umsetzen.

Abbildung: 5-Sterne-Open-Data-Modell nach Tim Berners-Lee



Quelle: <https://5stardata.info/de/>

Vom Ziel her denken und nicht verkomplizieren: **Der Berlin Data Hub muss auf Linked Open Data ausgerichtet sein und darf nicht überfrachtet werden**

Das Nutzungsversprechen von Open Data kann nur eingelöst werden, wenn mit einem modernen Datenmanagementsystem gearbeitet wird. Auf dem geplanten Berlin Data Hub⁴ lasten daher schon jetzt hohe Erwartungen. **Der Data Hub ist ein unabdingbarer und wichtiger Baustein für das Berliner Datenmanagement, aber keine "eierlegende Wollmilchsau"**. Er wird nur gut funktionieren, wenn das gesamte Daten-Ökosystem (insb. IT-Infrastruktur, Workflows, Zuständigkeiten, Personal, Ressourcen, politische Führung, technische Umsetzung) gut aufgestellt ist. Es gilt daher, den Data Hub mit realistischen Erwartungen zu flankieren. Drei Prämissen sind bei der Umsetzungen maßgeblich.

Erstens, der Data Hub sollte von Anfang an darauf ausgerichtet sein, verknüpfte Datensätze zu hosten - auch wenn es aktuell davon noch sehr wenige Fälle gibt. Es wäre ein Fehler, den Data Hub als "Cloudspeicher für große Excellisten" zu verzwingen. Der Aufbau von Speicher- und Verbreitungssystemen für Linked Open Data (sog. Triple Stores⁵) muss strategisch im Vordergrund stehen. Die Grundprinzipien von Linked Data erlauben hierbei auch eine dezentrale Architektur mit mehreren, verteilten Triple Stores, die dennoch auch für gemeinsame Abfragen verwendet werden können.

Beispielhaft sei hier das auf Seite 3 beschriebene Problem genannt: Bisher wurden Faktendaten zum Beispiel in Word veraktet und ausgedruckt. Dritte, die diese Daten benötigen, müssen sie händisch aus den Akten suchen. Ein Lösungsansatz wäre hierbei, den bisherigen Pfad der Aktenerstellung mit Office-Produkten zu verlassen und die notwendigen Informationen textbasiert zu erfassen, d.h. Inhalt und Form voneinander zu trennen. Mit einem zu schaffenden Workflow kann dann einerseits automatisiert das angestrebte formatierte Dokument samt aller Notwendigkeiten (Briefkopf, Datum etc) erstellt und als E-Akte veraktet werden. Andererseits können die Informationen parallel zu semantisch angereicherten Daten verarbeitet werden, die in einem Triple Store abgelegt werden. Somit kann bereits die Infrastruktur aufgebaut werden, mit der interne Dritte die so abgelegten Daten effizienter abfragen können, ohne hierfür händisch Papierakten konsultieren zu müssen. Die Entscheidung, welche der so erfassten Daten auch als Open Data bereitgestellt werden sollen, ist danach eine reine Kategorisierungsfrage.

⁴ Ein Konzept zur Entwicklung eines Berliner Data Hub wurde von KPMG 2019 erarbeitet; die [Abschlusspräsentation](#) mit den wichtigsten Ergebnissen ist hier verfügbar.

⁵ Ein Triple Store organisiert Informationen in einer Semantik mit den drei Bestandteilen „Subjekt – Prädikat – Objekt“ – deswegen Tripel. Als Beispiel kann das Wikidata-Datenobjekt für das Land Berlin herangezogen werden: „Berlin – ist Instanz von – Bundesland (Deutschland)“. So können beispielsweise Abfragen in der Form „zeige alle Bundesländer Deutschlands mit der zugehörigen Bevölkerungszahl in aufsteigender Reihenfolge an“ getätigt werden.

In der Vergangenheit wurden an mehreren Stellen bereits sogenannte ETL-Verfahren (Extraktion, Transformation, Laden) etabliert, mit denen die in Fachverfahren erhobenen Daten aus deren Datenbank exportiert und als Tabellen aufbereitet wurden, die danach über ein geeignetes Datenportal bereitgestellt werden. Diese Verfahren kamen bislang stets nur auf die Stufe von 3-Sterne-Daten, d.h. in Tabellen strukturierte Datensätze. Für die Weiterverwendung der Daten ist daher bislang stets eine händische Erstellung oder Anpassung von Auswertungsskripten notwendig. Um auch hier dem Ziel von 5-Sterne-Open-Data gerecht zu werden, ist es unerlässlich, diese ETL-Verfahren dahingehend anzupassen. Das heißt, dass die Semantik der Quelldatenbank durch die Verwendung standardisierter Vokabulare⁶ auch im Zieldatensatz kodiert und als Knowledge Graph⁷ abgelegt wird. Perspektivisch erscheint es auch sinnvoll, bei der Neubeschaffung von IT-Lösungen für Fachverfahren bereits deren interne Datenhaltung als Knowledge Graph zu fordern. Dies würde die bisherigen interimswiseen Extraktionsprozesse weitgehend ablösen, und bereits an der Quelle die Kompatibilität mit dem angestrebten Ziel sichern.

Zweitens, der geplante Data Hub soll kein monolithisches System sein, das alle erdenklichen Bedürfnisse und Anforderungen erfüllen kann. Vielmehr muss es sich um einen Verbund von Diensten und Systemen handeln, der verschiedene Aufgaben übernimmt. Hierbei kann modular vorgegangen werden, d.h. nicht alle Dienste und Systeme müssen am Anfang vorhanden sein. Sie können peu à peu dazukommen. So wächst der Data Hub in seinen Funktionalitäten und wird immer weiter verbessert. Folgende Aufgaben (teils grundverschieden) erscheinen sinnvoll:

- Auffindbarkeit und Katalogisierung
- Triple Store für Linked Data
- Speicherort für nicht-linked Data
- Zeitreihendatenbanken z.B. für Sensorikdaten
- Extraktions-, Transformations- und Ladeverfahren (ETL) für den automatisierten Import aus Datenbanken von Fachverfahren

Drittens, zur praktischen Umsetzung des Data Hub wurden in den Beteiligungsworkshops viele Hinweise und Ideen gegeben. Es ist daher erfolgversprechend, einen regelmäßigen Austausch zwischen Verwaltung und Dritten zu etablieren, um den Data Hub schrittweise aufzubauen und zu verbessern. Anekdotisch sind hier folgende Hinweise genannt:

6 Die notwendige Standardisierungsarbeit für den Löwenanteil dieser Vokabulare ist in den vergangenen Jahrzehnten bereits durch internationale Standardisierungsgruppen wie das World Wide Web Consortium (W3C) durchgeführt worden. Ohne weitere Vorarbeit kann ein großer Teil der notwendigen Semantik durch die Verwendung der Vokabulare auf schema.org umgesetzt werden.

7 Die Gesamtheit der in einem Triple Store abgelegten Informationen wird als Knowledge Graph ("Wissensgraph") bezeichnet. Da die Informationen als Tripel kodiert und somit untereinander vernetzt sind, bilden sie gemeinsam ein Netzwerk bzw. einen Graphen. Verschiedene Knowledge Graphs können miteinander verbunden werden und bilden dann einen größeren gemeinsamen Graphen. Ein gutes Beispiel ist der Knowledge Graph hinter Wikidata. Eine Einführung in das Konzept ist hier zu finden: <https://towardsdatascience.com/a-very-simple-guide-of-knowledge-graph-for-health-data-scientists-bb8cefc1db31>.

- Daten sollten möglichst granular und disaggregiert vorliegen, sie sollten eine zeitliche und räumliche Dimension haben.
- Datenformate sollten standardisiert vorgegeben werden; es sollten international gängige Lizenzen (z.B. Creative Commons) verwendet werden, um auch internationale Nutzung zu ermöglichen. Als Standard sollte die gemeinfrei-ähnliche Veröffentlichung unter CC-0 gewählt werden, zumal es in der Regel keinen urheberrechtlichen Schutz für Fakteninformationen gibt, der eine restriktivere Lizenz erlauben würde (z.B. mit Zwang zur Nennung der Quelle). Es ist wichtig, dass auch nicht tabellenbasierte Datensätze veröffentlicht werden können kann, insbesondere Video- und Bildformate (z.B. OpenParliamentTV).
- Es braucht viel mehr offene Schnittstellen zu den bestehenden Fachverfahren.
- Die Metadaten müssen deutlich verbessert werden; hier liegen massive Fehler, kaum ein Datensatz in Deutschland verfügt über „gute“ Metadaten (einfach, eindeutig, ohne Verwaltungssprache, nachvollziehbar; im Idealfall auch bilingual). Die Daten brauchen eine klare Identifikation sowie eine unverkennbare und unveränderbare ID sowie ein klares Vokabular (stabile, unveränderliche URIs).
- Der Data Hub sollte bilingual (deutsch und englisch) gestaltet werden, um den sowohl in Wissenschaft als auch Berliner Stadtgesellschaft stark vertretenen nicht-deutschen Muttersprachler:innen die Nutzung zu erleichtern.
- Upload- und Downloadprozesse im Data Hub sollten möglichst niedrigschwellig funktionieren (geringe Anforderungen an Datenformate, einfache Schnittstellen, etwaige Unterstützung). Der Download sollte auch ohne Limit funktionieren.
- Der Data Hub sollte die Möglichkeit haben, Datensätze zu kommentieren bzw. Verbesserungen zu melden.
- Zusätzlich zum Datenportal sollte es auch Visualisierungswerkzeuge (z.B. auch Dashboards) geben, mit denen Datensätze auch dargestellt und kontextualisiert werden können. Hier sind Opportunitätskosten zu berücksichtigen. Kann auf eine Linked-Open-Data-Infrastruktur zurückgegriffen werden, können Visualisierungen aufgrund der Verknüpfbarkeit der Daten mit wenig Aufwand und teilweise auch automatisiert erstellt werden. In den Vorstufen von Linked Open Data müssen diese Werkzeuge meist passgerecht angefertigt werden. Passende Dashboards können Verständnis und Akzeptanz für solide IT-Infrastruktur fördern, sollten aber nicht übermäßig mit den Ressourcen für den Aufbau der Linked-Data-Infrastruktur konkurrieren müssen.
- Der Data Hub sollte die Funktion haben, den Datennutzenden Stichworte oder Keywords zu empfehlen.
- Ziel des Data Hubs muss es auch sein, zwischen bestehenden Datenportalen verschiedener Ebene (Kommune, Land, Bund, EU) Daten kreuz-und-quer zu verlinken, Standards sind hier maßgeblich, bzw. als Dateninfrastruktur auch für zukünftige Projekte wie Urban Data Platform oder Digital Twins zu dienen.

Schlaglichter aus den Beteiligungsworkshops

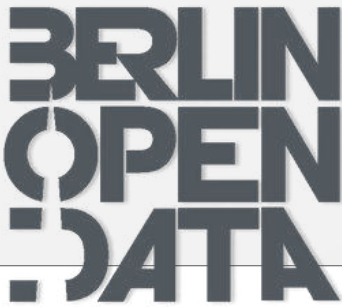
- Der Nutzen von Open Data ist für viele Beteiligte, insbesondere aus Verwaltung und landeseigenen Betrieben, nicht nachvollziehbar. **Der Aufwand ist konkret vorhanden, der Nutzen nur abstrakt in der Zukunft versprochen.** Das Potenzial von Open Data muss besser kommuniziert werden. Eine konsequent durchdachte Open-Data-Strategie führt zum Ausbau der in der Verwaltung ohnehin notwendigen Dateninfrastruktur samt der zugehörigen Prozesse. Diese haben das Potenzial, den Arbeitsalltag in der Verwaltung zu erleichtern und darüber auch den eigenen Nutzen direkt sichtbar zu machen.
- In der Verwaltung, aber auch darüber hinaus (Wissenschaft, Stadtgesellschaft, Wirtschaft), braucht es eine **„Open Data Mentality“**, um das Thema breiter zu verankern und voranzubringen.
- Es braucht **mehr politische Führung** beim Thema Open Data, besonders auf bezirklicher Ebene. Open-Data-Beauftragte haben teilweise unklare Zielstellungen und Mandate, es fehlt zudem an Ressourcen und Weiterbildungsmöglichkeiten. Es bedarf dazu, neben den erforderlichen IT-Ressourcen, auch einen **„Pioniergeist“** im Management, um die Datenöffnung und Datenteilung aktiv voranzutreiben.
- Die **Wirtschaft fordert den notwendigen Kulturwandel** nach dem Motto **„So kann es gehen!“**, dazu gehört auch ein aktives Experimentieren bei Open Data.
- **Datenmapping, Datenkooperation, Zuständigkeit für Öffnung:** Durch die Arbeits- und Organisationsweise der Verwaltung sind Zuweisungen von Zuständigkeit und Verantwortlichkeit zwischen allen Ebenen der behördlichen Verwaltungsstruktur für die Öffnung von Daten unumgänglich. In begrenztem Umfang kann auch die Erkundung vorhandener Datenbestände in der Verwaltung sinnvoll sein. Erkundung und Mapping dürfen jedoch kein Selbstzweck sein, sondern müssen unmittelbar von der Ertüchtigung der Dateninfrastrukturen hin zu Linked Open Data gefolgt werden. Gegebenenfalls ist die direkte Ertüchtigung sogar einem Mapping vorzuziehen (Opportunitätskostenprinzip).
- **Open Data wird nur funktionieren, wenn die bestehenden Fachverfahren Öffnung und Bereitstellung von Daten in die Arbeitsprozesse integrieren.** Fachverfahren auf allen Ebenen müssen daher überprüft und reformiert werden.
- Beispiele für Open Data und Linked Open Data sollten konkrete Anschauungen über **Effizienzgewinne für die Verwaltung, Kostenersparnisse und einfachere Zusammenarbeit** zwischen den Verwaltungen bieten. Akteur:innen aus der Zivilgesellschaft können helfen, diese ersten Beispiele zu erarbeiten.
- Das Berliner Energie- und Klimaschutzprogramm **BEK 2030⁸ ist ein gutes Beispiel für Daten** und sollte mehr für Auswertungen genutzt werden.

⁸ Vgl. Website BEK 2030: <https://www.berlin.de/sen/uvk/klimaschutz/klimaschutz-in-der-umsetzung/>.

- Die **Beantwortung parlamentarischer Anfragen** könnte durch passende Workflows begleitet werden, die eine Nachnutzung der hierfür herangezogenen Informationen in strukturierter Form ermöglichen.
- **Mehr Klasse als Masse:** Aktuell ist die Datenqualität der bereitgestellten Datensätze eher gering. Ziel sollte es sein, auf Qualität relevanter Datensätze zu setzen und in diesem Zuge die Automatisierung der Bereitstellung voranzutreiben, anstatt die Quantität minderwertiger und wenig genutzter Datensätze zu erhöhen.
- Ein **Bot-Tool** könnte dabei unterstützen, höhere Datenqualität zu erreichen, in dem der Bot Vorschläge macht, welche weiteren Maßnahmen es braucht, um die Daten in die nächste Sternekategorie von Linked Open Data zu heben.
- Wenn nur ein Bruchteil jener Ressourcen und Motivation, welche für die Umsetzung der DSGVO zur Verfügung standen, auch bei der Datenöffnung vorhanden wären, wären wir bei Open Data schon weiter. **Länder wie Großbritannien statt entsprechende Institutionen wesentlich und spürbar besser aus.** Der Ressourceneinsatz spiegelt hier auch einen politischen Willen wider.
- Mit der **Open Data Informationsstelle (ODIS)** verfügt Berlin über eine extrem wertvolle und nützliche Ressource, um das Thema in die Umsetzung zu bringen und dabei Wissensgrundlagen zu schaffen. Auch die landeseigenen Betriebe sollten die Möglichkeit erhalten, sich Beratung zur Umsetzung von Open Data einzuholen.
- Es besteht noch **Diskussions- und Erklärungsbedarf bei den Herausforderungen der landeseigenen Betriebe** in den Bereichen Rechtssicherheit, Datenschutz, Haftungsklausel, und der kritischen Infrastruktur - obwohl die meisten dieser Fragen bereits geklärt sind. Hier könnten Beratungsangebote hilfreich sein.

Ausblick

Mit der Beteiligung des Berliner Open-Data-Ökosystems im ersten Halbjahr 2022 ist sowohl eine umfassende Evaluation als auch eine Aufnahme der wichtigsten Bedarfe und Hinweise für die zukünftige Ausrichtung von Open Data in Berlin erfolgreich abgeschlossen. Es gilt nun, auf Basis dieser Ergebnisse und im weiteren Austausch mit alten und neuen Akteur:innen Open Data in Berlin weiterzuentwickeln. Dazu gehören die Formulierung von ambitionierten Zielen und die Konkretisierung von messbaren Handlungsempfehlungen für die nächsten Jahre. Dabei kann das Land Berlin auf die Erfolge der letzten Jahre wie zum Beispiel die Open Data Rechtsverordnung oder die Open Data Informationsstelle aufbauen. Open Data muss in die flächendeckende Umsetzung kommen, technisch und prozessual in die Abläufe integriert werden. Linked Open Data ist dabei der zentrale Baustein für eine digitale, gemeinwohlorientierte und effiziente Verwaltung.



09.06.2022

Inhaltsverzeichnis

1. Einführung durch das Open Data Strategieteam.....	2
2. Vorträge der OKFN, SenWEB und Jens Ohlig.....	4
3. Kleingruppenarbeit mit der Wardley-Mapping Methode	7
4. Ergebnisse der Gruppenarbeit an den “Wardley-Maps”	8
5. Ergebnisse der Kleingruppenarbeit: „Welche Anforderungen haben die Stakeholder:innen der Verwaltung?“	12
6. Aus den einzelnen Gruppen und aus der anschließenden Diskussion im Plenum kamen konkrete Vorschläge:	14
7. Zusammenfassung der Diskussionsergebnisse und Handlungsempfehlungen “Workshop Verwaltung”	17
8. Ausblick	18

1. Einführung durch das Open Data Strategieteam

Am 29. März 2022 fand der 1. Workshop im Rahmen des Partizipationsprozesses für die Berliner Open Data Strategie statt. Dazu luden Sebastian Askar, zentraler Open Data Beauftragter des Landes Berlin und Betül Özdemir, Referentin Open Data von der federführenden Senatsverwaltung für Wirtschaft, Energie und Betriebe (SenWEB) gemeinsam mit der Open Knowledge Foundation (OKF) und der Open Data Informationsstelle (ODIS) 20 Teilnehmer:innen aus der Berliner Verwaltung in das CityLAB Berlin ein.



Gruppenfoto der Teilnehmer:innen des Workshops mit den Stakeholder:innen der Verwaltung vor dem CityLAB Berlin



Sebastian Askar begrüßt die Teilnehmer:innen zum Auftaktworkshop der Berliner Verwaltung der Open Data Strategie

2. Vorträge der OKFN, SenWEB und Jens Ohlig

Im Anschluss an die Begrüßung durch Sebastian Askar von der SenWEB und Henriette Litta von der OKF, referierte Betül Özdemir, Referentin Open Data zum Status Quo in der Berliner Open Data Landschaft und stellte die Ziele und den Partizipationsprozess für die aktuelle Open Data Strategie vor. Alle aktuellen Entwicklungen zur Open Data-Strategie sind sowohl auf der [öffentlichen Informationsseite](#) zur Strategie als auch auf der [zentralen Open Data Seite des Landes](#) Berlin einzusehen.



