

Open ZDF Data

Konzeptpapier zur Vorlage im Rahmen der 3. Sitzung des Ausschusses Telemedien in der XV. Amtsperiode des Fernseh Rates am 09. Mai 2017.

Verfasser: Prof. Dr. Leonhard Dobusch, Mitglied des Fernseh Rates

Stand: 30. April 2017

Kernidee

In den letzten Jahren hat sich im öffentlichen Sektor die Erkenntnis durchgesetzt, dass öffentlich finanzierte und erstellte Datenbestände möglichst offen digital zugänglich gemacht werden sollen. Auf diese Weise sollen Transparenz, Innovation durch Dritte und ähnliche öffentliche Interessen gefördert werden. Auch für öffentlich-rechtliche Rundfunkanstalten wie das ZDF sind mit Open Data Potentiale, aber auch Aufgaben verbunden. Das vorliegende Konzeptpapier versucht sich in diesem Zusammenhang an ersten diesbezüglichen, strategischen und umsetzungspraktischen Überlegungen.

Was ist Open Data?

Unter Open Data werden online in maschinenlesbaren, nicht-proprietären Formaten für die Allgemeinheit frei nutzbare Datenbestände bezeichnet. Die freie Nutzbarkeit dieser Datenbestände wird durch die Verwendung freier bzw. offener Lizenzen¹ sichergestellt. Im Bereich von Regierungsdaten wurden diese Anforderungen beispielsweise von der Stadt Wien in Form von zehn Open-Data-Prinzipien wie Vollständigkeit, Maschinenlesbarkeit oder Diskriminierungsfreiheit operationalisiert (vgl. Tabelle 1).²

1. Vollständigkeit und Datenschutz	Von der Verwaltung veröffentlichte Datensätze sind so vollständig wie möglich, sie bilden den ganzen Umfang dessen ab, was zu einem bestimmten Thema dokumentiert ist. Metadaten, die die Rohdaten beschreiben und erklären, werden zusammen mit Formeln und Erklärungen zur Berechnung der Daten ebenfalls mitgeliefert. [...] Personenbezogene Daten sind von der Veröffentlichung grundsätzlich ausgenommen.
2. Primärquellen	Die Daten werden von der Verwaltung an ihrem Ursprung gesammelt und veröffentlicht. Dies geschieht mit dem höchstmöglichen Feinheitsgrad, nicht in aggregierten oder sonst wie modifizierten Formaten.
3. Zeitliche Nähe	Von der Verwaltung erstellte Datensätze stehen der Öffentlichkeit innerhalb eines angemessenen Zeitraums möglichst aktuell zur Verfügung. Sie werden veröffentlicht, sobald sie erhoben und zusammengestellt wurden. Daten, die in Echtzeit vorliegen, werden über eine Programmierschnittstelle (API) veröffentlicht.
4. Leichter Zugang	Von der Verwaltung veröffentlichte Datensätze sind möglichst einfach und barrierefrei zugänglich.
5. Maschinenlesbarkeit	Daten werden in etablierten Dateiformaten abgespeichert, die leicht maschinenlesbar sind (z.B. csv, xml, json), sodass eine automatisierte strukturierte Verarbeitung möglich ist.
6. Diskriminierungsfreiheit	Jede Person kann zu jeder Zeit auf die Daten zugreifen, ohne sich identifizieren oder eine Rechtfertigung für ihr Handeln abgeben zu müssen.
7. Die Verwendung offener Standards	Die Formate, in denen die Verwaltung Daten veröffentlicht sind möglichst offene Standards, über die keine juristische Person die alleinige Kontrolle hat.
8. Lizenzierung	Die Verwaltung veröffentlicht offene Verwaltungsdaten unter der Lizenz: Creative Commons Namensnennung (CC BY)
9. Dauerhaftigkeit	Von der Verwaltung veröffentlichte Informationen sind umfassend mit Metadaten dokumentiert und über lange Zeit hinweg zu finden. Einmal online gestellte Informationen werden mit angemessener Versionskontrolle versehen und dauerhaft archiviert.
10. Nutzungskosten	Durch die Festlegung der Verwendung der Lizenz: Creative Commons Namensnennung (CC BY) ist die Erhebung von Nutzungskosten derzeit nicht vorgesehen.

