

Open Data und GovData – ein Überblick

Was heißt „Open Data“?

International anerkannt ist die Definition von opendefinition.org, die besagt, dass „offene Daten [Daten sind], die von jedermann frei benutzt, weiterverwendet und geteilt werden können - die einzige Einschränkung betrifft die Verpflichtung zur Nennung des Urhebers.“

Der Begriff der Daten bezeichnet reine Fakten – unabhängig von Bedeutung, Interpretation und Kontext (vgl. z.B. [Begründung zu § 12 eGovG](#)). Dabei sind vor allem solche Fakten relevant, die in strukturierter Form vorliegen oder vorliegen könnten – also Statistiken, Geodaten, Haushaltsdaten, Adress-/Kontaktlisten oder Organigramme. Unstrukturierte Informationen wie Vermerke, Akten, Studien, Berichte oder andere Fließtexte fallen nicht darunter. Gleichwohl gilt: Sollten solche Texte auf (strukturiert darstellbaren) Daten beruhen, sind diese „Grundlagen-Fakten“ durchaus relevant.

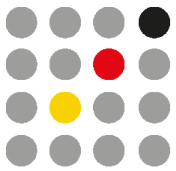
Um eine Nutzung der Daten zu vereinfachen, sollten dabei verschiedene Rahmenbedingungen eingehalten werden: So sollten die Daten maschinenlesbar und -verarbeitbar sein. PDF-Dateien oder Webseiten, über die manuell Datenbank-Einträge abgerufen werden können, machen eine Weiterverwendung zum Beispiel in anderen IT-Anwendungen unnötig schwer. Darüber hinaus sollten die Daten ohne Registrierung oder ähnliches zugänglich und geldleistungsfrei nutzbar sein. Sie sollten weiterhin aktuell, vollständig und unbearbeitet sein, soweit rechtliche Regelungen dagegen sprechen. Einfache Nutzungsbestimmungen, zum Beispiel auf Basis der für diesen Zweck durch das Bundesministerium des Innern entwickelten [Datenlizenz Deutschland](#) machen es dem Nutzer leichter zu erkennen, was mit den Daten gemacht werden darf. Auch die Auffindbarkeit der Daten spielt eine Rolle: Ein zentraler Katalog wie [GovData](#) hilft dabei, Daten zu finden.

Der Nutzen von Open Data

Die offene Bereitstellung von Daten der Verwaltung kann in ganz verschiedenen Bereichen einen Nutzen haben. Einen guten Überblick über die Nutzergruppen und den jeweiligen Nutzen bietet zum Beispiel die Studie [„Digitales Gold – Nutzen und Wertschöpfung durch Open Data für Berlin“](#) (Kap. 2.3, S. 14ff.). Im Kern lassen sich drei verschiedene Bereiche zusammenfassen.

Transparenz

Das Schaffen von mehr Transparenz im Sinne der *accountability* von Regierung und Verwaltung ist insbesondere seitens der zivilgesellschaftlichen Akteure ein Kernziel von Open Data. Die Veröf-



fentlichung von Haushaltszahlen im Ist und Soll, von Aktenplänen, IFG-Statistiken und anderen Daten kann dazu beitragen, das Verwaltungshandeln überprüfbarer und nachvollziehbarer für den Außenstehenden zu gestalten und damit demokratiefördernd zu wirken. Klar ist hier aber auch: Transparenz kann nicht umfassend sein, sondern muss immer auch Raum lassen für interne Diskussions- und Entscheidungsprozesse. Gleichwohl zeigen empirische Studien wie zum Beispiel die im Auftrag der Weltbank erstellte [„Do More Transparent Governments Govern Better?“](#), dass Transparenz und Informationszugang positiv auf die Qualität der Regierungsführung wirken, was wiederum mit einem höheren Wirtschaftswachstum einhergeht.

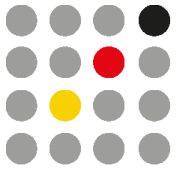
Volkswirtschaftliche Potentiale

„Daten sind das Öl der Informationsgesellschaft des 21. Jahrhunderts“, ist ein derzeit häufiger gehörter Ausspruch. Er bezieht sich oftmals auf Unternehmen wie Facebook, Google etc., die mit den Daten ihrer Nutzer Geld zu verdienen versuchen. Er gilt aber ebenso für solche Daten, die die Verwaltung aufgrund ihrer gesetzlichen oder anderer Aufträge erhebt. Es gibt mittlerweile eine Vielzahl an Studien, die das (volks-) wirtschaftliche Potential dieser Verwaltungsdaten abzuschätzen versuchen. Ein aktuelles Beispiel ist die oben erwähnte Studie [„Digitales Gold – Nutzen und Wertschöpfung durch Open Data für Berlin“](#), die alleine für Berliner Verwaltungsdaten ein jährlichen Nutzen zwischen 21,6 Mio. € und 54,0 Mio. € prognostiziert. Die vielfach zitierte [„Vickery-Studie“](#) prognostiziert für die EU einen direkten wirtschaftlichen Nutzen von bis zu 40 Mrd. €.