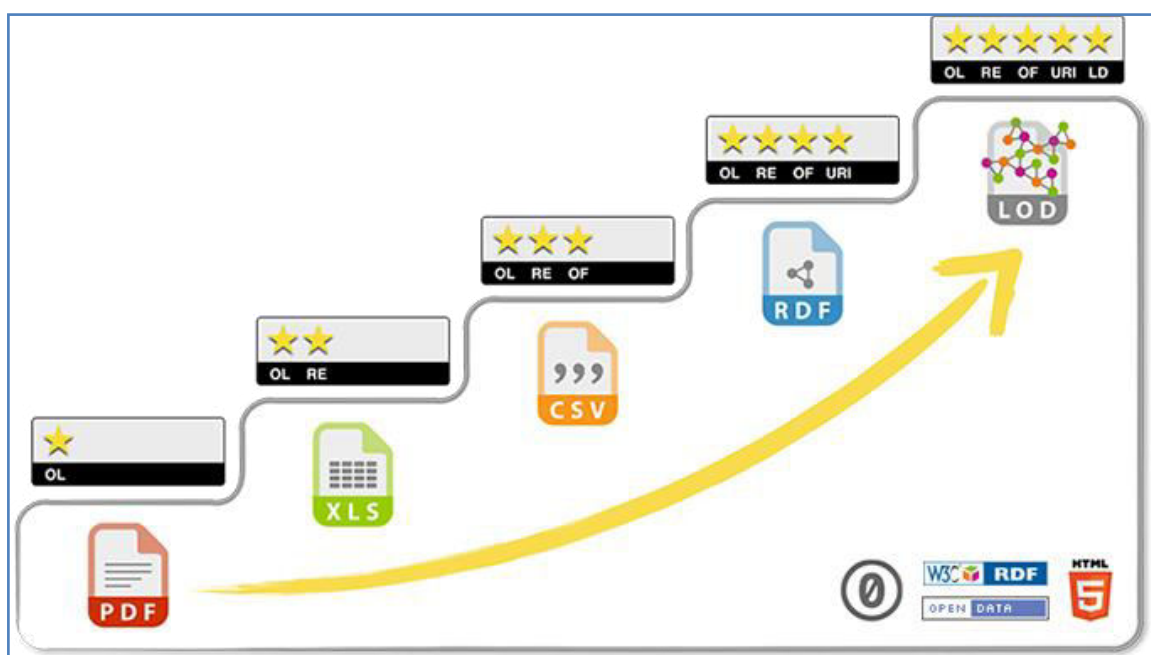
Betül Özdemir, Referentin Open Data, stellt Ziele und Partizipationsprozess der Strategiefindung vor

Danach stellte Walter Palmethofer von der OKF die Ergebnisse der Umfrage der Online-Beteiligung für Open Data vor, die auf meinberlin.de vom 18. Februar bis 18 März 2022 lief, an der sich das ganze Open Data Ökosystem beteiligen konnte.

Bei den Ergebnissen der Umfrage waren die verschiedenen Erwartungshaltungen an Open Data von der Stakeholder:innen Gruppe der Verwaltung und Zivilgesellschaft spannend, ebenso die Einschätzung über die vermuteten Datenverwender*innen des Open Data Portals. So nennen beispielsweise die Verwaltungsmitarbeiter:innen die Wissenschaft als primäre Anwender:in von Open Data, während die Teilnehmer:innen der Umfrage insgesamt die Zivilgesellschaft als hauptsächliche Nutzer:in von Open Data sehen. Die aufbereiteten Ergebnisse der Online-Umfrage liegen [hier](#) ab.

Anschließend ging es weiter mit den Vorträgen und der Datenexperte Jens Ohlig stellte in seinem Impulsvortrag "Linked Open Data und der Weg nach Hause" die Möglichkeiten und Perspektiven „verlinkbarer“ offener Daten vor: In dem Moment, in dem Daten aus verschiedensten Quellen

automatisch miteinander kombinier- und verknüpfbar sind, können Anfragen an diese Datenbestände gestellt werden, die vorher gar nicht denkbar waren. Und obgleich – beispielsweise im Umgang mit dem Bürgeramt – die meisten typischen Anfragen relativ gleichförmig sein können, gibt es immer wieder Fragen, die vielleicht nur für wenige Menschen und nur einmal in deren Leben überhaupt wichtig sind. Für diese Menschen könnte genau diese Frage aber die wichtigste ihres ganzen Lebens sein. Der Aufruf, so Ohlig: „Baut viele Datenendpunkte, die durch Linked Open Data miteinander verknüpft werden können und abfragbar sind – denn nur so können auch diese wichtigen Fragen eines Tages problemlos und schnell lösbar sein.“ Die Präsentationsfolien von Jens Ohlig liegen [hier](#) abrufbar bereit.



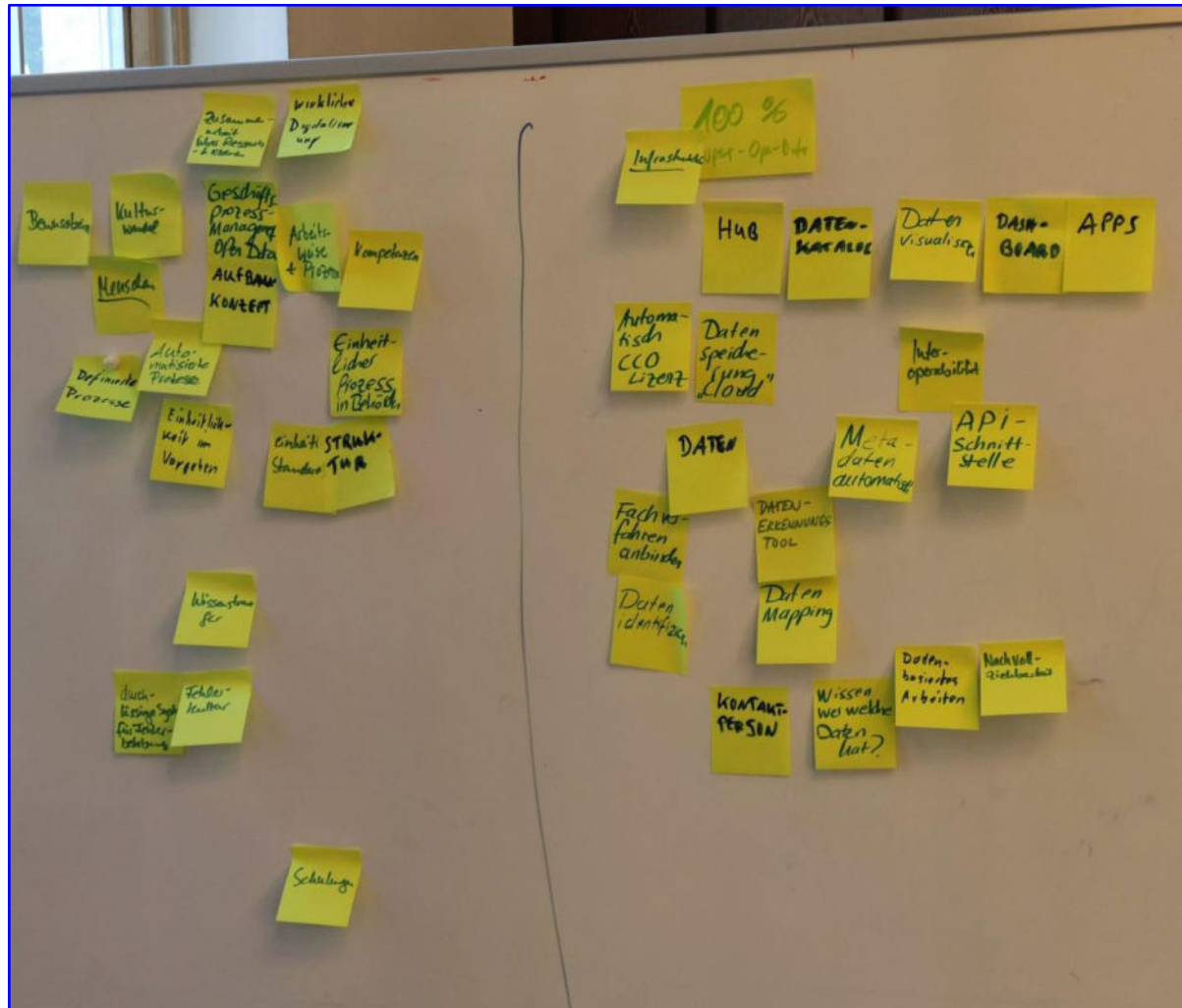
Ausschnitt aus dem Vortrag von Jens Ohlig zum 5* Open Data Modell

3. Kleingruppenarbeit mit der Wardley-Mapping Methode

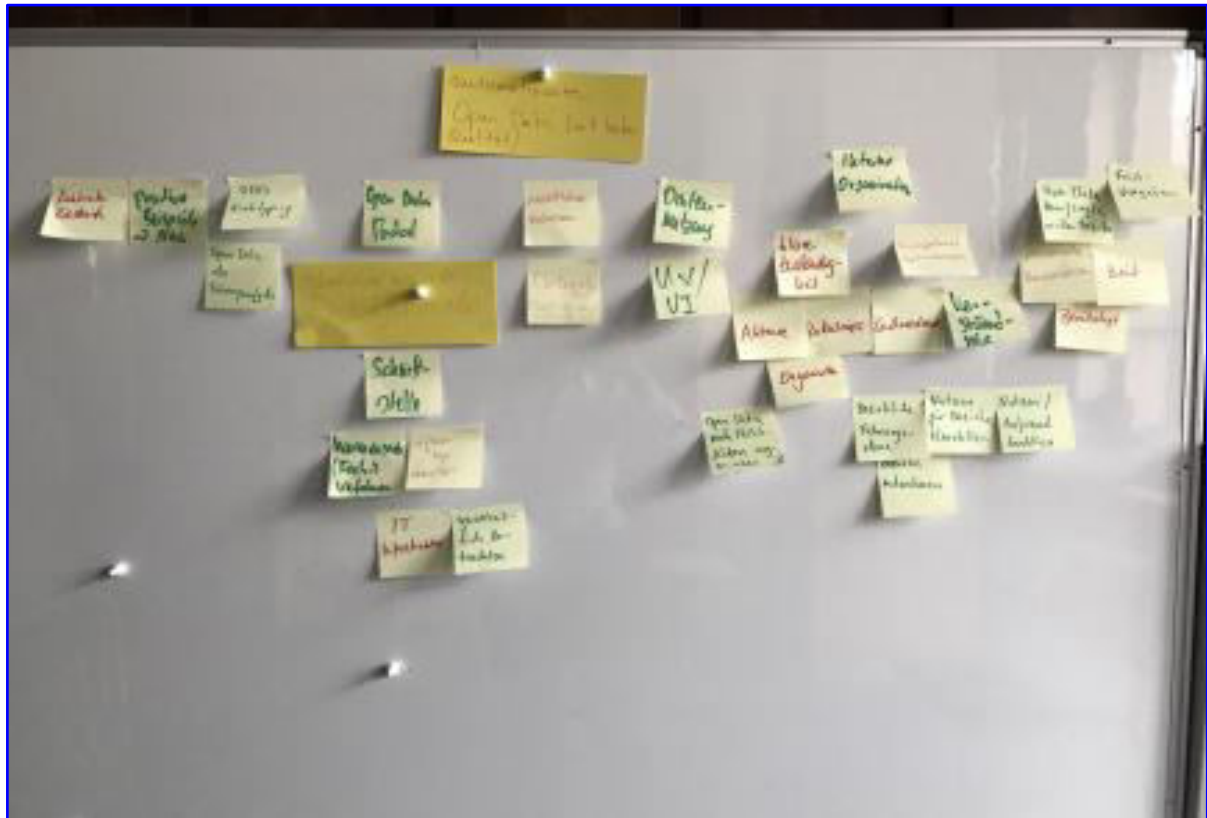
Im darauf folgenden Workshop-Format sammelten die Teilnehmer:innen in moderierten Kleingruppen Ideen, Vorschläge und Maßnahmen aus der Perspektive der Verwaltung für die zukünftige Open Data Strategie. Dazu wurde eine angepasste Wardley-Mapping-Methode genutzt, die in allen Stakeholder:innen Workshops für das Brainstorming verwendet wird. Die Kleingruppen brachten ihre "Map" auf Whiteboards und stellten diese im Anschluss dem Plenum vor. Neben der reinen "Map" als Ergebnis der Kleingruppenarbeit hielten die Moderator:innen auch die fruchtbaren und intensiven Gespräche und Diskussionen fest. Die Ergebnisse der Kleingruppenarbeit werden als Handlungsempfehlungen der Stakeholder:innen Gruppe der Verwaltung in die Open Data Strategie einfließen.

4. Ergebnisse der Gruppenarbeit an den "Wardley-Maps"

Wardley-Map Gruppe 1:



Wardley-Map Gruppe 2:



Wardley-Map Gruppe 3:



Wardley-Map Gruppe 4:



5. Ergebnisse der Kleingruppenarbeit: „Welche Anforderungen haben die Stakeholder:innen der Verwaltung?“

Gruppenübergreifend sind in der Diskussion und Herausarbeitung zentrale Schwerpunkte für die Open Data Strategie aus Sicht der Verwaltung mit Blick auf ein 5* Open Data zusammengekommen:

- ★ Vision Linked Open Data: Es ist notwendig, dass konkrete, einfache Beispiele gezeigt werden müssen, was Linked Open Data bedeutet und was die Vorteile für eine Verwaltung sind, wenn die Daten in dieser „5 Sterne“ Qualität vorhanden wären. Dazu sollten Narrative mit konkreten Nutzen-Stories wie z.B. der Effizienzgewinnung für die Verwaltung, Kosteneinsparnisse und einfachere Zusammenarbeit zwischen Verwaltungen (Bezirks- und Senatsverwaltungen) geschaffen, sowie positive Beispiele aus dem Verwaltungsalltag herausgearbeitet werden.
- ★ Klärung der finanziellen und personellen Ressourcen, Zuständigkeiten & Rollenbilder. Konkret gibt es aktuell viele Open-Data-Beauftragte, die in den Senats- und Bezirksverwaltungen benannt wurden. Jedoch haben diese aber oft gar keine zeitlichen Ressourcen für das Thema Open Data. Die Hauptaussage ist, dass sie diese Aufgaben „nebenbei als Zusatzaufgabe“ machen. Die benannten Open Data Beauftragten sind mit ihren Stellenbezeichnungen auch oft thematisch eher weiter weg von dem Thema Open Data bzw. dem Datenmanagement. In einigen Verwaltungen ist die Aufgabe Open Data bei der Kommunikation bzw. Social Media, etc. angesiedelt, die keine Datenkompetenz haben, sondern lediglich über das Imperia System des Webaufttritts des Landes Berlin die Daten veröffentlichen. Auch die erwarteten Ziele an diese Beauftragten sind teilweise ungeklärt und werden verschieden aufgefasst. Wir brauchen ein klares und einheitliches Rollkonzept für die Open Data Beauftragten mit Aufgaben und Zielformulierungen für ihre tagtägliche Arbeit mit Daten und für die Datenveröffentlichung.
- ★ Nutzen von Open Data mehr zeigen: Durch mangelnde personelle Ressourcen bleibt Open Data vielfach eine Mehrarbeit, deren wahrgenommener Nutzen und Vorteil für die Verwaltung nicht immer greifbar ist. Dies spiegelt sich dann auch bei den Diskussionen über die bezirkliche Ebene wider, wo einerseits die Führung zu Open Data teilweise als fehlend empfunden wurde, andererseits auch hier noch Bedarf zu bestehen scheint, die Vorteile von Open Data auch für die Verwaltung selbst klarer herauszuarbeiten.

- ★ **Rechtliche Rahmenbedingungen:** Der bestehende Rechtsrahmen mit dem §13 EGovG und der Open Data-Verordnung, die seit dem 1.1.2021 in Kraft getreten ist, wurde als eine gute Ausgangslage eingeschätzt, um Open Data in der Berliner Verwaltung umzusetzen. Auszuarbeiten sind noch die notwendigen Voraussetzungen einer möglichst reibungsarmen und effizienten Umsetzung des Prinzips „Open Data by Default“.

- ★ **Erweiterte technische Möglichkeiten durch den Data Hub:** Mehrere Diskussionen drehten sich um die Frage, ob die technische Lösung für alle Bereiche der Berliner Verwaltung mit einheitlichen Prozessen und Workflows vorliegt. Der Wunsch nach einer zentralisierten Plattform, wie dem geplanten Data Hub traf auf anekdotische Erzählungen, dass aktuell vielfach Datensilos parallel entwickelt werden würden und man sich einen gemeinsamen Ort für alle Verwaltungen für die Datenspeicherung wünscht. Hier scheint viel Potenzial zu liegen, den Wunsch nach zentralen Instanzen mit einer Architektur zu verbinden, die von Anfang an auf Verlinkbarkeit auch mit verbleibenden oder gar hinzukommenden dezentralen Datenspeichern setzt.

- ★ **Datenqualität:** Es muss bei den veröffentlichten Daten auch eine Datenqualität gewährleistet werden. Bei den aktuell veröffentlichten Datensätzen gibt es keine einheitlichen Qualitätsstandards, die als Leitfaden oder Empfehlung herangezogen werden können. Es wird gewünscht, dass der zukünftige Data Hub ebenfalls beim Hochladen auch die Qualität der Daten prüft und auch die Qualität verbessern kann, bevor diese für Dritte veröffentlicht werden.

- ★ **Metadaten bzw. Metadatenkatalog:** Es wurde die Frage diskutiert, worin die Vorteile einer separaten Metadatenbeschreibung liegen, die nicht allein in der vorgeschriebenen Anforderung für den Import der Daten aus dem Berliner Open Data Portal z.B. in das Govdata-Portal des Bundes begründet liegen. Wichtig sind auch die Themen Formate und Lizenzen, die möglichst klar und einheitlich sein sollten. Verbunden wurde ebenfalls die Frage, wann Lizenzen mit Anforderungen über CC-0 hinaus, wie z.B. CC BY, CC BY SA überhaupt anwendbar sein sollen, da weitere Einschränkungen in den Nutzungsbedingungen die Wiederverwendbarkeit der Daten einschränken würden.



Blitzlicht aus der Vorstellung der Gruppenarbeit im Plenum

6. Aus den einzelnen Gruppen und aus der anschließenden Diskussion im Plenum kamen konkrete Vorschläge:

- ★ Vorschlag für ein praktisches Anwendungsbeispiel für Linked Open Data: Die Beantwortung parlamentarischer Anfragen könnte durch ein passendes Prozessmanagement Open Data begleitet werden. Das bedeutet, dass die für die Beantwortung der Fragen herangezogenen Informationen auch in strukturierter Form als Excel- oder CSV-Datei aufbereitet werden und als Open Data veröffentlicht werden können.

Dadurch ergäbe sich ein direkter praktischer Nutzen auf mehreren Ebenen: Gleichzeitig wird der Prozess der Beantwortung der Anfragen stromlinienförmiger und die Nachnutzung der Daten als Open Data wird ebenfalls ermöglicht. Kernpunkt ist die praktische Nutzung bereits bestehender Prozesse und dafür die Verbesserung und eine gute Usability der Informationen aus den parlamentarischen Anfragen für das Open Data Portal. Open Data soll nicht als Mehraufwand, sondern als Ergebnis eines sowieso angenehmeren Prozesses entstehen.

- ★ Als akute Herausforderung wurde die Verbesserung des Wissensmanagements in der Verwaltung genannt. Aufgebaute Erfahrung werde vielfach nur mündlich weitergegeben. Selbst Informationen über VAK-Schulungen und Fortbildungen z.B. die von SenWEB über den Crashkurs Open Data würden nicht automatisch an die damit befassten Stellen im Land

Berlin weitergeleitet. Die Informationen zu Weiterbildungen, die es im Bereich Open Data gibt, müssen an die jeweiligen Personen, die zukünftig Open Data Beauftragte ihrer Verwaltung sind, weitergeleitet werden. Es müsste eine Art „Willkommenspaket für neue Open Data Beauftragte“ geben mit Weiterbildungen, die sie besuchen müssen.