Tabelle 1: Gekürzte und leicht adaptierte Fassung der Open-Data-Prinzipien der Stadt Wien

¹ Für Details zum hier verwandten Verständnis von Offenheit, vgl. <http://opendefinition.org/od/1.1/de/> [29.04.2017]

² Vgl. <https://open.wien.gv.at/site/10-open-data-prinzipien/> [29.04.2017]

Warum Open Data?

In einer zunehmenden Zahl an Studien und (populär-)wissenschaftlichen Veröffentlichungen zu Open Data werden eine Vielzahl von Gründen für Engagement im Bereich Open Data angeführt (vgl. für einen Einstieg z.B. das Open-Data-Dossier der Bundeszentrale für politische Bildung).³ Ohne Anspruch auf Vollständigkeit seien deshalb nur einige der meistgenannten Gründe angeführt:

- *Öffentlich finanziert, öffentlich zugänglich*: zumindest bislang fokussierte die Open-Data-Debatte auf den öffentlichen Sektor, weshalb ein Anspruch auf Datenzugriff für die Allgemeinheit unmittelbar aus der öffentlichen Finanzierung dieser Datenbestände abgeleitet wird.
- *Transparenz und Accountability*: frei zugängliche (Roh-)Datenbestände erlauben Dritten eigenständige Auswertungen ebendieser Daten vorzunehmen und verbessern so Transparenz und Accountability öffentlich finanzierter Stellen.
- *Innovationsförderung (extern)*: das volle Innovationspotential von digitalen Datenbeständen, die teilweise ja automatisiert anfallen, lässt sich nicht alleine durch die datenspeichernde Einrichtung ausschöpfen. Open Data ermöglicht niederschweligen Zugang zu Daten für innovative Verwendungszusammenhänge.
- *Innovationsförderung (intern)*: Dem kanadischen Open-Government-Experten David Eaves zu Folge⁴ sind es häufig die verschiedenen Abteilungen und Einrichtungen der Verwaltung selbst, die von dem mit Open Data verbundenen, besseren Zugang zu (den eigenen) Datenbeständen am meisten profitieren, weil z.B. Abstimmungsaufwände entfallen.
- *Integration und Verknüpfung mit anderen Datenbeständen*: offen verfügbare

³ Vgl. <http://www.bpb.de/gesellschaft/medien/opendata/> [29.04.2017]

⁴ Vgl. Forsterleitner, C./Gegenhuber, T. (2011): Lasst die Daten frei! Open Government als kommunale Herausforderung und Chance. In: Dobusch et al. (Hrsg.): Freiheit vor Ort: Handbuch kommunale Netzpolitik. München: Open Source Press, [http://www.freienetze.at/pdfs/Forsterleitner-Gegenhuber\(2011\)Lasst-die-Daten-frei_FvO.pdf](http://www.freienetze.at/pdfs/Forsterleitner-Gegenhuber(2011)Lasst-die-Daten-frei_FvO.pdf) [29.04.2017]

Datenbestände lassen sich einfach mit anderen Datenbeständen verknüpfen und bieten damit u.a. auch völlig neue Grundlagen für wissenschaftliche Forschung oder neue Wissensplattformen (vgl. z.B. Wikidata).⁵ Ein Beispiel im Bereich des öffentlich-rechtlichen Rundfunks ist das Projekt „Research and Education Space“ der BBC.⁶

- *Dezentrale Archivierung:* Die Veröffentlichung von Daten unter offener Lizenz stellt ein Sicherheitsnetz dar, um öffentlich-rechtlich erstellte Informationen dezentral zu archivieren auch wenn für den Datenersteller die Notwendigkeit zur Depublikation auftritt. Diese Szenario wurde beispielsweise in Großbritannien für den Fall einer Rezeptdatenbank des BBC diskutiert.⁷

Sämtliche dieser Punkte treffen auch auf beitragsfinanzierten, öffentlich-rechtlichen Rundfunk zu. Daraus folgt aber unmittelbar die Frage, warum auf den verschiedensten Ebenen öffentlicher Verwaltungen – kommunaler, regionaler und (supra-)nationaler Ebene – längst (miteinander vernetzte) Open-Data-Portale (vgl. z.B. govdata.de, daten.rlp.de, daten.berlin.de) bestehen, während ein vergleichbares Angebot öffentlich-rechtlicher Rundfunkanstalten bislang fehlt.

Hinzu kommt, dass es nicht nur klassische Verwaltungsbehörden sind, die über Open-Data-Angebote verfügen, sondern längst auch öffentliche Dienstleister wie z.B. die Deutsche Bahn,⁸ die BSR⁹ oder – besonders für den ZDF