Verwaltungs-interner und -externer Praxisnutzen

Aus offenen Daten entstehenden IT-Anwendungen und Visualisierungen bieten einen ganz praktischen Nutzen für ihre Nutzer, so zum Beispiel die Nutzung von Verkehrs-, Wetter-, Umwelt-, Geo- oder ähnlicher Daten in mobilen *Apps*. Die Verwaltung verfügt in den verschiedensten Bereichen über Daten, die – eine einfachere Nachnutzung vorausgesetzt – auf diese Weise mit Mehrwert genutzt werden können. Im letzten Abschnitt dieser Broschüre sind einige Beispiele aufgeführt.

Auch für die Verwaltung selbst kann es nützlich sein, wenn Daten andere Behörden besser und einfacher wiederverwendbar angeboten werden. So kann es sein, dass die eigenen Daten in anderen Behörden gar nicht bekannt sind, obwohl sie dort durchaus Relevanz haben. Ebenso kann eine Entwicklung zu einer „standardmäßigen“ Daten-Veröffentlichung dazu führen, dass die Bereitstellungsprozesse insgesamt effizienter gestaltet werden können.



Um welche Daten geht es?

Einige Beispiele wurden bereits weiter oben genannt. Es gibt jedoch auch Einschränkungen, denn nicht all Daten der Verwaltung können veröffentlicht werden: Dazu zählen insbesondere personenbezogene Daten oder Daten, die die nationale Sicherheit betreffen. Ebenso können Daten, bei denen andere als die Verwaltung selbst nicht das Urheberrecht innehaben, nur im Ausnahmefall veröffentlicht werden – nämlich dann, wenn dies vertraglich vereinbart wurde.

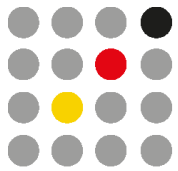
Darüber hinaus geht es jedoch um die verschiedensten Daten. So wird in einer Vielzahl von Organisationseinheiten – oftmals auch unbewusst – mit oder auf Basis von eigens erhobenen Daten gearbeitet. Dies sind auf der einen Seite Organisationseinheiten, die für zentrale und Querschnittsaufgaben zuständig sind, und auf der anderen Seite die mit Fachaufgaben betrauten Organisationseinheiten. Beispiele und Wege, um ein Bewusstsein für die eigenen Daten zu finden, sind im Artikel „Verwaltungsdaten öffnen – aber welche?“ in der [Bibliothek von GovData](#) zusammengestellt.

Die Daten-Veröffentlichung auf GovData

GovData wurde als Pilot im Auftrag des Bundesministeriums des Innern und in Abstimmung mit den Ländern entwickelt. Das Portal soll bis 2015 als Bund-Länder-übergreifend finanzierte und koordinierte *Anwendung des IT-Planungsrats* in den Regelbetrieb überführt werden.

Ziel des Portals ist es, Verwaltungsdaten aus Bund, Ländern und Kommunen an zentraler Stelle im Sinne eines „Katalogs“ auffindbar zu machen. Die Datensätze werden jedoch weiterhin von den Datenbereitstellern dezentral vorgehalten und gepflegt – üblicherweise wie bislang auch auf eigenen Webseiten oder Portalen. Grundlage für diesen so genannten *Metadatenkatalog* bildet dabei ein einheitliches Schema, in dem die Daten beschrieben werden und das Links zu den tatsächlichen Daten enthält. Dieses Schema, die *Metadatenstruktur* von GovData, stellt sicher, dass die Daten über das Portal durch die angebotenen Filter- und Suchmechanismen bestmöglich gefunden werden können.