- ★ Der Vorteil einer strukturierten Datenerfassung sei meist für die ausführenden Verwaltungsbeschäftigten nicht nachvollziehbar. Daraus folgt, dass die Daten eben nicht nachnutzbar abgelegt werden, weil das für ihren eigenen Zweck ausreiche (z.B. Rechtssicherheit bei Bescheiden gewährleisten; Akte wird ggf. nur hervorgeholt, wenn es zu Klagen kommt). Die so abgelegten Daten sind nicht qualitativ hochwertig, so dass sie für die Nutzung als Open Data nicht ausreichen. Es muss deshalb auch Standards für die Datenerfassung innerhalb des Datenmanagements der Verwaltung geben, nicht für die Veröffentlichung der Daten auf dem Open Data Portal.
- ★ Geodaten als Vorreiter-Datensegment: Es ist historisch bedingt, dass die Daten des Geodatenportals (FIS-Broker) 80% der Daten aus dem Open Data Portal ausmachen, da sie über eine API-Schnittstelle automatisch an das Open Data Portal angebunden sind. Geodaten könnten daher im Open Data Prozess eine stärkere Vorbildrolle spielen. Hier muss jedoch darauf geachtet werden, dass in der Folge nicht die für Geodaten sinnvollen und notwendigen Formate und Paradigmen, auch für alle anderen Datenarten als logisch und gegeben gesehen und angewendet werden.
- ★ Inventur von Daten und Metadaten und deren Verantwortlichkeit.
Es bedarf einer Karte als „Daten Mapping“ mit den Informationen, welche Daten in welchen Behörden vorliegen und wer die jeweiligen Datenverantwortlichen bzw. Open Data Beauftragten sind. Die Frage ist vor allem auch: Wann würde eine solche Inventur überhaupt Sinn ergeben und wann wären die hierfür notwendigen Ressourcen sinnvoller eingesetzt, wenn nicht gleich die notwendige Infrastruktur für die automatisierte Bereitstellung von Daten aufgebaut werden kann?
- ★ Zum Datenportal:
 - im Open Data Portal gibt es leider noch keine standardisierten Datenformate für die zu veröffentlichenden Daten mit Metadaten

- es braucht zudem mehr offene Schnittstellen (zu den bestehenden Fachverfahren), über die direkt im Open Data Portal veröffentlicht werden kann
- es gibt keine bis wenige automatisierte Möglichkeiten, um die Daten der noch nicht herstellerseitig mit Schnittstellen ausgestatteten Fachverfahren auf anderem Wege für die automatische Veröffentlichung vorzubereiten. Es müssen die dafür notwendigen Sach- und Personalressourcen vorhanden sein, um die Fachverfahren mit API-Schnittstellen anzubinden, bzw. die Veröffentlichung aufzubereiten (z.B. [ETL-Skripte](#))
- Alle Fachverfahren müssen auf allen Ebenen harmonisiert werden
- wenn die Verwaltung (egal welche Ebene) bereits mit Daten arbeitet, müssen diese Daten schon von Grund auf für eine mögliche Veröffentlichung auf dem Open Data Portal so gut wie möglich qualitativ vorbereitet werden (open by default), ohne dass dies für die Ausführenden als unangenehmer Mehraufwand empfunden wird
- die IT-Programme, mit denen die Verwaltung arbeitet, müssen die Verarbeitung von Daten und deren Veröffentlichung einfach unterstützen (open by design)

7. Zusammenfassung der Diskussionsergebnisse und Handlungsempfehlungen “Workshop Verwaltung”

- ★ Aus Sicht der teilnehmenden Stakeholder:innen aus der Verwaltung sollten anhand konkreter Berliner Beispiele die Vorteile von Linked Open Data klar aufgezeigt werden, damit die Akzeptanz und Unterstützung der Vision Linked Open Data erhöht wird.
- ★ Der geplante “Berliner Data Hub” soll die bereits vorhandenen Silos aufbrechen und langfristig 5* verlinkte Daten an einem Ort bereitstellen.
- ★ Der Arbeits- und Datenfluss soll verbessert werden. Dazu sollen die Fachverfahren standardisiert und geprüft werden.
- ★ Zwischen den Bezirken und der Senatsverwaltung soll die Datenkooperation, auch bei der Datenerhebung bzw. -inventur erhöht werden. Es müssen auch klare Regelungen bei Daten getroffen werden, die von den Bezirksverwaltungen an die Senatsverwaltungen zur Speicherung oder für die Aufbereitung weitergegeben werden, welche Verwaltung (Bezirks- oder Landesebene) diese Daten als Open Data veröffentlichen soll oder muss. Das könnten zum Beispiel „Auftragsdatenverarbeitungsverträge“ zwischen den Senats- und Bezirksverwaltungen sein.
- ★ Um das Open Data Ökosystem in Berlin angemessen zu unterstützen, braucht es die notwendigen Ressourcen für die Verwaltung. Das bedeutet personelle Ressourcen in Form von Open Data Beauftragten, die mit den notwendigen zeitlichen Ressourcen und der notwendigen Aus- und Fortbildung vorbereitet sind und auch ihre Rollen und Aufgaben kennen.

8. Ausblick

Die Ergebnisse des Workshops werden den Teilnehmer:innen zur Kommentierung zurück gespielt. Im weiteren Prozess folgen Workshops mit den Zielgruppen Zivilgesellschaft, Wirtschaft und Wissenschaft. Alle aktuellen Entwicklungen und Ergebnisse zur Open Data Strategie können auf der [öffentlichen Informationsseite](#) eingesehen werden.

Die Open Knowledge Foundation freut sich auf Anregungen, Ergänzungen oder Kommentierungen: opendataberlin@okfn.de

Für die Open Data Strategiefindung wurde die OKF von der Senatsverwaltung für Wirtschaft, Energie und Betriebe im Jahr 2022 beauftragt, gemeinsam mit der Open Data Informationsstelle den Online-Beteiligungsprozess auf meinberlin.de und die Partizipation des Open Data Ökosystems bestehend aus Verwaltung, Wirtschaft, Wissenschaft und Zivilgesellschaft in Form von Workshops durchzuführen.

Ziel ist es, die Erkenntnisse aus dem breiten Partizipationsprozess in das zukünftige Open Data Berlin Konzept des Landes Berlin einfließen zu lassen.



Ergebnisbericht

zum **Strategieworkshop** mit der
Stakeholder:innen-Gruppe der **Wissenschaft**

am 19. Mai 2022 im CityLAB Berlin

im Rahmen des Partizipationsprozesses für die
Open Data Strategie Berlin 2022

Inhaltsverzeichnis

1. Einführung durch das Open Data-Strategieteam	3
2. Vorträge der OKFN, SenWEB, Jens Ohlig und Heiko Richter	4
3. Kleingruppenarbeit mit der Wardley-Mapping Methode	7
4. Ergebnisse der Gruppenarbeit an den "Wardley-Maps"	7
5. Ergebnisse der Kleingruppenarbeit: „Welche Anforderungen haben die Stakeholder:innen der Wissenschaft?“	11
6. Zusammenfassung der Diskussionsergebnisse und Handlungsempfehlungen „Workshop Wissenschaft“	15
7. Ausblick	16

1. Einführung durch das Open Data-Strategieteam

Zum Abschluss der Workshop Reihe zur Open Data Strategie kamen am **19. Mai 2022** im CityLAB Vertreter:innen der Wissenschaft zusammen, um ihre Visionen, Ziele und Maßnahmen für die neue Open Data Strategie zu evaluieren. Dazu begrüßten die Senatsverwaltung für Wirtschaft, Energie und Betriebe (SenWEB) gemeinsam mit der Open Knowledge Foundation (OKF) und der Open Data Informationsstelle (ODIS) 15 Teilnehmer:innen vor Ort sowie 5 digital zugeschaltete Teilnehmer:innen.



Gruppenfoto der Teilnehmer:innen der Stakeholder:innen Gruppe der Wissenschaft vor dem CityLAB

2. Vorträge der OKFN, SenWEB, Jens Ohlig und Heiko Richter

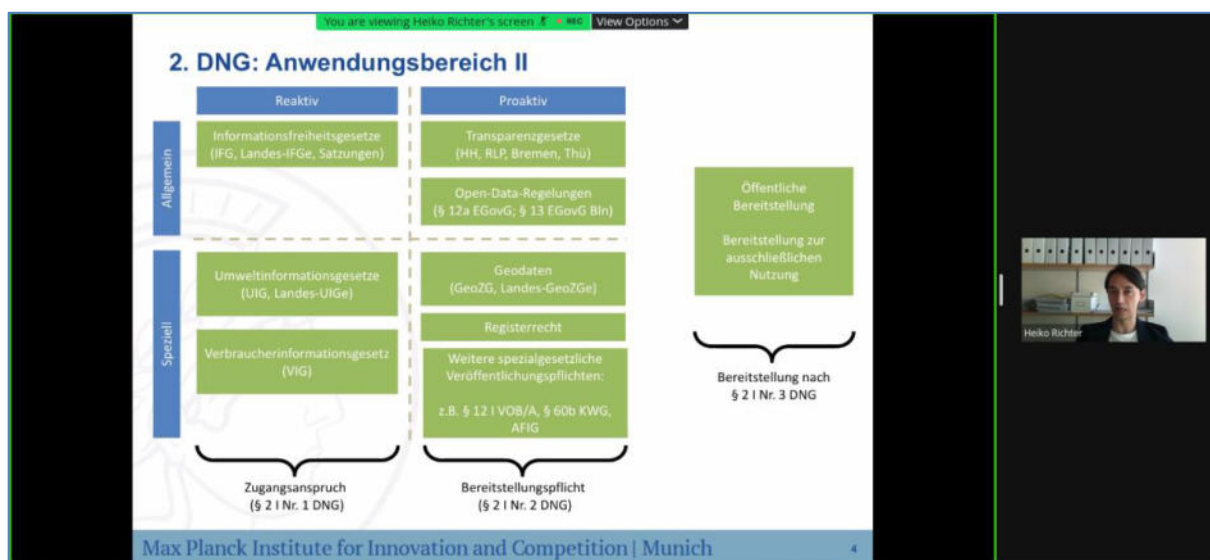
Im Anschluss an die Begrüßung durch Henriette Litta (OKF) und einer Vorstellungsrunde aller Teilnehmer:innen, stellte Betül Özdemir, Referentin Open Data (SenWEB), den Status Quo in der Berliner Open-Data-Landschaft, sowie die Ziele als auch den aktuellen Prozess der Open Data Strategie vor. Alle aktuellen Entwicklungen zur Open Data-Strategie sind sowohl auf der [öffentlichen Informationsseite](#) zur Strategie als auch auf der [zentralen Open Data Seite des Landes](#) Berlin einzusehen.

Eine Einstiegsfrage aus dem Plenum, ob sich die Berliner Behörden an die Open Data Verordnung halten und Daten bereitstellten, bezog sich auf die Selbstwahrnehmung der Berliner Verwaltung im Themenbereich Open Data. Hier sei laut Betül Özdemir bei der Umsetzung durch die einzelnen Behörden noch Luft nach oben. Mit der gemeinsam mit der ODIS entwickelten Broschüre zu den Kerndatensätzen gibt es einen ersten wichtigen Katalog zur Identifikation relevanter und wertvoller Datensätze, die in der Berliner Verwaltung liegen. Diese Kerndatensätze werden Schritt für Schritt veröffentlicht. Allerdings befänden sich eine Vielzahl an Informationen und Daten in einzelnen Fachverfahren, aus denen heraus sie aufgrund der datenschutzrechtlichen Bedenken nicht immer automatisch auf dem Open Data Portal bereitzustellen seien. Hier fehlten laut Knud Möller von BerlinOnline oftmals fertige Export-Schnittstellen (APIs), und die Aufgabe der Anbindung der Fachverfahren läge bei privaten Dienstleistern des Landes Berlin. Ebenfalls aus dem Plenum kam der Hinweis, dass Berlin sich bei seiner Strategie mit anderen deutschen und europäischen Städten abstimmen und Erfahrungen teilen solle. Betül Özdemir bekräftigte, dass Berlin zum Beispiel über die Urban Data Partnership von Fraunhofer Morgenstadt bereits im Bereich Open Data vernetzt und offen für einen weiteren Austausch sei. Zudem wurde aus dem Plenum angemerkt, dass Gewerbedaten interessant seien, diese aber oft explizit für eine Veröffentlichung angefragt werden müssen. Hierzu konnte Frau Özdemir berichten, dass derzeit ein Projekt bei der Senatsverwaltung für Wirtschaft, Energie & Betriebe angestoßen wurde, bei dem Gewerbedaten von der Firma Polyteia in Dashboards präsentiert und demnächst der Öffentlichkeit zur Verfügung gestellt werden. In der weiteren Diskussion stellten die teilnehmenden Wissenschaftler:innen dar, dass gute Datenstandards und passende Datenformate für die wissenschaftliche Arbeit mit Daten sehr relevant seien, um zum Beispiel räumliche oder zeitliche Zusammenhänge zu analysieren.

Danach stellte Walter Palmetshofer (OKF) die Ergebnisse der Onlinebeteiligung auf [meinberlin.de](#) vor. Vom 18. Februar bis 18 März 2022 konnte sich die Berliner Öffentlichkeit an der Umfrage beteiligen. Die aufbereiteten Ergebnisse der Online-Umfrage können [hier](#) eingesehen werden.

Anschließend folgten mit Dr. Heiko Richter und Jens Ohlig zwei spannende Impulsvorträge. Heiko Richter warf einen [kritischen Blick \(Video\)](#) auf das Datennutzungsgesetz (DNG) und diskutierte mit dem Plenum einen Anspruch auf Datenbereitstellung.

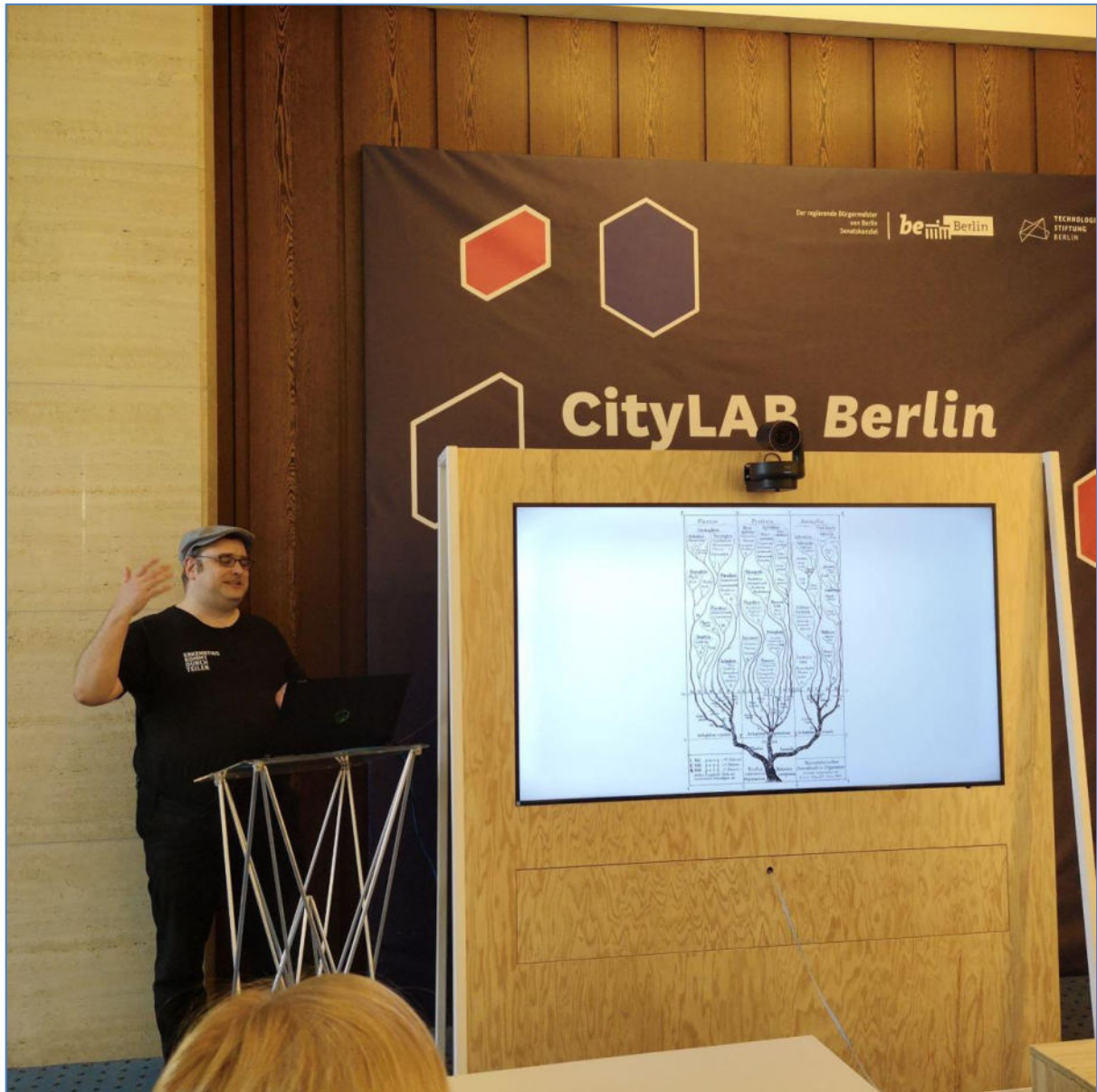
Dabei stellte er zunächst die historische Entwicklung und die Anwendungsbereiche des DNG vor. Demnach gebe es nach seiner Lesart einen grundsätzlichen Anspruch auf Open Data. In der Praxis seien jedoch viele Ausnahmen zu beobachten. Auch ein Blick auf die Open-Data-Landschaft in Deutschland zeige, dass viele Städte und Kommunen ihrer Pflicht nicht nachkommen, offene Daten bereitzustellen. Ein rechtlicher Anspruch auf Open Data könne sogar Anreize für den Staat setzen, den Zugang zu offenen Daten noch weiter zu verengen. Zur Durchsetzung von Open Data als Default fehle es auch an Sanktionsmechanismen. Mit Blick auf die perspektivische Weiterentwicklung der Rechtslage zu Open Data führte Dr. Richter aus, dass im Koalitionsvertrag der derzeitigen Regierungskoalition ein Rechtsanspruch als Ziel gesetzt sei: *“Wir führen einen Rechtsanspruch auf Open Data ein und verbessern die Datenexpertise öffentlicher Stellen.”* (Koalitionsvertrag Mehr Fortschritt wagen, S. 17).



Dr. Heiko Richter präsentiert digital

Als zweiten Impulsvortrag nahm Jens Ohlig die Teilnehmer:innen auf eine Reise durch Indien mit und stellte auf diese Weise die Idee einer Wissensgesellschaft mit 5*-Linked-Open-Data vor. [Das 5-Sterne-Modell](#) wurde 2009 aus [einem Vortrag von Tim Berners-Lee](#) entwickelt und skizziert die Organisation von Wissen in semantisch angereicherten, verlinkbaren offenen Daten. Die bisherige Herangehensweise ende meist bei der Bereitstellung von Datensätzen als einzelne Dateien und teilweise immerhin die Veröffentlichung von Daten über Schnittstellen. Dies erlaube zwar die Nutzung dieser Informationen in bestimmten, auf diesen Zweck zugeschnittenen Anwendungen und

Apps, jedoch erst die Organisation dieses Wissens als Linked Open Data erlaube auch die Beantwortung von Fragen, die sehr spezialisiert und selten vorkämen. Am Beispiel der Wissensdatenbank Wikidata zeigte Jens Ohlig, wie solche hochkomplexen und individuellen Anfragen bereits heute möglich sind und warum das Ziel der 5 Sterne auch für die Verwaltung selbst viele Erleichterungen bringen kann. Die [Präsentationsfolien von Jens Ohlig](#) sind ebenfalls öffentlich zugänglich.



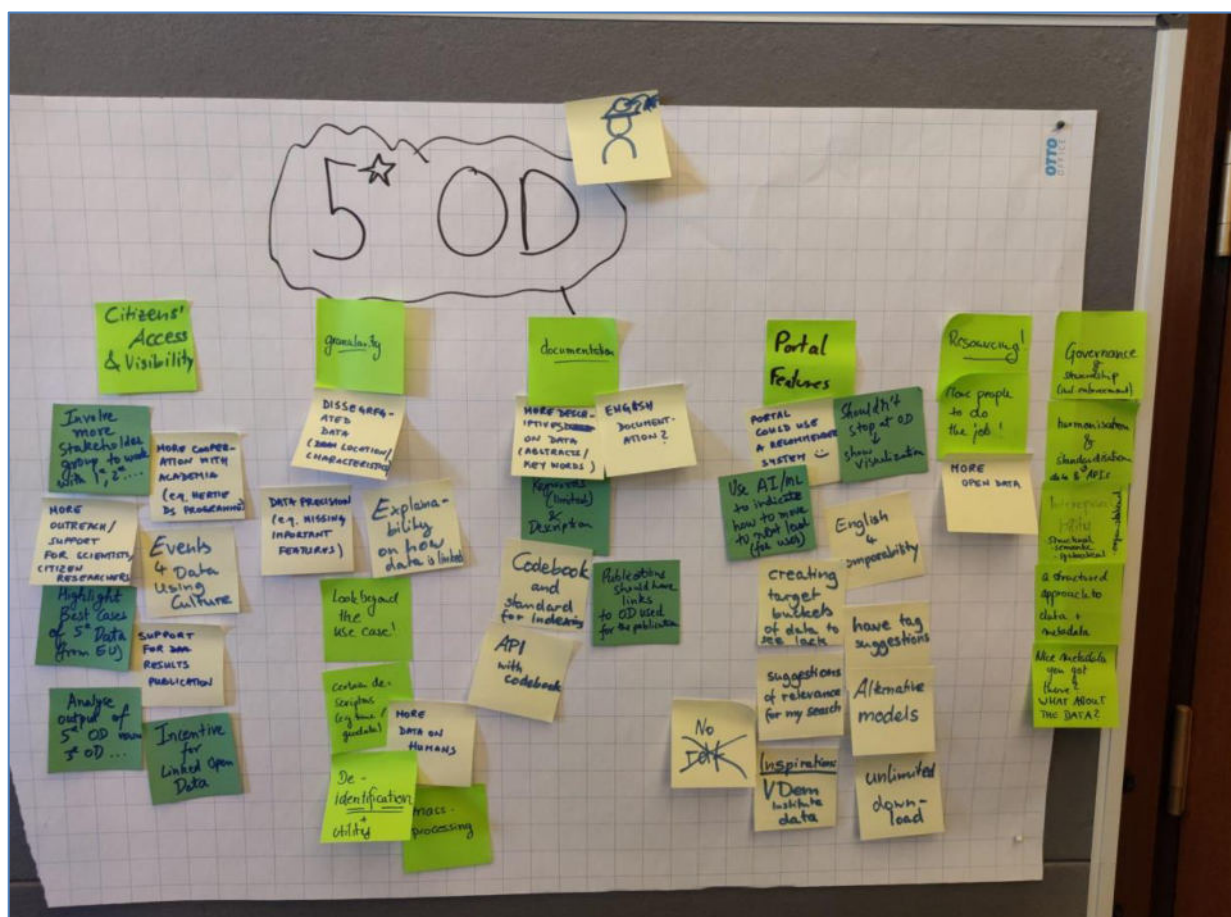
Jens Ohlig und sein leidenschaftliches Plädoyer für Linked Open Data

3. Kleingruppenarbeit mit der Wardley-Mapping Methode

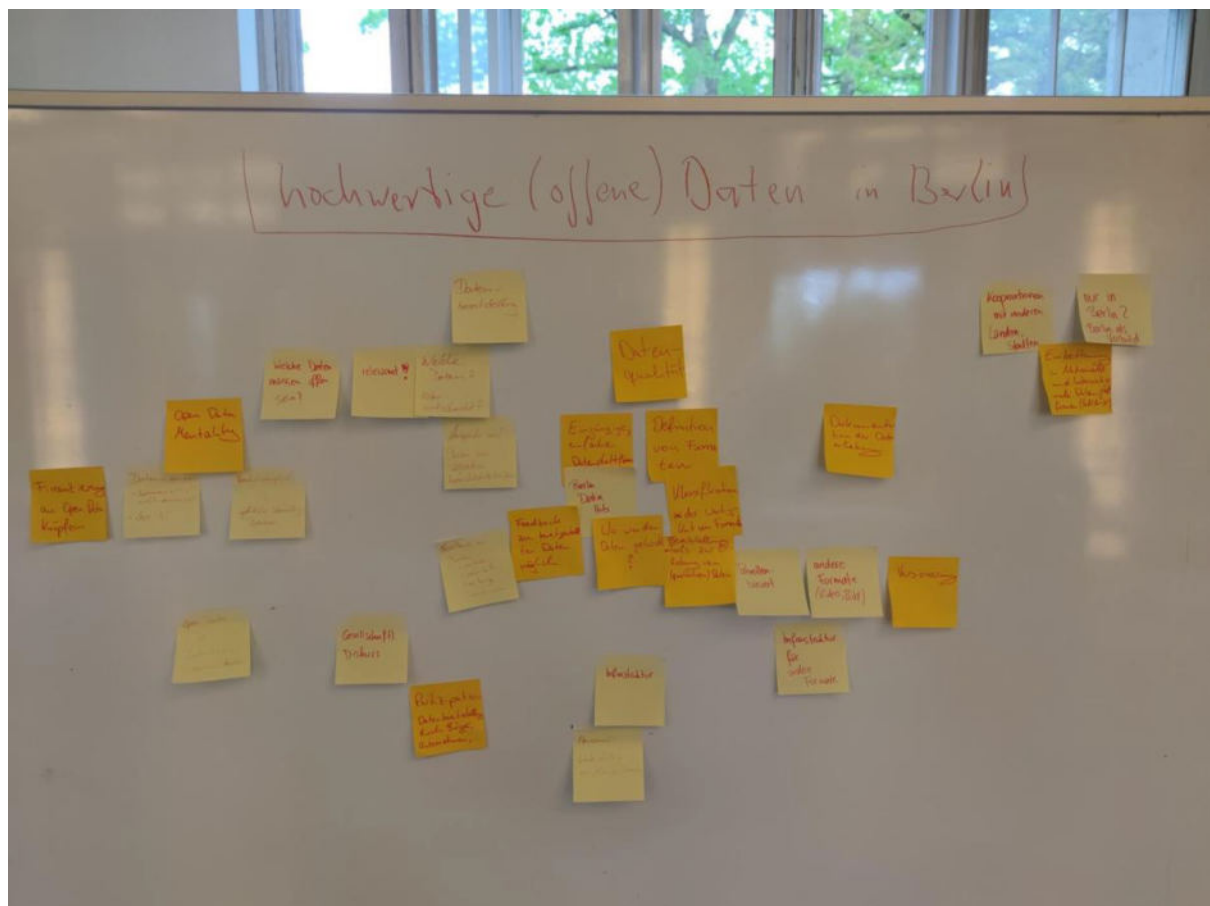
Im darauf folgenden Workshop-Format sammelten die Teilnehmer:innen in moderierten Kleingruppen Ideen, Vorschläge und Maßnahmen für die neue Open Data Strategie. Dazu wurde eine angepasste Wardley-Mapping-Methode genutzt, die in der Strategieentwicklung großen Anklang findet. Die Kleingruppen (drei vor Ort/eine remote digital per Zoom) brachten ihre Map auf Whiteboards bzw. eine Miroboard und stellten diese im Anschluss dem Plenum vor. Neben der reinen Map als Ergebnis der Kleingruppenarbeit hielten die Moderator:innen auch die fruchtbaren und intensiven Gespräche und Diskussionen fest. Die Ergebnisse der Kleingruppenarbeit können so in die Open-Data-Strategie einfließen.

4. Ergebnisse der Gruppenarbeit an den “Wardley-Maps”

Wardley-Map Gruppe Nummer 1:



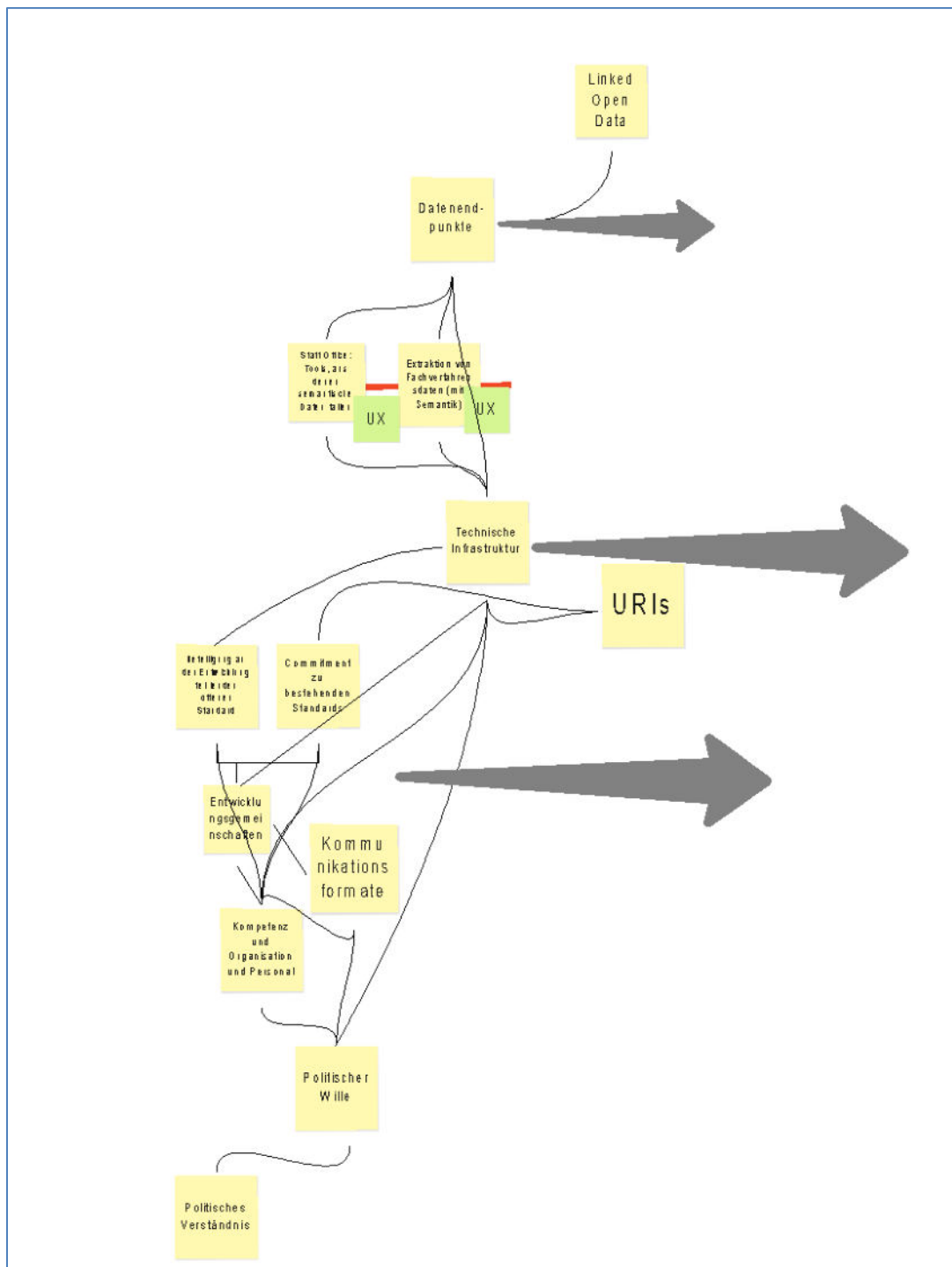
Wardley-Map Gruppe Nummer 2:



Wardley-Map Gruppe Nummer 3:



Wardley-Map Gruppe Nummer 4:



5. Ergebnisse der Kleingruppenarbeit: „Welche Anforderungen haben die Stakeholder:innen der Wissenschaft?“

Gruppenübergreifend sind in der Diskussion und Herausarbeitung zentrale Schwerpunkte für die Open Data Strategie aus Sicht der Wissenschaft mit Blick auf ein 5* Open Data zusammengekommen:

- ★ **Ausreichend Personal sowohl in der Verwaltung selbst als auch bei beratenden Institutionen** wie z.B. die Open Data Informationsstelle (ODIS) sind ein wesentlicher Faktor, um die infrastrukturellen Grundlagen und das Wissen zur Bereitstellung von Open Data voranzutreiben. In anderen Ländern wie zum Beispiel Großbritannien sind entsprechende Institutionen wesentlich und spürbar besser ausgestattet. Der personelle Ressourceneinsatz spiegelt hier auch einen politischen Willen wider.
- ★ **Data Governance**, das heißt in diesem Fall klare Zuständigkeiten und Verantwortlichkeiten im Umgang mit Daten sorgen für eine Standardisierung und Vereinheitlichung von Daten und Zugriffswegen.
- ★ **Interoperabilität** der Daten ist sehr wichtig. Standardvokabulare wie z.B. die vom W3C spezifizierten sorgen für eine weltweite Nutzbarkeit.
<https://www.w3.org/People/Bos/DesignGuide/interoperability.html>
- ★ **Mehrsprachiges Angebot des Open Data Portals**: Das bestehende Berliner Open-Data-Portal sollte bilingual (deutsch und englisch) gestaltet werden, um den sowohl in Wissenschaft als auch in der Berliner Stadtgesellschaft stark vertretenen nicht-deutschen Muttersprachler:innen die Nutzung des Portals zu erleichtern.
- ★ **Dashboards**: Zusätzlich zum Datenportal sollte es auch Visualisierungswerkzeuge geben, mit denen Datensätze auch dargestellt und kontextualisiert werden können.
- ★ **Keyword Suche**: Das Portal sollte die Funktion haben, den Datennutzenden Stichworte oder Key-Words zu empfehlen
- ★ **Downloadmöglichkeiten**: Ein kommendes Datenportal sollte die Möglichkeit haben, Datensätze direkt und ohne Limit herunterzuladen.

- ★ **Veröffentlichung auch von Bildern und Videos** auf dem Open Data Portal: Es ist wichtig, dass auch nicht tabellenbasierte Datensätze veröffentlicht werden können, insbesondere Video- und Bildformate (z.B. OpenParliamentTV)
- ★ **Metadaten/Beschreibungen zu den veröffentlichten Daten auf dem Open Data Portal:** Die genaue Dokumentation bzw. Beschreibung von Daten ist aus Sicht der Wissenschaft sehr wichtig, eine bilinguale Beschreibung (deutsch/englisch) wäre sehr hilfreich. Auch passende Keywords sind hilfreich, um Datensätze zu finden
- ★ **Codebook zu der API-Schnittstelle des Open Data Portals:** API sollten ein Codebook enthalten, es wäre auch zu begrüßen, wenn es zu dem Codebook auch einen standardisierten Index für die Beschreibung gibt.
- ★ **(Wissenschaftliche) Publikationen** sollten Links zu Open Data haben
- ★ **Granularität der Daten:** Für die Wissenschaftler:innen maßgeblich sind auch die Granularität bzw. die Genauigkeit der Daten. Weniger um einzelne Use-Cases sollte die Nutzbarkeit der Daten selbst im Vordergrund einer Open-Data-Strategie liegen.
- ★ **Datenqualität:** Eine möglichst hohe Datenqualität ist entscheidend für den Erfolg von Open Data. Daten sollten so genau wie möglich, dauerhaft auffindbar, maschinenlesbar und vollständig sein.
- ★ **Disaggregierter Zustand von Daten:** Daten sollen in einem disaggregierten Zustand vorliegen. Disaggregation bezeichnet die Aufschlüsselung von statistischen Daten nach bestimmten Merkmalen in unterschiedlichen Einzelgrößen.
- ★ **Bessere Analysen** mit großen Datenmengen: Daten sollten in einer gewissen Größe vorhanden sein, um Analysen machen zu können
- ★ **Räumliche & zeitliche Dimensionen:** Daten sollten zeitliche und räumliche Dimensionen haben

- ★ **Gesellschaftsbezogene Daten als Open Data veröffentlichen:** Es gibt ein “lack of data on humans” Es sollte auch ein stärkerer Fokus auf Daten und Informationen über Menschen und Gesellschaft im Generellen liegen. Derzeit liegt der Fokus stark auf Infrastruktur wie z.B. Verkehr.
- ★ **Die Sichtbarkeit und der Zugang zu Open Data** insbesondere für die **Wissenschaft** sollte verbessert werden.
- ★ Es braucht mehr **Open Data Kooperationen** mit der **Wissenschaft** z.B. Universitäten (Hertie School of Governance), um das Bewußtsein für wissenschaftliche Daten in der Zukunft auch als offene Daten zur Verfügung zu stellen.
- ★ **Events mit der Wissenschaft** für Open Data: Es bedarf an Veranstaltungen, die das Interesse an einer Datennutzungskultur in der Wissenschaft wecken.
- ★ Die Reichweite von **Open-Data-basierenden Forschungsergebnissen** sollte vergrößert werden bzw. sollte es eine Plattform für Forschungsergebnisse als Open Data aber auch für den Austausch von zivilgesellschaftliche Gruppen Open Data geben.
- ★ **Mehr Anreize** für die Wissenschaft zur Arbeit mit offenen Daten schaffen und Mehrwerte von Open Data verdeutlichen: Es braucht mehr Anreize, Open Data (idealerweise 5* Open Data) bereitzustellen oder mit Open Data zu arbeiten
- ★ **Bundesweite Open Data Verfügbarkeit:** Daten dürfen nicht nur für Berlin erhoben und bereitgestellt werden, sondern müssen mindestens bundesweit vorliegen, weil im besonderen Maße Vergleiche zwischen Regionen und Städte im Rahmen von Datenanalysen interessant sind
- ★ **Auch Unternehmensdaten als Open Data fordern:** Gerade Unternehmensdaten z.B. aus dem Bundesanzeiger sind interessant und veröffentlichungspflichtig. Man muss aber derzeit für den Zugang bezahlen. Eine zur Verfügung Stellung von Unternehmensdaten als Open Data (kostenfrei und zur freien Weiterverwendung) wäre wichtig.

- ★ Mehr **Partizipation und Abfragen**, welche Daten der Verwaltung veröffentlicht werden sollen: Über Fragen, welche Daten veröffentlicht werden bzw. relevant für die Allgemeinheit sind, braucht es mehr gesellschaftlichen Diskurs und dauerhafte partizipatorische Ansätze

- ★ „**Open Data Mentality**“: In der Verwaltung, aber auch darüber hinaus (Wissenschaft, Stadtgesellschaft, Wirtschaft), braucht es eine “Open Data Mentality”, um das Thema breiter bekannt zu machen, um sowohl die Vorteile der Nutzung von Daten Dritter zu verdeutlichen und auch die zur Verfügung Stellung von eigenen Daten als Open Data zu fördern.

- ★ **Datenschutzrechtliche Aufklärung**: Immer wieder, wenn wir über Open Data sprechen, kommt es früher oder später zu einer Diskussion um das Thema Datenschutz: dieses Bedürfnis, beide Themen zusammen zu diskutieren und auch voneinander abzugrenzen, sollte mitgedacht werden; dafür braucht es unter Umständen Formate, aber auch die politische Aufhängung von Datenschutz und Open Data Beauftragung (Querverbindungen, Stärke des Mandats) ist Ausdruck dieses Spannungsverhältnisses

6. Zusammenfassung der Diskussionsergebnisse und Handelsempfehlungen „Workshop Wissenschaft“

- ★ 5 Sterne Open Data sollte als ein langfristiges Ziel gesehen werden, um sowohl die Qualität zu steigern als auch die Kontextualisierung von Daten z.B. für die Wissenschaft zu gewährleisten.
- ★ Nutzen strukturierter Daten behördenintern viel mehr in den Vordergrund stellen, damit Open Data nicht nur als eine aufwändige Zuarbeit für externe Dritte verstanden wird.
- ★ Wissenschaft soll auch Links zu Open Data in Publikationen aufnehmen und auch die Reichweite der Open Data Forschungsergebnisse erhöhen.
- ★ Qualität der Daten und die Granularität muss erhöht werden, damit die Wissenschaft mit den Daten auf dem Open Data Portal arbeiten kann.
- ★ Mehr Ressourcen für Wissens- und Kompetenzaufbau in der öffentlichen Hand aber auch in der Wissenschaftslandschaft zu Nutzen und Mehrwerten von Open Data und der Arbeit mit Daten.

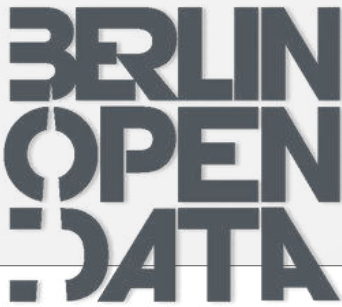
7. Ausblick

Mit dem Wissenschaftsworkshop ist die Workshopreihe zur Open Data Strategie beendet, aber nicht die Möglichkeit der Beteiligung. Die Ergebnisse des Workshops werden den Teilnehmer:innen zur Kommentierung zurück gespielt. Alle aktuellen Entwicklungen und Ergebnisse zur Open Data Strategie können auf der [öffentlichen Informationsseite](#) eingesehen werden.

Die Open Knowledge Foundation freut sich auf **Anregungen, Ergänzungen oder Kommentierungen:** opendataberlin@okfn.de

Für die Open Data Strategiefindung wurde die OKF von der Senatsverwaltung für Wirtschaft, Energie und Betriebe im Jahr 2022 beauftragt, gemeinsam mit der Open Data Informationsstelle den Online-Beteiligungsprozess auf www.meinberlin.de und die Partizipation des Open Data Ökosystems bestehend aus Verwaltung, Wirtschaft, Wissenschaft und Zivilgesellschaft in Form von Workshops durchzuführen.

Ziel ist es, die Erkenntnisse aus dem breiten Partizipationsprozess in das zukünftige Open Data Berlin Konzept des Landes Berlin einfließen zu lassen.