Ein detaillierter Vorgehensplan zur Datenbereitstellung findet sich zum Beispiel in der [Bibliothek von GovData](#). Dieser bezieht sich insbesondere auf solche Daten, die bislang nicht veröffentlicht unter Berücksichtigung der im Abschnitt „Was heißt Open Data“ genannten Rahmenbedingungen veröffentlicht wurden. Sind Daten jedoch bereits in einem mehr oder minder geeigneten Format, zum Beispiel als Excel-Tabelle, veröffentlicht, so ist das Procedere relativ aufwandsarm: Es muss sichergestellt werden, dass die Nutzungsbestimmungen für diese Daten eindeutig angegeben sind. Hierfür eignen sich besonders Standard-Lizenzen wie die *Datenlizenz Deutschland*. Darüber



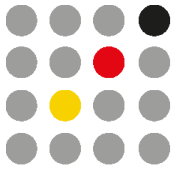
hinaus muss der Datensatz mit Hilfe der Metadatenstruktur beschrieben werden. Der einfachste Weg der Veröffentlichung auf GovData ist dann, dort einen entsprechenden Zugangs-Account einzurichten und anschließend die Datenbeschreibung über ein Webformular einzutragen. Davon ausgehend, dass die Beschreibung durch die zuständigen Fach-Experten erstellt wird, ist hier von einem Aufwand von weniger als 30 Minuten pro Datensatz auszugehen. Sollte die Daten dann zu einem späteren Zeitpunkt aktualisiert werden, so ist durch den zuständigen Datenbereitsteller nur der entsprechende Link sowie der Stand der Daten in der Beschreibung anzupassen.

Sollten mittelfristig mehr als nur ein paar wenige Datensätze eingetragen werden, lohnt es sich, eine Nutzung der zusätzlich existierenden Programmierschnittstellen zu überdenken. Sie erlauben ein automatisiertes Einspielen neuer oder aktualisierter Datenbeschreibungen; ihre Nutzung ist unter bestimmten Voraussetzungen mit relativ wenig Aufwand möglich. Weitere Informationen hierzu finden sich ebenfalls in der [Bibliothek von GovData](#).

Beispiele für die Daten-Nutzung

Im Folgenden sind einige Beispiele zusammengestellt, die dabei helfen können, ein Verständnis für die Bandbreite potentieller Daten-Weiterverwendung zu gewinnen:

- Das österreichische Datenportal [data.gv.at](#) zeigt mittlerweile über 200 Anwendungen, die auf Basis offener Daten in Österreich entstanden sind.
- Die Website [opendata500.com](#) sowie die zugehörige Studie zeigt die Nutzung offener Daten in den USA in verschiedensten Branchen und Themengebieten.
- Die offiziell über [gesetze-im-internet.de](#) bereitgestellten Bundesgesetze können auch – teilweise mit zusätzlichen Leistungen – über andere Portal wie zum Beispiel [dejure.org](#) abgerufen werden. eine Vielzahl von Apps bieten auch die Gesetzestexte auch für die Offline-Nutzung auf Smartphone oder Tablet PC an.
- Unter Nutzung der [Softwareentwicklungs-Plattform Github](#) wurde die zeitliche Entwicklung von Gesetzen nachvollziehbar gemacht. [docpatch.org](#) ermöglicht im Sinne der politischen Bildung das Nachvollziehen aller Veränderungen am deutschen Grundgesetz mit einer Vielzahl an Möglichkeiten.
- Die von der Wasser- und Schifffahrtsverwaltung veröffentlichten Echtzeit-Pegelstände der deutschen Bundeswasserstraßen werden in einer Vielzahl an mobilen Apps und Webseiten genutzt. Ähnliches gilt für Umwelt- und Wetterdaten, z.B. Ozonwerte, Luftqualität etc.



- Die Website eatsure.ca zeigt die Ergebnisse von Lebensmittelkontrollen in Reestaurants in der kanadischen Stadt London; die Bewertungsplattform yelp.com (ehemals *qype.de*) nutzt ähnliche Daten in San Francisco
- Verkehrs- und Mobilitätsdaten zu Staus, Baustellen, Benzinpreisen, Fahrplänen von Bus und Bahn etc. werden, soweit sie vorliegen, in Navigations- und ähnlichen Apps genutzt.
- Statistiken aus beliebigen Themenbereichen finden insbesondere in der Forschung, aber auch unter den *Datenjournalisten* immer mehr Nutzung. Die Maschinen-Lesbarkeit ermöglicht hier im Vergleich zu Druck-Exemplaren ganz neue Möglichkeiten der Erkenntnis-Gewinnung.

STAND September 2014

LIZENZ Diese Informationen sind vom Bundesministerium des Innern erstellt worden und stehen unter der Lizenz „Creative Commons Namensnennung 3.0 Deutschland“.

VERSION 1.0