09.06.2022

Inhaltsverzeichnis

1. Einführung durch das Open Data Strategieteam.....	3
2. Vorträge der OKFN, SenWEB und Polyteia	3
3. Diskussion zu Vorträgen.....	5
4. Kleingruppenarbeit mit der Wardley-Mapping Methode	5
5. Die Ergebnisse des Kleingruppenarbeit als Wardley-Maps	6
6. Ergebnisse der Kleingruppenarbeit: „Welche Anforderungen haben die Stakeholder:innen der Wirtschaft?“	9
7. Zusammenfassung der Diskussionsergebnisse und Handelsempfehlungen “Workshop Wirtschaft”	13
8. Ausblick	14

1. Einführung durch das Open Data Strategieteam

Nach dem ersten Workshop im Rahmen des Open Data Partizipationsprozesses mit der Stakeholder:innen Gruppe der Berliner Verwaltung fand am **28. April 2022** der Strategieworkshop mit der Stakeholder:innen Gruppe der Wirtschaft statt. Zu diesem Workshop luden die Senatsverwaltung für Wirtschaft, Energie und Betriebe (SenWEB) gemeinsam mit der Open Knowledge Foundation (OKF) und der Open Data Informationsstelle (ODIS) 20 Teilnehmer:innen aus der Wirtschaft in die DB mindbox, um ihre Visionen, Ziele und Maßnahmen für die neue Open Data Strategie zu evaluieren.



Gruppenfoto der Teilnehmer:innen der Stakeholder:innen Gruppe der Wirtschaft vor der DB mindbox an der Spree

2. Vorträge der OKFN, SenWEB und Polyteia

Im Anschluss an die Begrüßung durch Henriette Litta, Geschäftsführerin der OKF, referierte Betül Özdemir, Referentin Open Data der Senatsverwaltung für Wirtschaft, Energie und Betriebe zum Status Quo der letzten 10 Jahre Open Data in Berlin und stellte die Ziele und den Partizipationsprozess der aktuellen Strategiefindung vor. Alle aktuellen Entwicklungen zur Open Data Strategie sind sowohl auf der [öffentlichen Informationsseite zur Strategie](#) als auch auf der [zentralen Open Data Seite des Landes Berlin](#) einzusehen.

Anschließend stellte Stefan Kaufmann von der OKF die Ergebnisse der Umfrage der Online-Beteiligung für Open Data vor, die auf meinberlin.de vom 18. Februar bis 18 März 2022 lief, an der sich das ganze Open Data Ökosystem beteiligen konnte. Die aufbereiteten Ergebnisse der Online-Umfrage können [hier](#) eingesehen werden.

Im weiteren Programmverlauf gab es einen Best Practice Vortrag von Svenja Stäbler und Cord Maximilian Siebke von der Polyteia GmbH, die ihre Perspektiven als Wirtschaftsakteur:innen für die notwendigen technischen Voraussetzungen für Datenbereitstellende und Datennutzer:innen anhand eines Modellsystems von Polyteia vorstellten. Außerdem präsentierte die Polyteia GmbH beispielhaft ein Open Data Projekt, in dem sie im Rahmen eines Auftrages für die Senatsverwaltung für Wirtschaft Energie und Betriebe die Daten der Gewerbedatenbank mit Dashboards visualisiert haben. Die [Präsentationsfolien von Polyteia](#) sind ebenfalls öffentlich zugänglich.



Polyteia stellt das Projekt mit den Gewerbedaten des Landes Berlin vor

3. Diskussion zu Vorträgen

Anschließend diskutierten die Teilnehmer:innen die Frage, wie bei kritischen Infrastrukturen, die etwa bei den landeseigenen Betrieben liegen, Fragen des Datenschutzes oder Datensicherheit vor der Veröffentlichung von Daten gewährleistet werden können. Hier wurden von den Teilnehmer:innen unterschiedliche Sichtweisen auf die Herangehensweise auf diese Problematik diskutiert. Dabei ist auch deutlich geworden: Es fehlt noch immer das Bewusstsein für die Potentiale durch die Datenbereitstellung, die notwendige Rechtssicherheit bei der Verfügungstellung von Daten und die positiven Beispiele für die Datennutzung aus Sicht der Wirtschaft.

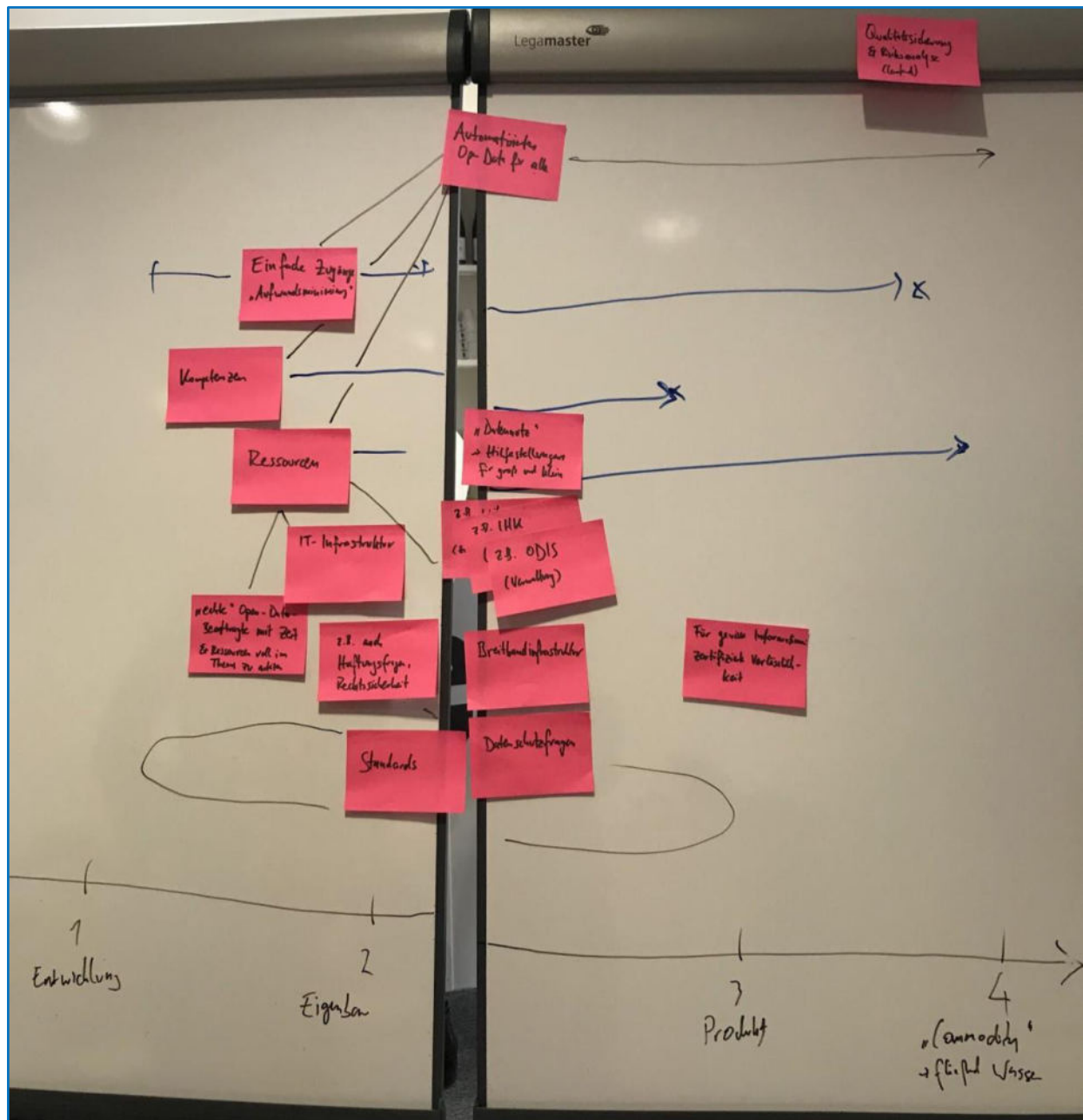
4. Kleingruppenarbeit mit der Wardley-Mapping Methode

Anschließend teilten sich die Teilnehmenden in Kleingruppen auf. Im darauffolgenden Workshop-Format sammelten die Teilnehmenden in moderierten Kleingruppen Ideen, Vorschläge und Maßnahmen aus der Perspektive der Akteure der Wirtschaft für die neue Open Data Berlin Strategie.

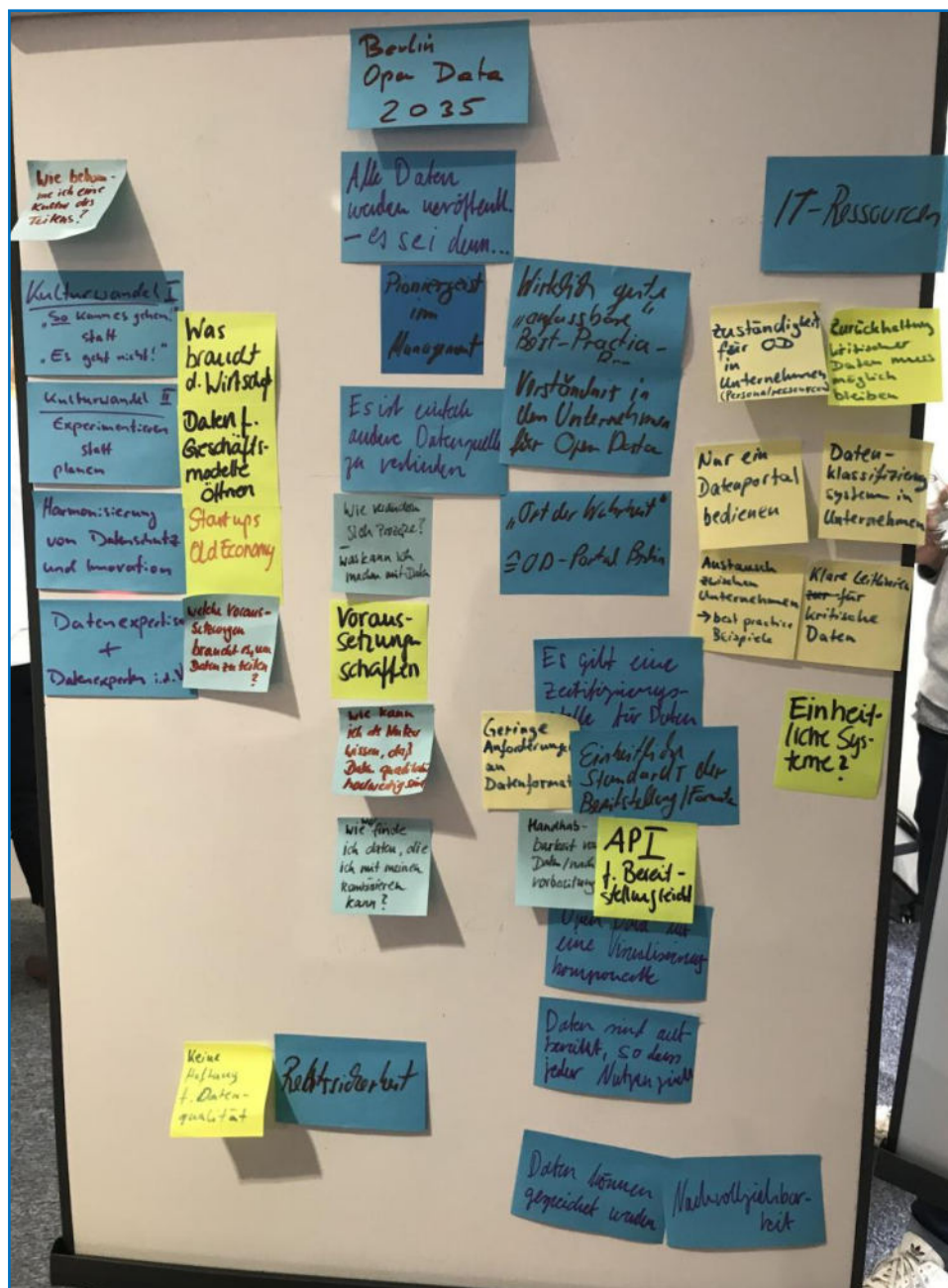
Dazu wurde eine angepasste **Wardley-Mapping-Methode** genutzt, die bei allen Workshops des Partizipationsprozesses verwendet wurde. Die Kleingruppen brachten ihre Vision für die Zukunft von Open Data in Berlin mit Maßnahmenvorschlägen auf Whiteboards und stellten diese im Anschluss dem Plenum vor. Neben der reinen "Wardley Map" als Ergebnis der Kleingruppenarbeit hielten die Moderator:innen auch die fruchtbaren und intensiven Gespräche und Diskussionen, die in den Kleingruppen entstanden sind, fest. Die Ergebnisse der Kleingruppenarbeit werden als Handlungsempfehlung der Stakeholder:innen Gruppe der Wirtschaft in die Open Data Strategie Berlin einfließen.

5. Die Ergebnisse des Kleingruppenarbeit als Wardley-Maps

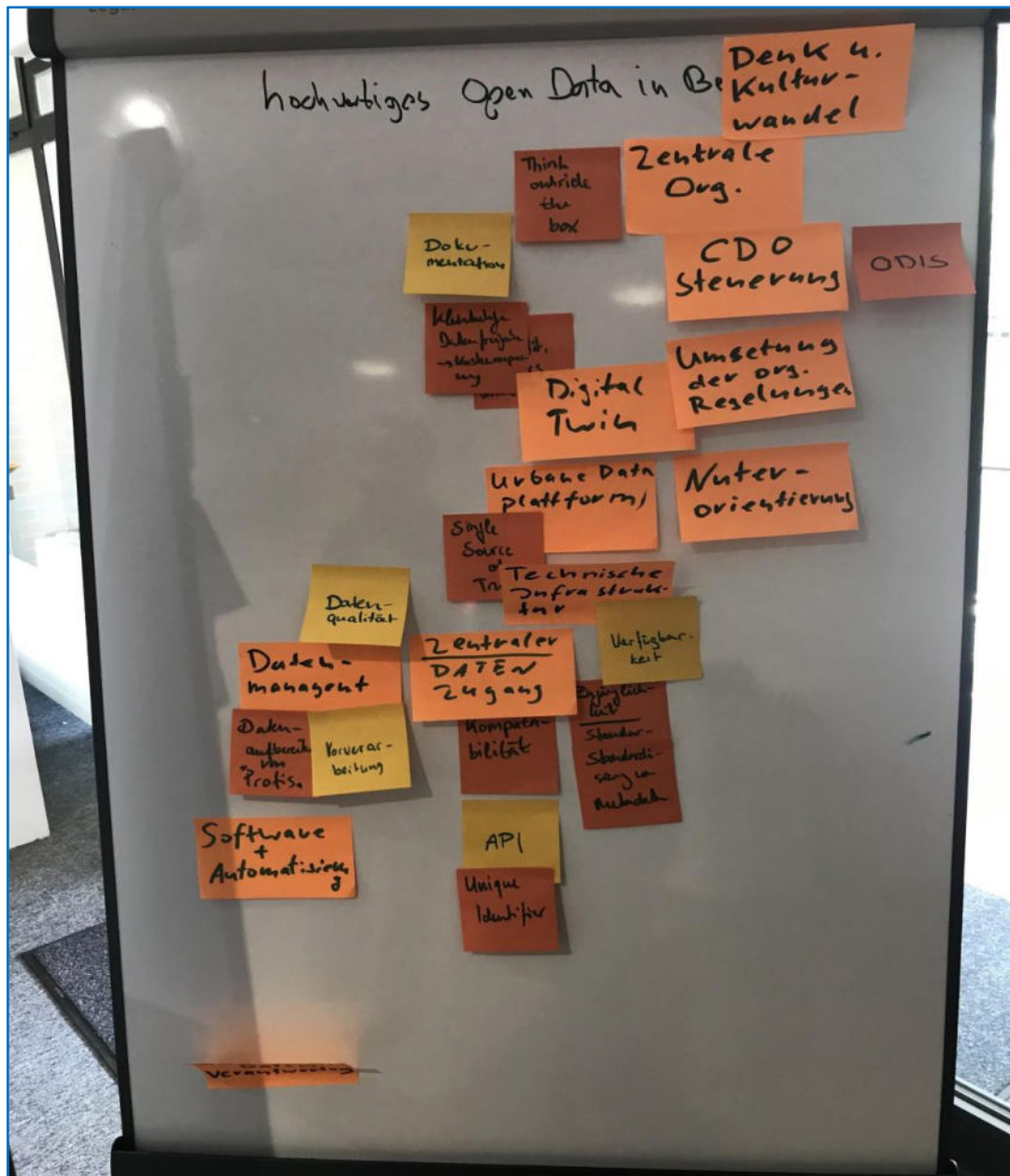
Wardley-Map Gruppe Nummer 1:



Wardley-Map Gruppe Nummer 2:



Wardley-Map Gruppe Nummer 3:



6. Ergebnisse der Kleingruppenarbeit: „Welche Anforderungen haben die Stakeholder:innen der Wirtschaft?“

Gruppenübergreifend sind in der Diskussion und Herausarbeitung zentrale Schwerpunkte für die Open Data Strategie aus Sicht der Wirtschaft mit Blick auf ein 5* Open Data zusammengekommen:

★ **Fragestellung der Finanzierung der Zurverfügungstellung von Daten als Open Data bei**

„landeseigenen Betrieben“: Wie werden Aufwände finanziert, die bei landeseigenen Betrieben für die Bereitstellung von Daten entstehen, die sie erst erheben oder zusätzlich für Open Data aufbereiten müssen? Müssen die Aufwände über Gebühren refinanziert werden oder wird das vom Land Berlin finanziert?

- Zusammenfassend: Open Data bei landeseigenen Betrieben wird als ein Mehraufwand gesehen, der nicht im Sinne von „sowieso gemachter Digitalisierungsvorhaben“ verstanden wird. Der Aufwand für Open Data wird als eine zusätzliche Arbeit gesehen, deren Mehrwert nicht direkt erkennbar ist. Aufwände müssen in Form von finanziellen Mitteln abgegolten werden. Es wird gewünscht, dass die neue Open Data Berlin Strategie auch eine Handlungsempfehlung für die Umsetzung von Open Data bei landeseigenen Betrieben gibt.
- Es fehlt an einer guten „Story“, bei der ein mittelständisches Unternehmen mit Open Data zusätzlichen Umsatz generiert hat („Potenzial nicht griffig“).
- Best Practice Beispiel mit „BIM für einen Gebäudeabbruch oder eine Modernisierung“. Liegen die BIM-Daten vor, kann eine Auftragnehmer:in ein werthaltigeres Angebot erstellen. Das wäre auch eine Abfrage, die im Lebenszyklus eines Gebäudes vielleicht nur einmal gemacht wird → sind die Voraussetzungen für die Zugänglichkeit dieser Daten jedoch vorhanden (LOD), wird dieser Schritt erheblich vereinfacht.

★ **Personelle Ressourcen in Form von Open Data Beauftragten:** Es besteht die Frage, ob die Benennung von Open-Data-Beauftragten wirklich einen Schritt nach vorne für die Veröffentlichung von Daten bedeutet, oder ob hiermit vor allem der Vollzug einer Handlung signalisiert werden soll, während die dafür benannten Personen teilweise gar keine Zeit für diese Aufgaben haben oder sich gar nicht einarbeiten können.

- ★ **„Haftung als psychologischer Faktor“:** Insbesondere Personen mit juristischem Hintergrund denken gerne zuerst an die Abwehr von Haftungs- und Rechtsfragen bevor sie Daten als Open Data veröffentlichen. Die Frage nach der Rechtssicherheit und Haftung bei der Datenbereitstellung spielt in den Unternehmen noch immer eine große Rolle, obwohl jene bereits geklärt wäre. Hier herrscht Informationsbedarf.
Hier könnten **Handbücher und Handreichungen** hilfreich sein, die hier auf Standardvorgehen und die Ausnahmetatbestände der Open Data Verordnung verweisen und anhand von Beispielen die „Open Data Fähigkeit“ von Daten veranschaulichen.
- ★ **Grundverständnis für das Potenzial von Open Data** ist auch in kleinen Unternehmen notwendig: Was haben wir an Daten? Was können wir mit Daten machen? Hier wäre auch eine Empfehlung zur Nutzung von Daten für KMUs mit Best Practice Beispielen, wie mit Daten zusätzlicher Umsatz generiert werden kann, notwendig.
- ★ **Open Data ist Führungsaufgabe:** Für mehr Open Data in Unternehmen braucht es, neben den erforderlichen IT-Ressourcen, auch einen „Pioniergeist“ im Management bzw. In der Führungskultur, um die Datenöffnung und Datenteilung proaktiv voranzutreiben. Es wird in der Wirtschaft weniger als „ein technisches Thema“ gesehen, sondern mehr als ein Thema für das Management. Es braucht klare Leitlinien vom Management, um den „Mindset“ für Open Data in Unternehmen zu öffnen.
- ★ **Technische Anforderungen an die Datenplattform für Open Data**, die von der Wirtschaft genutzt wird: Es sollte möglich sein, dass Daten auf dem Open Data Portal hochgeladen werden, und nicht mehr nur als eine Metadaten-Plattform genutzt wird, die lediglich auf die Daten hinweist. Der Upload-Prozess der Daten auf dem Open Data Portal sollte einfach und einheitlich sein. Dafür sind geringe Anforderungen für Datenformate bzw. deren Anpassung und eine vereinfachte Klassifizierung von Daten notwendig. Es wird eine Datenplattform mit einfacher Schnittstelle zu externen Datenplattformen gewünscht für eine vereinfachte automatische Bereitstellung von Daten mit „wenigen zusätzlichen Handgriffen“. Dazu wurde die Metapher der Wasserleitung (BWB und BSR) verwendet: man dreht den Hahn auf, das Wasser fließt in der gewünschten Qualität, ohne dass man als Person lange darüber nachdenken bzw. die Wasserqualität überprüfen muss.

- ★ **Personelle Zuständigkeiten klären:** Es muss in Unternehmen analog wie in der Verwaltung geklärt werden, wer für Open Data zuständig ist und bei Fragen zu Daten und Open Data unterstützen kann.

- ★ **Netzwerken zu Open Data in der Wirtschaft:** Der Austausch mit anderen Unternehmen, auch mit Startups, z.B. zur Datenveredelung oder Erstellung von neuen Geschäftsmodellen wäre begrüßenswert. Es wird empfohlen, dass es Veranstaltungen für Unternehmen, KMUs und Startups gibt, um sich über Open Data und die Datennutzung der Daten der Verwaltung zu informieren. Welche Daten der Verwaltung gibt es und was kann daraus an Geschäftsmodellen generiert werden? Welche Erkenntnisse wünscht sich die Berliner Verwaltung von der Wirtschaft aus diesen Daten?

- ★ **„Open Data vs. DSGVO“.** Open Data ist und war schon immer eine Prioritätenfrage in Unternehmen und in der Verwaltung: Wenn nur ein Bruchteil jener Ressourcen und Motivation, welche für die Umsetzung der DSGVO zur Verfügung gestellt wurde, auch bei der Datenöffnung vorhanden gewesen wäre, wären wir bei Open Data schon viel weiter.

- ★ **Rechtliche Rahmenbedingung von der Bundes- und Landesebene:** Der Rechtsrahmen und die Notwendigkeit der Umsetzung von Richtlinien und Rechtsverordnungen sind für die Teilnehmer:innen noch nicht gänzlich bekannt. Hier herrscht Informationsbedarf in Form von Weiterbildungen und Veranstaltungen auch für die Wirtschaftsunternehmen. Die Digitalagentur könnte zum Thema Open Data auch die KMUs beraten und zu den vorhanden rechtlichen Rahmenbedingungen und den Potenzialen der Nutzung von Daten für Geschäftsmodelle informieren.

- ★ **Open Data Wissen für Wirtschaft:** Insgesamt war der Wissensstand zum Thema Open Data in Berlin eher gemischt. Wirtschaftliche Akteur:innen beschäftigen sich aktuell wenig mit den Potenzialen von eigenen Daten und Open Data der Verwaltung. Sie nutzen wohl auch kaum die Verwaltungsdaten auf dem Open Data Portal. Es braucht für diese Zielgruppe auch deutlich mehr Unterstützung; Wäre es denkbar, dass ODIS auch städtische Betriebe berät? Den Wunsch einer ODIS nicht nur für die Verwaltung sondern auch für die Wirtschaft fand breite Unterstützung bei den Teilnehmer:innen.

- ★ **Berlin als Vorreiterstadt für Open Data in Deutschland:** Die Berliner Verwaltung und der Chief Digital Officer sollen aktiv das Thema Open Data für das Land Berlin steuern. Aufgabe wäre es, Berlin auf der Open Data Landkarte als Vorreiterstadt zu verankern und den Unternehmen den Zugang zu den offenen Daten der Verwaltung ermöglichen.

- ★ **Qualitätssicherung:** Damit die Wirtschaft, KMUs, Startups und Unternehmen die Daten der Verwaltung nutzen können, muss ein bestimmter Qualitätsstandard für die Daten gewährleistet werden.

- ★ **Urban Data Platform Berlin:** Wir brauchen ein übergreifendes Konzept für eine urbane Datenplattform. Auf der Urban Data Platform würden Daten aus den Bereichen Ver- und Entsorgung, Verwaltung, Gesundheit, Bauwesen, Mobilität, Logistik, Transport und Wirtschaft zur Verfügung gestellt und miteinander vernetzt werden.
Die Integration und Vernetzung von städtischen Daten und standardisierten Schnittstellen würde durch die Urban Data Platform schnell und einfach realisiert werden. Es gibt den Wunsch von den Akteur:innen der Wirtschaft, dass es nur eine Plattform gibt als „Single Point of Truth“, wo die Daten der Berliner Verwaltung durch Dritte abgerufen werden können.

- ★ **Digital Twin für die Stadt Berlin:** Der Digitale Zwilling soll das digitale Herzstück der Zukunftsstadt sein. Damit kann auch den Herausforderungen der Smart City mit innovativen Lösungen begegnet werden. Die Urban Data Plattform soll dabei die zentrale Datendrehscheibe des Digitalen Zwillings darstellen. Mit ihr sollen ehemals separierte Insellösungen zu einem gemeinsamen Ökosystem der Stadt vernetzt werden. Zentrale Zukunftsthemen wie Klimaschutz, eine zukunftsorientierte Mobilität oder die integrierte Stadtentwicklung könnte Berlin mit dem „Digitalen Zwilling“ bestmöglich umsetzen.

7. Zusammenfassung der Diskussionsergebnisse und Handelsempfehlungen “Workshop Wirtschaft”

- ★ Die Wirtschaft fordert den notwendigen Kulturwandel in den Unternehmen, KMUs nach dem Motto “So kann es gehen!”, dazu gehört auch ein aktives “Experimentieren” bei den Möglichkeiten der Nutzung von eigenen Daten aber auch beim Abschöpfen der Potentiale von Open Data der Verwaltung für eigene neue Geschäftsmodelle.
- ★ Die Vorteile von „Linked Open Data“, im Sinne der Vernetzung von Daten und Zusammensetzung in einen Kontext Rahmen, sollten mit konkreten Anschauungsbeispielen und Erfolgstories aus Berliner Betrieben/Verwaltungen aufgezeigt und auch standardmäßig umgesetzt werden.
- ★ Es besteht noch Diskussions- und Erklärungsbedarf bei den Herausforderungen der landeseigenen Betriebe in den Bereichen Rechtssicherheit, Datenschutz, Haftungsklausel im Rahmen von Open Data und der kritischen Infrastruktur.
- ★ Der Auf- und Ausbau des Berliner Open Data Ökosystems benötigt Ressourcen. Dazu sollen einerseits die richtigen IT-Tools zur Verfügung gestellt werden, andererseits die technische, personelle und auch finanzielle Ausstattung bzw. Unterstützung gewährleistet sein.
- ★ Die Einrichtung einer ODIS für die Wirtschaft, soll die Potentiale der Datennutzung auf dem Open Data Portal aufzeigen, aber auch Wirtschaftsunternehmen dazu befähigen, selbst offene Daten auf dem Open Data Portal zur Verfügung zu stellen.
- ★ Für die Smart City Strategie wird die Bereitstellung der Daten, nicht nur als Excel- oder CSV-Datei auf dem Open Data Portal zielführend sein, sondern auch eine Abbildung der Daten als Digital Twin von der Stadt Berlin bzw. der Aufbau einer "Urban Data Platform“, um die Daten aus Klimaschutz, Mobilität oder einer integrierten Stadtentwicklung bestmöglich zu verdeutlichen.

8. Ausblick

Die Ergebnisse des Workshops werden den Teilnehmer:innen zur Kommentierung zurück gespielt. Im weiteren Prozess folgt der Workshop mit der Zielgruppe Wissenschaft. Alle aktuellen Entwicklungen und Ergebnisse zur Open Data Strategie können auf der [öffentlichen Informationsseite](#) eingesehen werden. Die Open Knowledge Foundation freut sich auf **Anregungen, Ergänzungen oder Kommentierungen**: opendataberlin@okfn.de

Für die Open Data Strategiefindung wurde die OKF von der Senatsverwaltung für Wirtschaft, Energie und Betriebe im Jahr 2022 beauftragt, gemeinsam mit der Open Data Informationsstelle den Online-Beteiligungsprozess auf www.meinberlin.de und die Partizipation des Open Data Ökosystems bestehend aus Verwaltung, Wirtschaft, Wissenschaft und Zivilgesellschaft in Form von Workshops durchzuführen.

Ziel ist es, die Erkenntnisse aus dem breiten Partizipationsprozess in das zukünftige Open Data Berlin Konzept des Landes Berlin einfließen zu lassen.



Ergebnisbericht

zum **Strategieworkshop** mit der
Stakeholder:innen-Gruppe der **Zivilgesellschaft**

am 31. März 2022 im CityLAB Berlin

im Rahmen des Partizipationsprozesses für die
Open Data Strategie Berlin 2022

Inhaltsverzeichnis

1. Einführung durch das Open Data Strategieteam.....	2
2. Vorträge der OKFN, SenWEB und Jens Ohlig.....	2
3. Kleingruppenarbeit mit der Wardley-Mapping Methode	3
4. Ergebnisse der Kleingruppenarbeit als Wardley-Maps	4
5. Ergebnisse der Kleingruppenarbeit: „Welche Anforderungen haben die Stakeholder:innen der Zivilgesellschaft?“	9
6. Zusammenfassung der Diskussionsergebnisse und Handlungsempfehlungen “Workshop Zivilgesellschaft”	12
7. Ausblick	13

1. Einführung durch das Open Data Strategieteam

Am **31. März 2022** gingen die Partizipationsworkshops zur neuen Open Data Strategie Berlins in die zweite Runde. 28 Teilnehmer:innen vor allem aus den Gruppen des Digitalen Ehrenamts trafen sich zum Zivilgesellschaft-Workshop im CityLAB.

2. Vorträge der OKFN, SenWEB und Jens Ohlig

Im Anschluss an die Begrüßung durch Henriette Litta, Geschäftsführerin der OKF, referierte Betül Özdemir, Referentin Open Data der Senatsverwaltung für Wirtschaft, Energie und Betriebe zum Status Quo der 10 Jahre Open Data in Berlin und stellte die Ziele und den Partizipationsprozess der aktuellen Strategiefindung vor.

Nach der Vorstellung der [Ergebnisse des Online-Beteiligungsprozess](#) durch Walter Palmetshofer von der OKF e.V. konnten die Anwesenden im Rahmen der Vorstellungsrunde auch ihre persönliche Einschätzung zu den Entwicklungen der vergangenen 10 Jahre Open Data in Berlin und Deutschland einbringen. Mehrfach wurde festgestellt, dass die Entwicklungen unerwartet langsam voranschreiten und häufig nicht erkennbar auf eine weitgehende Automatisierung zuarbeiten würden. Betül Özdemir von SenWEB machte das Angebot eines Hackathons für Linked Open Data, um weiter auf die Herausforderung eines einfacheren “Literacy” des Linked Open Data Ansatzes (dies war ein zentrales Feedback vom 1. Workshop am 29.03.2022 mit der Verwaltung) *und dessen Vorteile besser darzustellen*. Der Vorschlag wurde auch nach dem Workshop noch in Einzelgesprächen diskutiert und weiterentwickelt und sehr begrüßt. Als vielversprechender Veranstaltungsrahmen könnte das Barcamp-Format gewählt werden.



Die Workshop-Teilnehmer:innen im CityLAB Berlin

Jens Ohlig hielt daraufhin seinen [Impulsvortrag „Linked Open Data und der Weg nach Hause“](#) zu den Chancen und Möglichkeiten verlinkter offener Daten, und veranschaulichte, wie Linked Open Data bereits im Projekt WikiData ermöglicht wurde.

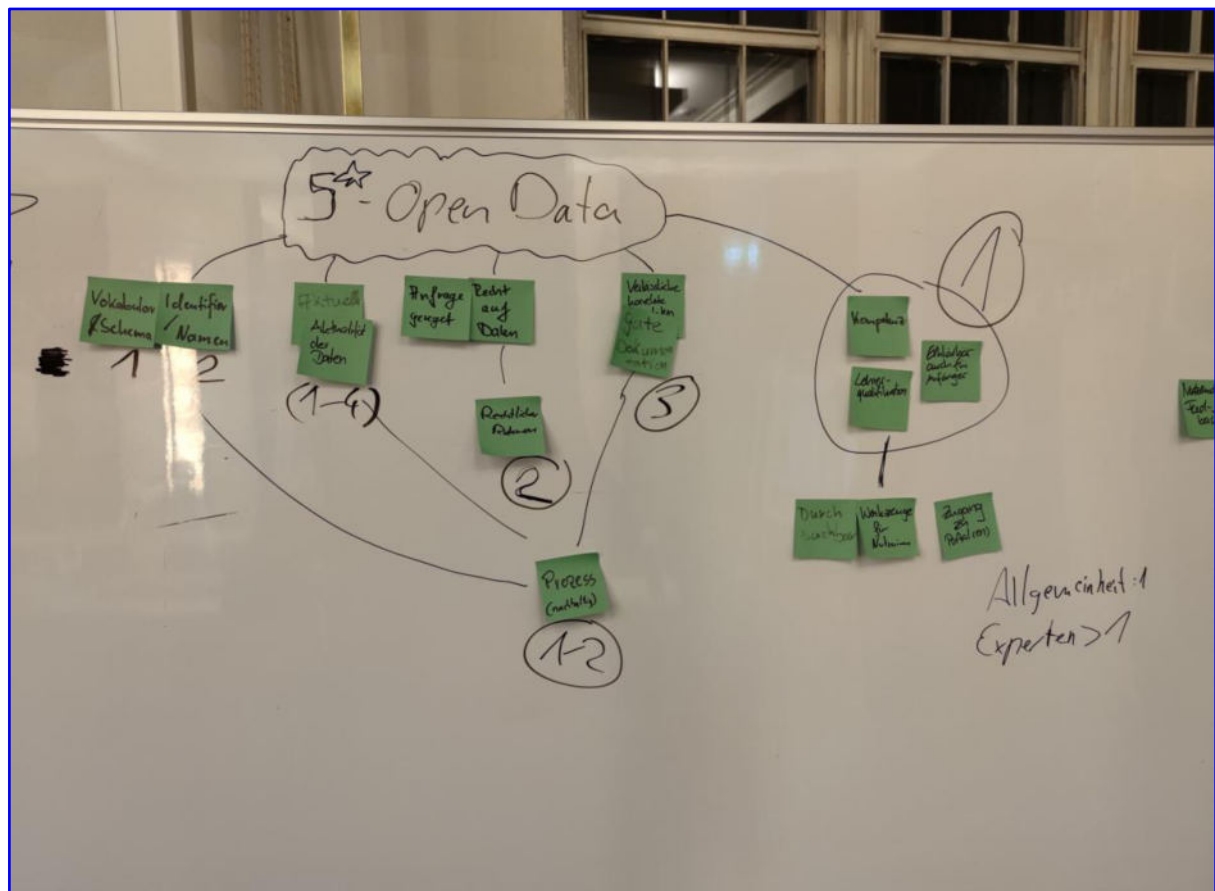
3. Kleingruppenarbeit mit der Wardley-Mapping Methode

Anschließend sind die Teilnehmer:innen in Kleingruppen zusammengekommen. Zu den 4 Gruppen vor Ort schaltete sich eine Gruppe digital zu. Analog zum Verwaltungsworkshop wurde eine angepasste Wardley-Mapping-Methode genutzt, die auch bei den anderen Stakeholder:innen Workshops verwendet wurde. Die Kleingruppen brachten ihre „Map“ auf Whiteboards bzw. bei der digitalen Gruppe auf einem Miro-Board und stellten diese im Anschluss dem Plenum vor. Neben der reinen „Map“ als Ergebnis der Kleingruppenarbeit hielten die Moderator:innen auch die fruchtbaren und intensiven Gespräche und Diskussionen fest. Die Ergebnisse der Kleingruppenarbeit werden mit Handlungsempfehlungen der Stakeholder:innen Gruppe der Zivilgesellschaft in die Open Data

Strategie einfließen. So wurde die Wardley Methode in den Kleingruppen umgesetzt und folgende Inputs erzielt:

4. Ergebnisse der Kleingruppenarbeit als Wardley-Maps

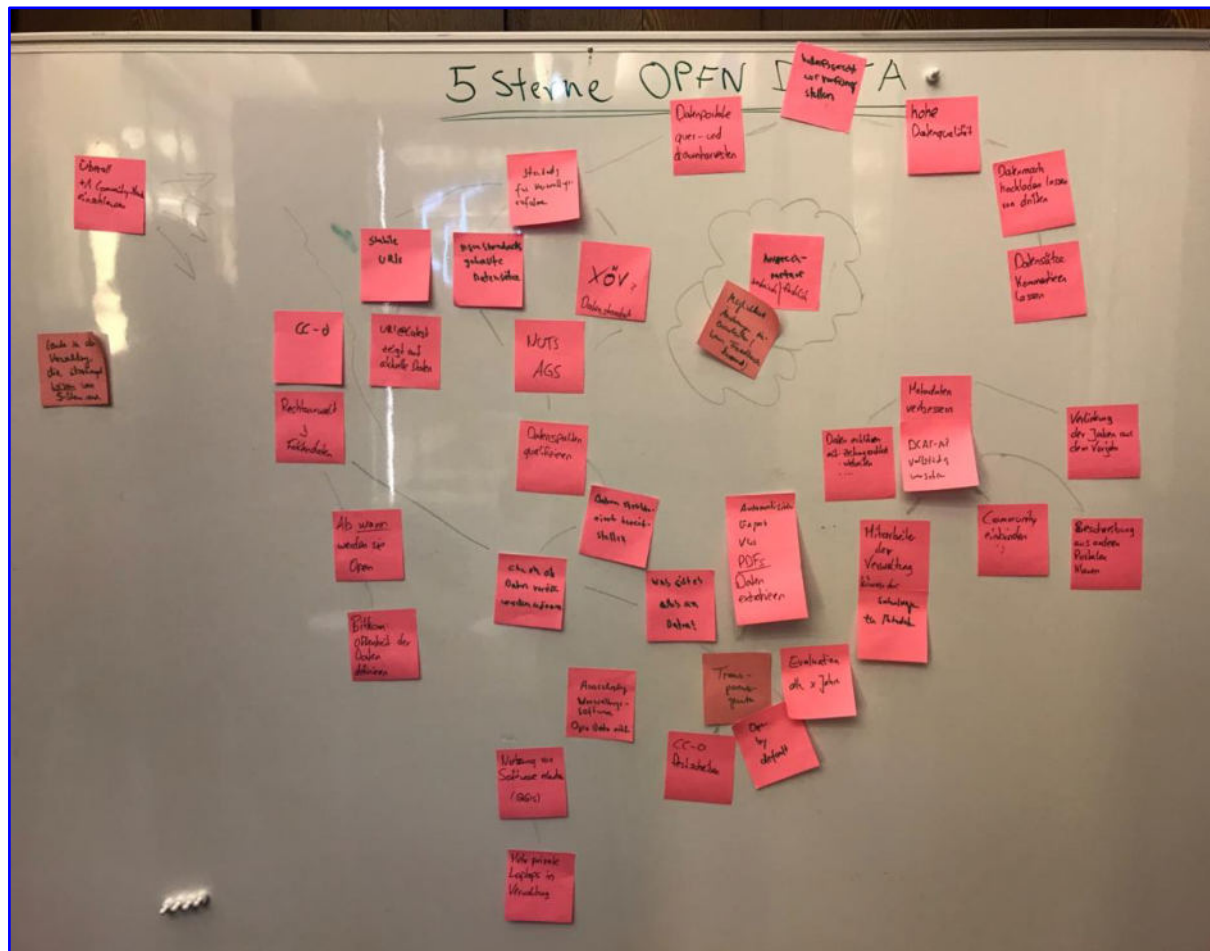
Wardley-Map Gruppe Nummer 1:



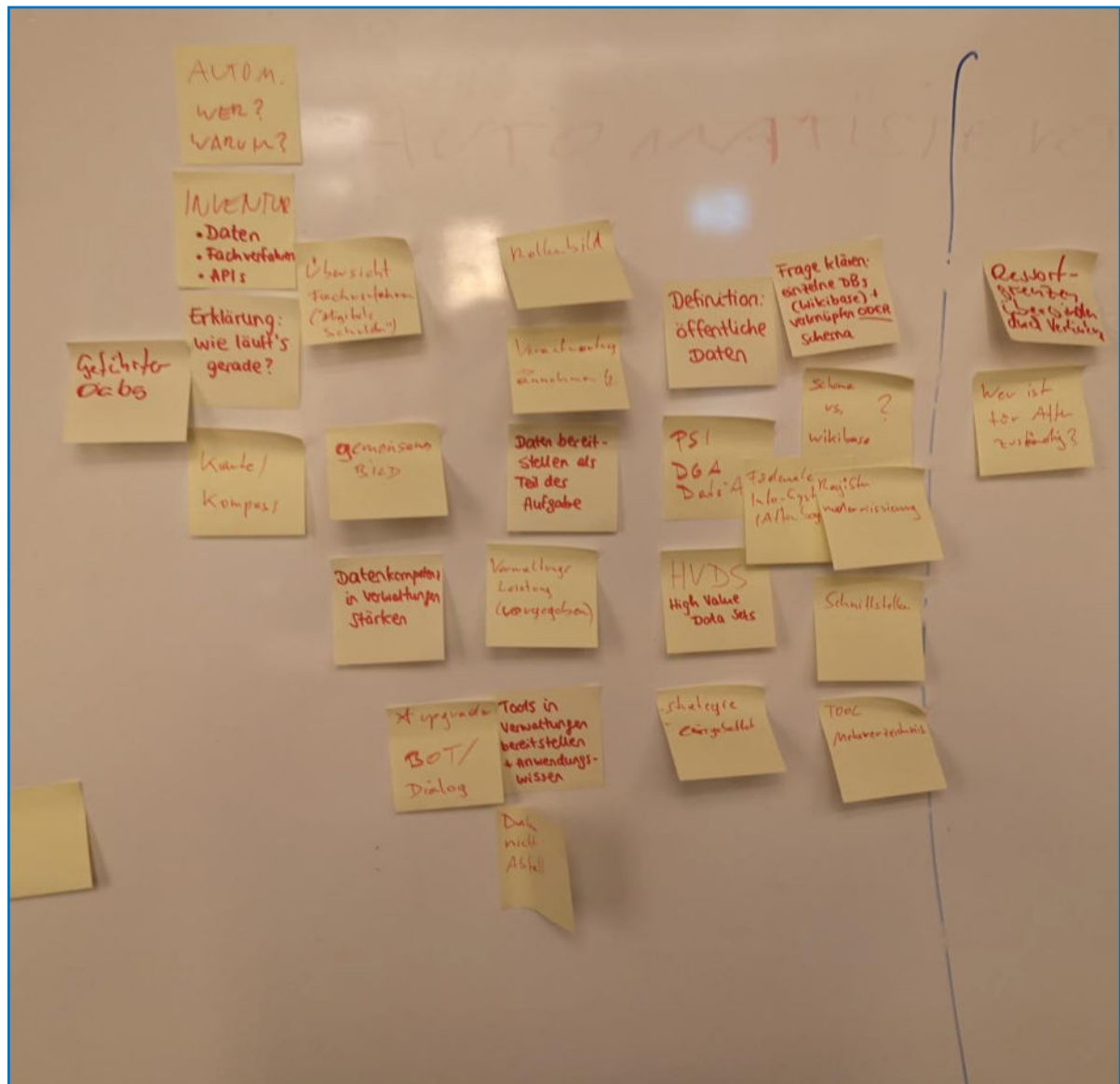
Wardley-Map Gruppe Nummer 2:



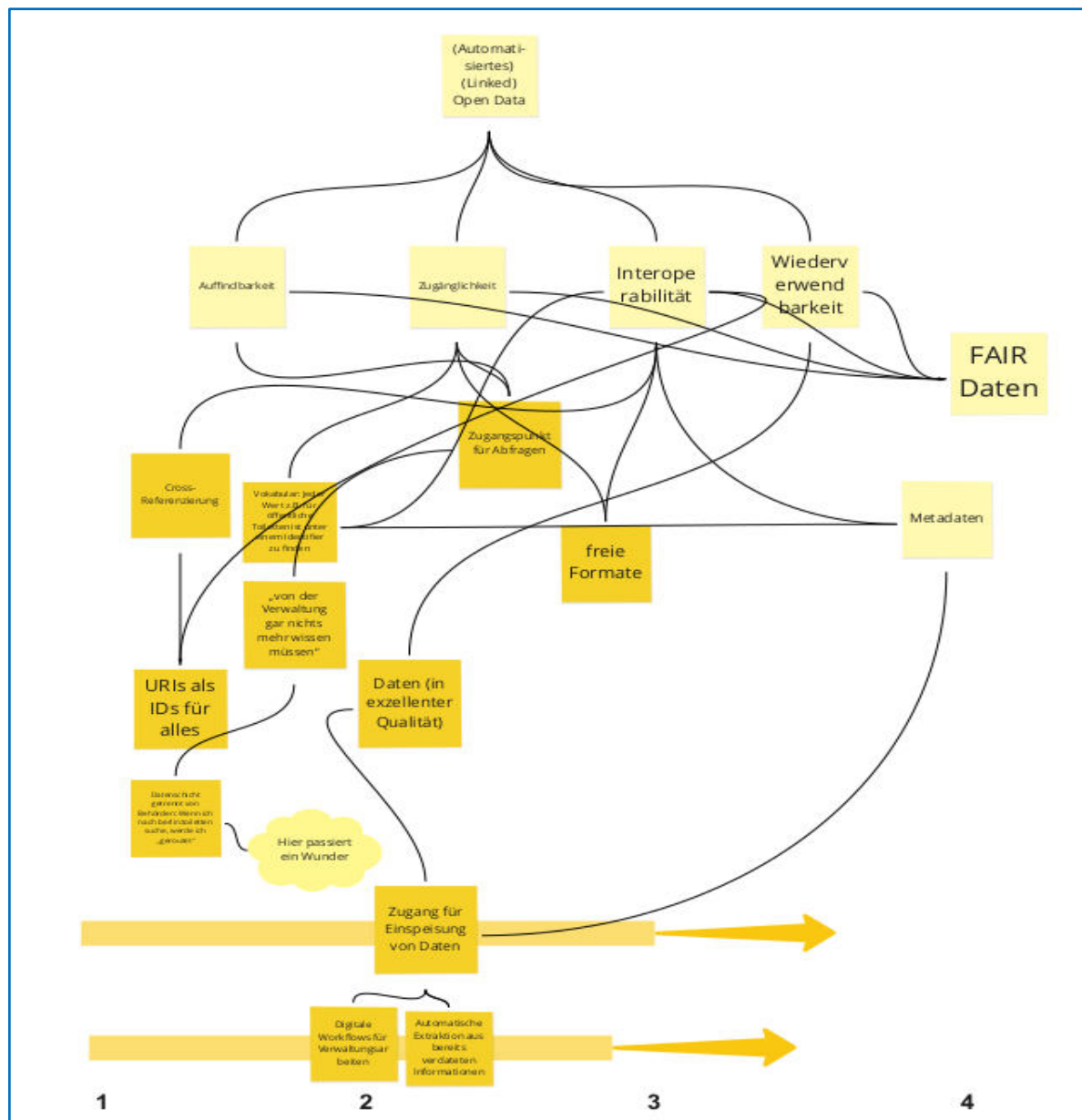
Wardley-Map Gruppe Nummer 3:



Wardley-Map Gruppe Nummer 4:



Wardley-Map Gruppe Nummer 5 (remote):



5. Ergebnisse der Kleingruppenarbeit: „Welche Anforderungen haben die Stakeholder:innen der Zivilgesellschaft?“

Gruppenübergreifend sind in der Diskussion und Herausarbeitung zentrale Schwerpunkte für die Open Data Strategie aus Sicht der Zivilgesellschaft mit Blick auf ein 5* Open Data zusammengekommen:

- ★ **Linked Open Data** ist ein **sehr gutes Ziel**, aber wir sind vom notwendigen Unterbau für die ersten vier Sterne aus dem „5* Sterne-Modell“ noch sehr weit entfernt.
- ★ Es sollte ein **Open Data Institut in Berlin** eröffnet werden: Ein Open-Data-Institut bereitet Gesellschaft, Politik und Wirtschaft auf einen ganzheitlichen und chancenorientierten Umgang mit offenen Daten vor. Es erarbeitet Konzepte, technische Lösungen und Studien und unterstützt als unabhängige, überparteiliche und interdisziplinäre Organisation den Abbau von Barrieren für eine chancenorientierte Datennutzung. Ein Open-Data-Institut verstärkt die Veröffentlichung und evidenzorientierte Nutzung von Daten, fördert die Zivilgesellschaft als Ideenpartner und schafft durch die Unterstützung bei Datenveröffentlichungen Vertrauen in den öffentlichen Sektor.
- ★ Ziel muss nicht nur sein, einen einzelnen Datensatz zu verbessern oder das Berliner Open Data Portal, sondern generell zu ermöglichen, zwischen allen bestehenden Datenportalen auf lokaler, Landes-, Bundes-, EU-Ebene interessante **Daten „quer- und downverlinken“** zu können: Standards sind hier maßgeblich!
- ★ **Wir brauchen einen Rechtsanspruch für Open Data:** Es gibt zwar eine Open Data Rechtsverordnung, aber anscheinend wird sie nicht oder nicht ausreichend umgesetzt. Sonst gäbe es ja „open by default“. Es fehlt an Sanktionen bei „Nicht-Umsetzung der Open Data Verordnung“ und offensichtlich Verantwortlichkeiten und Zuständigkeiten für Open Data. Eventuell kann man den Rechtsrahmen anpassen und verbessern.
- ★ Die **Datenqualität** muss gewährleistet sein. Bessere Datenqualität fördert auch die Weiterverarbeitung der Daten, da die mögliche Frustration durch notwendige Nachbearbeitung der Daten abnimmt.

- ★ Wichtig ist für **5-Sterne Open Data** auch eine klare Identifikation der Daten sowie eine unverkennbare und **unveränderbare ID für die Daten** sowie ein klares Vokabular (z.B. brauchen wir stabile, unveränderliche URIs auf dem Open Data Portal)
- ★ Die **Metadaten** müssen massiv verbessert werden; hier liegen massive Fehler, kaum ein Datensatz in Deutschland verfügt über „gute“ Metadaten (Wir brauchen einfache, eindeutige, ohne Kenntnis der Verwaltungssprache nachvollziehbare Metadaten)
- ★ Weiteres Thema war die mangelnde Daten-**Kompetenz** der Datenveröffentlichenden: Eine Erkenntnis war, dass es an Personen in der Verwaltung mangelt, die wissen was „**5-Sterne Open Data**“ sind. Dazu braucht es bessere Erzählungen und Beispiele aus dem Verwaltungsalltag. Auch der Vorschlag des Brückenschlags mit der Zivilgesellschaft zum regelmäßigen Austausch, um den Mehrwert von Linked Open Data zu verdeutlichen, wurde angedacht.
- ★ Die Daten auf dem Open Data Portal sollten **grundsätzlich als CC0-Lizenz** veröffentlicht werden. Restriktive Lizenzen oder Lizenzen mit einer Namensnennung machen es Datennutzenden schwieriger die Daten weiter zu verwenden, bzw. auch bei der Verwendung der Verwaltungsdaten richtig zu zitieren z.B. für Kartenanwendungen.
- ★ Die Veröffentlichung von Daten als **Echtzeitdaten** (z.B. bei Umwelt- oder Mobilitätsdaten) ist ein Ziel für die Weiterentwicklung des Datenbestandes auf dem Open Data Portal.
- ★ Es sollten die „**High Value Datasets**“ auch auf Landesebene in Berlin veröffentlicht werden. Es gibt bereits die Publikation „Kerndatensätze der ODIS“, die die wertvollsten Daten der Berliner Verwaltung auflistet. Ziel sollte es sein, dass diese Kerndatensätze analog zu den „High Value Datasets“ veröffentlicht werden.
- ★ Die **Klimadaten vom Berliner Energie und Klimaschutz Konzeptes 2030** sind ein gutes Daten-Beispiel in Berlin, deren Veröffentlichung zur Lösung der klimapolitischen Herausforderungen nützlich sein könnte.

- ★ Vernetzung mit anderen Bundesländern als Best Practice zum aktuellen Stand für Open Data soll angedacht werden. Dies betrifft einerseits den Vergleich der Standards für Open Data in anderen Städten, andererseits auch für den Vergleich, inwiefern andere Städte Open Data in weitere rechtliche Rahmenbedingungen einbetten und wie sie ihre Open Data Strategien umsetzen.

- ★ Vorschlag für ein praktisches Anwendungsbeispiel: Zur Unterstützung bei der Bereitstellung von zukünftigen Linked Open Data soll **ein Bot-Tool**, der erklärt welche weiteren Maßnahmen es braucht, um die Daten in die nächste „Sternekategorie“ von Linked Open Data zu heben, die Verwaltung unterstützen.

- ★ Herangehensweise für Linked Open Data wie beim 5* Sterne Schema: Es sollten die **einzelnen Schritte klar vorher definiert** werden, um die einzelnen Sterne für Open Data zu erreichen. Oder man nutzt die vorhandene Grundlage von wikibase, so dass man mit diesem Modell erst einmal loslegen kann und dann überlegt, wie man später die Ansätze für 5* Sterne Open Data wieder zusammenführt.

- ★ Auf dem Open Data Portal sollte es die Möglichkeit geben, die **Daten zu kommentieren**. Es sollten auch ältere Versionen veröffentlichter Daten auffindbar sein, so dass man die Entwicklung der Daten nachvollziehen kann. (Nachvollziehbarkeit)

- ★ Um eine bessere Interpretierbarkeit der veröffentlichten Datensätze zu ermöglichen, sollte es neben den maschinenlesbaren Formaten auch mehr **Visualisierungen und Dashboards** der Daten geben, um auch so den Mehrwert von Daten aufzuzeigen.

6. Zusammenfassung der Diskussionsergebnisse und Handelsempfehlungen “Workshop Zivilgesellschaft”

- ★ Die Zivilgesellschaft sieht den Linked Open Data Ansatz als ein sehr begrüßenswertes langfristiges Ziel. Idealerweise wird die Linked Open Data Vision in Berlin bis spätestens 2030 umgesetzt, bzw. werden die dafür notwendigen Rahmenbedingungen endlich geschaffen.
- ★ Die Bereitstellung der Daten ist Teil der Aufgabe der Verwaltung, im Rahmen der Open Data Verordnung gehört auch der Rechtsanspruch auf Daten dazu. Die dabei verwendete Lizenz auf dem Open Data Portal für alle Daten CC0 sein.
- ★ Die Zivilgesellschaft fordert die dafür notwendigen Ressourcen in der Verwaltung (Personal, klare Zuständigkeiten und Ansprechpersonen mit z.B. genügend Zeitbudget für die Open Data Beauftragten, finanzielle Ausstattung der IT Infrastruktur, die notwendige Fortbildung zum besseren Datenverständnis und die Erhöhung der Datenkompetenz), um endlich den Datenzugang für die Berliner Bürger:innen zu den Daten der Verwaltung auszubauen.
- ★ Neben der Open Data Verordnung wäre auch eine proaktivere Herangehensweise der Stadtverwaltung und der Führungsspitzen bei der Datenöffnung begrüßenswert, welche über die Mindesterfüllung der Open Data Richtlinie der EU hinaus geht (z.B. Bereitstellung von Echtzeitdaten im Umwelt- und Mobilitätsbereich).

Zum Abschluss fand unter Berücksichtigung der erforderlichen Infektionsschutzmaßnahmen nach 2 Jahren Corona endlich wieder einmal ein informeller Austausch der ehrenamtlich Aktiven statt.



Im Anschluss des Workshops: Gemeinsames diskutieren mit Pizza und Kaltgetränken

7. Ausblick

Die Ergebnisse des Workshops werden den Teilnehmer:innen zur Kommentierung zurück gespielt. Im weiteren Prozess folgen Workshops mit den Zielgruppen Wirtschaft und Wissenschaft. Alle aktuellen Entwicklungen und Ergebnisse zur Open Data Strategie können auf der [öffentlichen Informationsseite](#) eingesehen werden.

Die Open Knowledge Foundation freut sich auf **Anregungen, Ergänzungen oder Kommentierungen:** opendataberlin@okfn.de

Für die Open Data Strategiefindung wurde die OKF von der Senatsverwaltung für Wirtschaft, Energie und Betriebe im Jahr 2022 beauftragt, gemeinsam mit der Open Data Informationsstelle den Online-Beteiligungsprozess auf meinberlin.de und die Partizipation des Open Data Ökosystems bestehend aus Verwaltung, Wirtschaft, Wissenschaft und Zivilgesellschaft in Form von Workshops durchzuführen.

Ziel ist es, die Erkenntnisse aus dem breiten Partizipationsprozess in das zukünftige Open Data Berlin Konzept des Landes Berlin einfließen zu lassen.

Open Data in der Berliner Verwaltung

Status quo, Bedarfe und Perspektiven

Impressum

© Technologiestiftung Berlin 2018

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

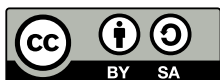
info@technologiestiftung-berlin.de · technologiestiftung-berlin.de

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-Go-vernment-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.

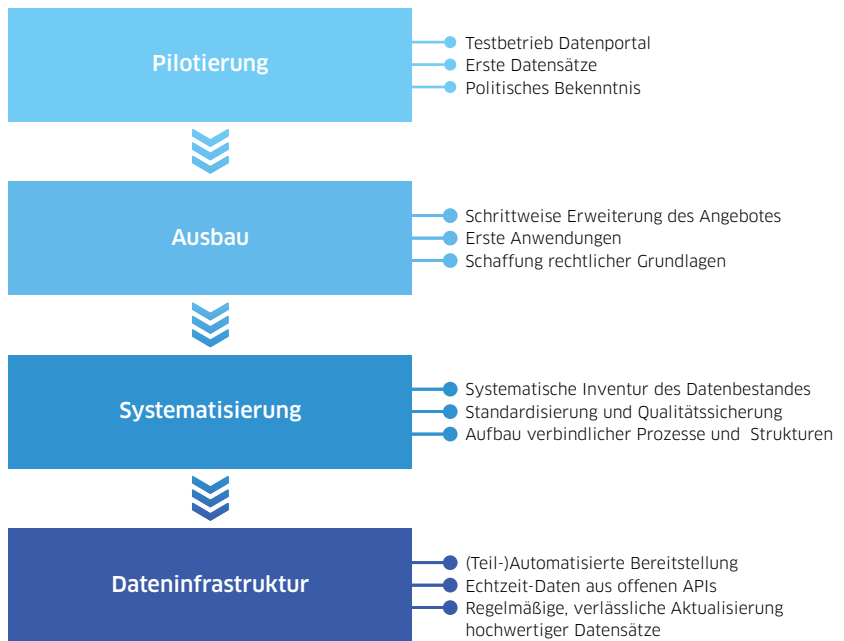
4 www.oecd.org/sti/ieconomy/PolicyNote-DDI.pdf

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, featuring a prominent red brick church with a clock tower in the foreground. The city's dense urban landscape, including various apartment buildings and commercial structures, is visible. A large, semi-transparent blue geometric shape, composed of two overlapping triangles, is positioned over the upper half of the image. A smaller, semi-transparent red triangle is located in the lower right corner, partially overlapping the blue shape and the cityscape. The sky is bright blue with scattered white clouds.

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.

[illegible]

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 selbst 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

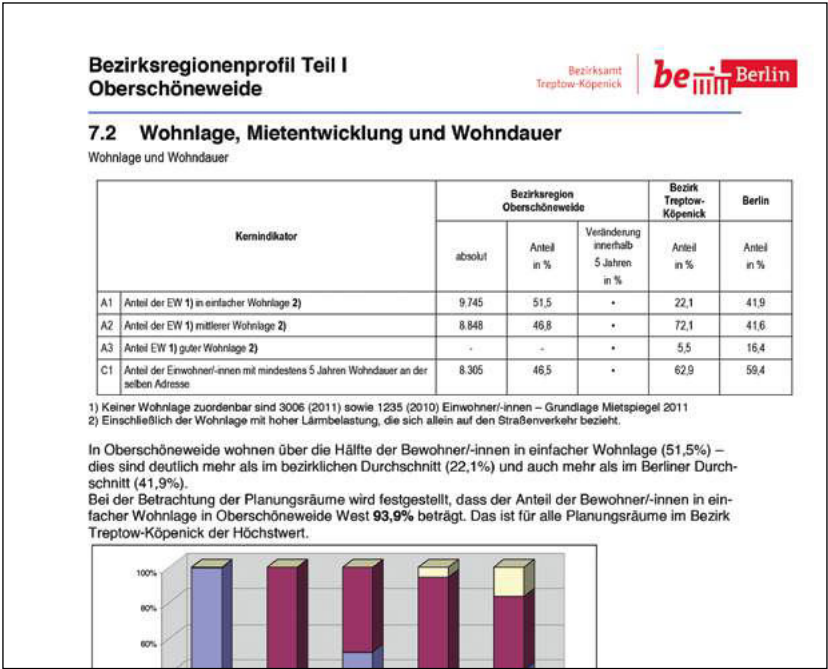
Eine wichtige Rolle in diesem Zusammenhang spielt die Position der „bezirklichen Datenkoordination“, die in mittlerweile sechs Bezirken¹³ als Teil der „Organisationseinheit Soziale Planungskoordination“ (OE SPK)¹⁴ eingerichtet wurde. Die Datenkoordination übernimmt nach Bedarf die Beschaffung, Zusammenführung, 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 Planungskoordination“ (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 Oberschöneweide (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.

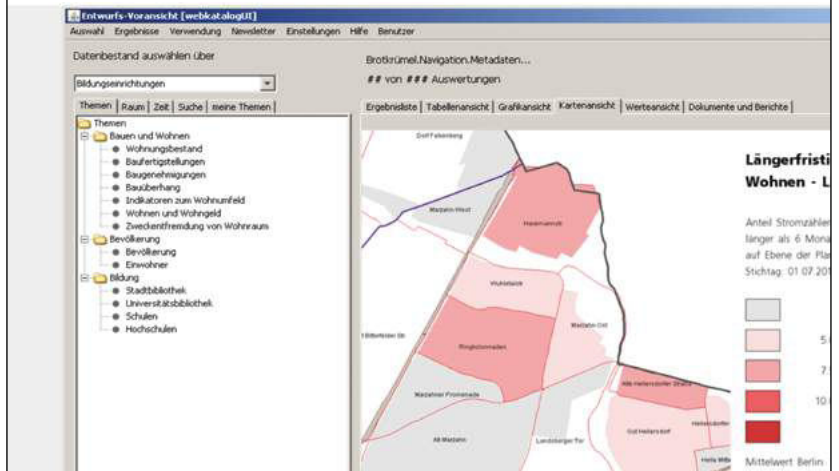
Im Rahmen dieser Tätigkeiten fragen die Senatsverwaltungen regelmäßig Datenbestände von Bezirken, Landesämtern und Betrieben an und werten diese aus. Verglichen mit den Bezirken werden also weniger Primärdaten erhoben. Gemäß der ministeriellen Aufgaben werden vielmehr bestehende Daten aus verschiedenen 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 Berl
 Start Sitemap Kontakt Newsletter Hilfe Impressum mit Berlin

Einfache Suche

Erweiterte Suche

Überblick

Hinweise

Aktuell

RSS-Feed

Ordner

Gesundheitsdaten

GBE-Indikatoren

Erläuterungen

Sozialdaten

Sozialstatistiken

Berichte

Publikationsreihen

Schwort

Info zur Suche

FAQ (häufige Fragen)

Abkürzungen

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 (79)	Gesetzliche Krankenversicherung (GKV) (23)
		Gesundheitsausgaben nach Ausgabenträgern und Leistungsarten (1)
		SGB XII (BSHG), BVG und LPIfGG (PflegeG) (36)
		weitere Ausgabenträger (13)
	Indikatoren zur Sozialversicherung (42)	Beitragsätze, Grundlohsummen und GBA (42)

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

Katalog Meine Tabellen Diagrammansicht Tabellenansicht Tabellenansicht Tabellenansicht Tabellenansicht Tabellenansicht

Datenquelle: Binnenwanderung im Land Berlin bis 2013

Tabelle herunterladen: Excel 2007 (xlsx) max 16.384 Spalten x 65.000 Zeilen

Daten abrufen Tabelle leeren Tabelle speichern Tabelle drucken Papiermodus

Jahre nach Herkunftsgebiet nach Zielgebiet

Filter: Standardsumme: Anzahl

Schichten: 2013 2

Zeilenanzahl: 308 (12 x 12 x 12) insgesamt, 144 (12 x 12 x 12) 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	2.425	1.326	1.769	1.548	1.847	684	731
02 Friedrichshain-Kreuzberg	2.794	-	2.953	1.021	371	820	2.096	3.828	1.228	668
03 Prenzlau	3.097	2.506	-	841	378	544	805	1.021	813	969
04 Charlottenburg-Wilmersdorf	2.425	1.127	1.029	-	1.845	2.555	2.406	1.031	341	291
05 Spandau	952	431	820	1.379	-	534	582	387	390	385
06 Steglitz-Zehlendorf	1.110	525	498	1.895	558	-	2.578	803	322	213
07 Tempelhof-Schöneberg	1.752	1.427	827	2.291	721	2.988	-	2.426	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	306	233	262	498	982	-	1.096
10 Marzahn-Hellersdorf	486	435	851	253	234	198	318	280	1.047	-
11 Lichtenberg	978	1.158	1.561	480	311	303	457	509	1.258	2.635
12 Reinickendorf	1.068	425	1.090	875	998	470	598	478	258	381

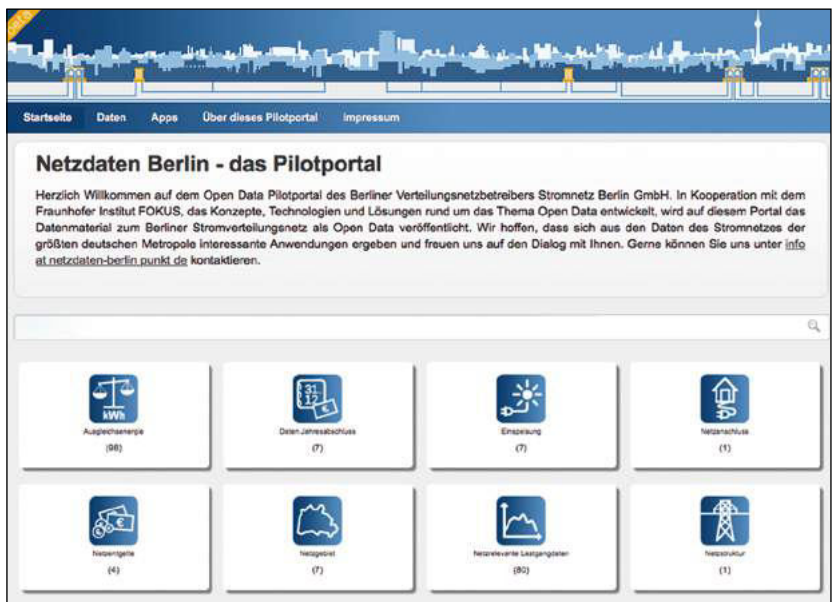
Statistisches Informationssystem Berlin-Brandenburg (StatIS-BBB)

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.



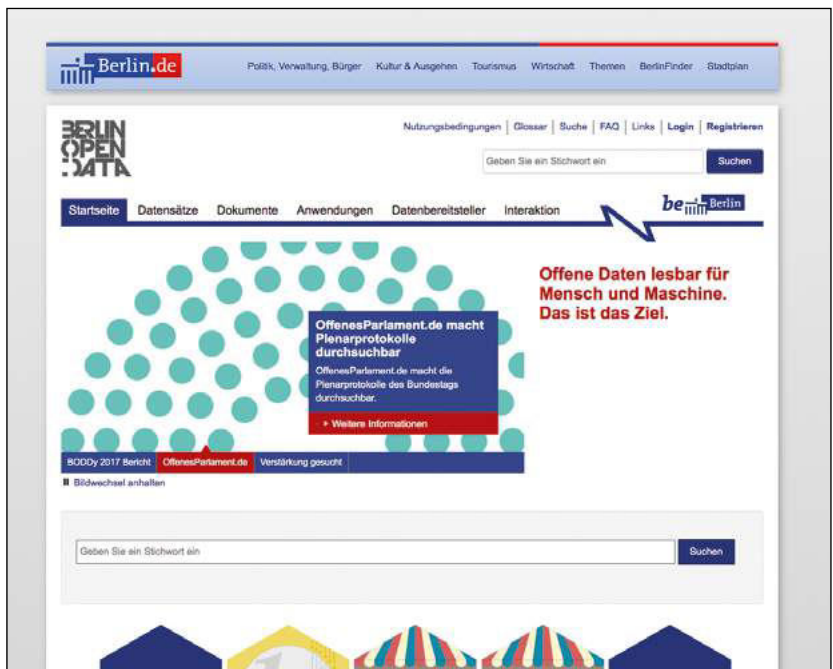
Netzdaten Berlin

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.



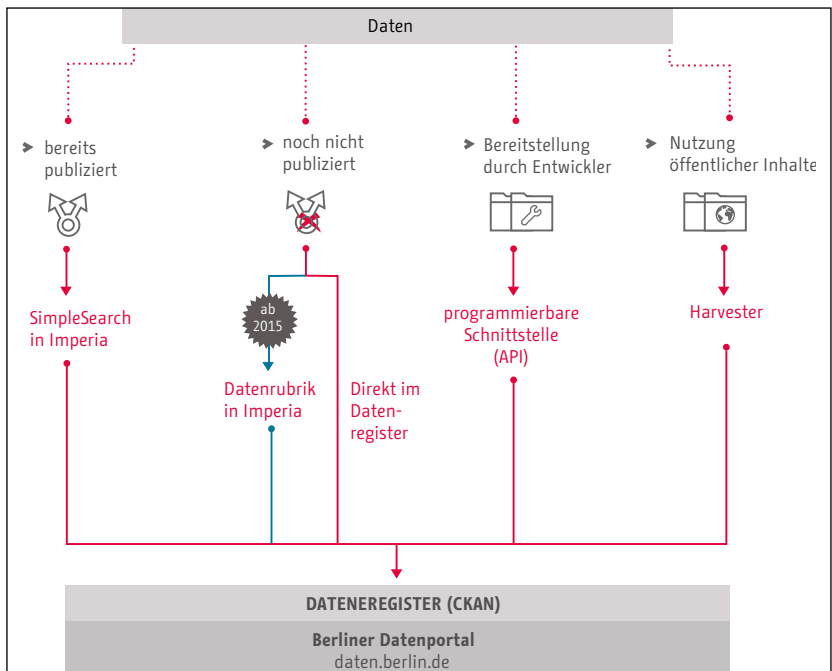
Berliner Datenportal

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 (StatIS 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.

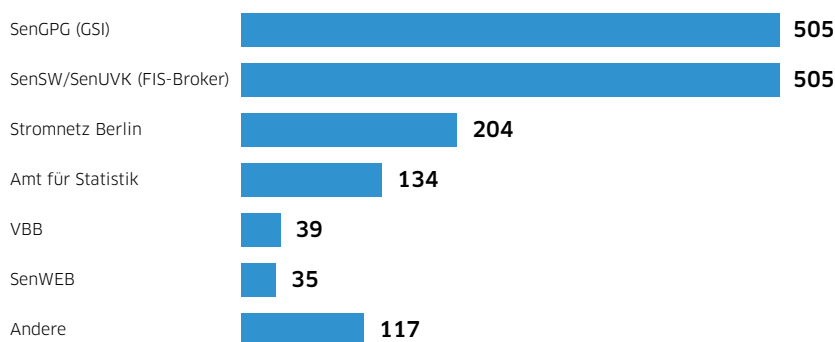
Durch eine zusätzlich geschaffene Schnittstelle lassen sich die Tabellen, die hinter SimpleSearch-Formularen liegen, in wenigen Schritten auch im Berliner Datenportal veröffentlichen. Dazu müssen lediglich in einem separaten Reiter die zugehörigen Metadatenfelder ausgefüllt und eine Freigabe erteilt 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.

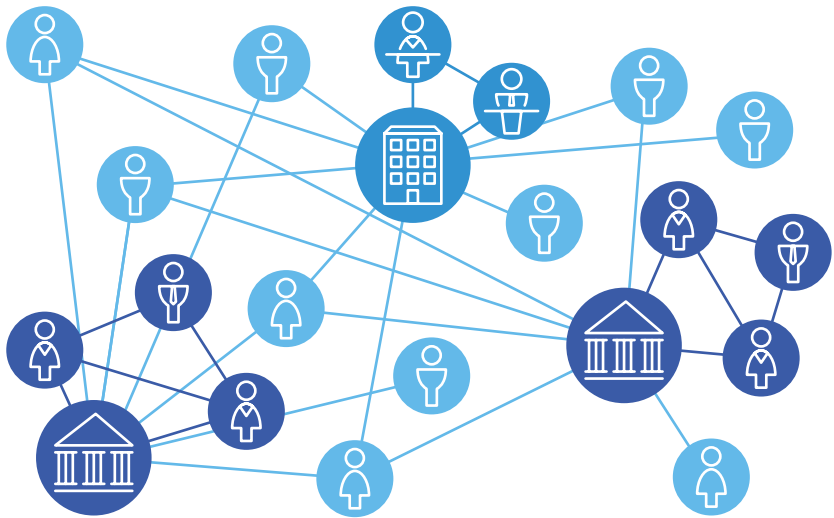
Im Rahmen zahlreicher Interviews mit Beschäftigten von Bezirks- und Hauptverwaltungen haben wir versucht, die zentralen Herausforderungen auf dem Weg zu einer systematischeren Datenbereitstellung zu identifizieren. Drei Handlungsfelder, die sich während unserer Gespräche mit den Mitarbeiter*innen als 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.

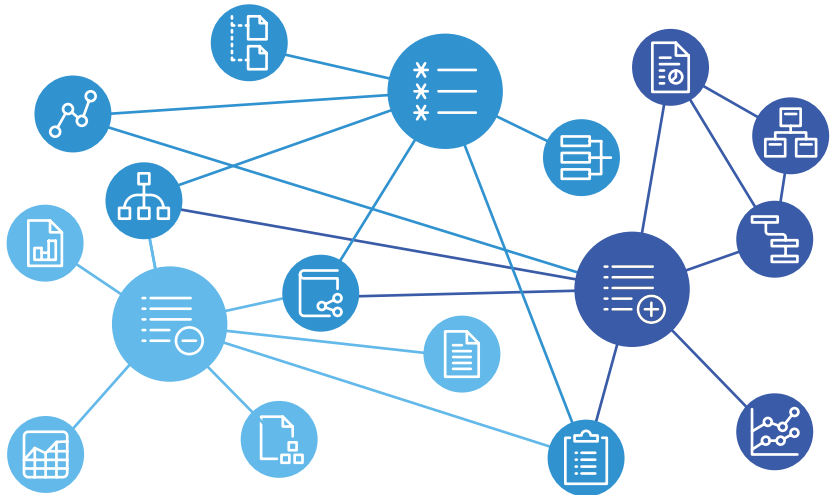
Bei der qualitativen Verbesserung der Datenveröffentlichung spielt auch der Austausch zwischen Verwaltung und der Open Data-Community eine wichtige Rolle. Um eine Weiternutzung von offenen Daten zu fördern, sollten Nutzergruppen möglichst früh in den Veröffentlichungsprozess einbezogen werden. Bei der 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.

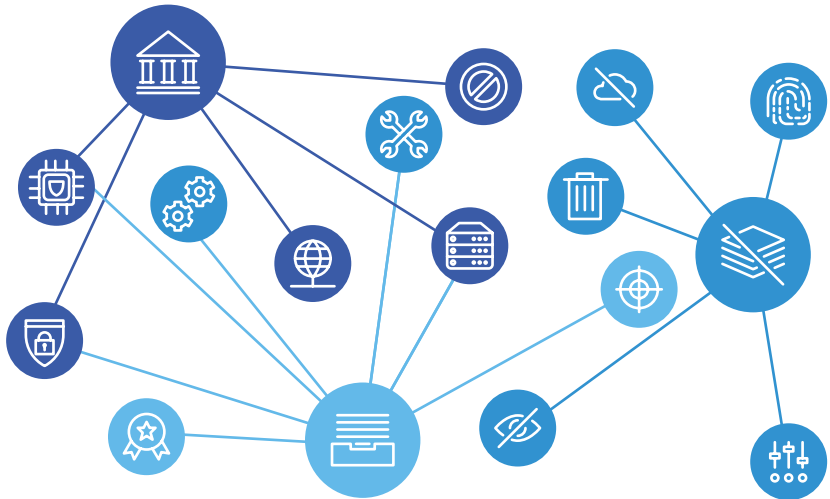
27 <https://index.okfn.org>

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 draufschaut, 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/datenberesteller> 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 Konsolidierung 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. Bislang 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.

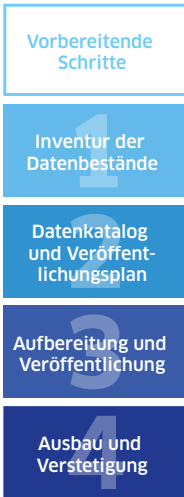
Mit Blick auf die in diesem Bericht geschilderten Herausforderungen wäre der Aufbau eines solchen Systems ein komplexes und zweifellos ambitioniertes Vorhaben. Mit der Einrichtung von Strukturen zur Standardisierung und Qualitätssicherung offener Verwaltungsdaten können jedoch schon heute wichtige Grundlagen 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>



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.

Angeichts 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](http://www.w3.org/2013/share-psi/bp/pomd) 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

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	Stadtplan	Kostenpunkt	Bewertung	NGP/DB	Ansprechpartner
Geolayer Kinderwagen-Infostafette (amstern - Hyperlink)	Stadtplan, WMIS	10	Phase noch offen	MA 11	JA	A	38	JA	Anna Mader
Geolayer Kulturlag. Historische Karten	WMIS	10	Phase noch offen	MA 8			23		Dr. Ingrid Fischer, MA 8
Geolayer Marktgebiete	Stadtplan, WMIS	10	Phase noch offen	MA 21 B			25		Martin Martin, MA 21B, WMIS
Geolayer Flurkarten, Flurstücks-Adressen	Stadtplan, WMIS	10	Phase noch offen	MA 21 B	JA		39		Heide-Martin, MA 21B, WMIS
Geolayer Ständige Geschäfter	Stadtplan, WMIS	10	Phase noch offen	WM 04		B	31	JA	Florian Thomas, WM 04
Geolayer Wohnungskategorie	Stadtplan, WMIS	10	Phase noch offen	MA 21A, MA 21B		B	31		Martin Martin
Geolayer Behinderteneinrichtungen - Erweiterung	Stadtplan, WMIS	10	Phase noch offen	MA 48			80		Wolfgang Winkler
Geolayer Constanst	Stadtplan, WMIS	10	Phase 20.6.2016	MA 48		A	38		Stefan Schmalzberger
Geolayer Fahrschulen Wien	Stadtplan, WMIS	10	Phase 20.6.2016	MA 65	Nein		26	Nein	Martin Hoffmann
Geolayer Vorkommnisse (Hochwasser - Kulturlag.)	Stadtplan, WMIS	10	Phase 20.6.2016	MA 7			32		Barbara Kirsch
Geolayer Kunstwerke im öffentlichen Raum - Kulturlag.	Stadtplan, WMIS	10	Phase 20.6.2016	MA 7			32		Barbara Kirsch
Geolayer Biotopkartierung	Stadtplan, WMIS	10	Phase 20.6.2016	MA 22	JA	B	33	JA	Dr. Klaus Strohm

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“)

41 www.ogdcockpit.eu

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.
- Erbstößer, 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.

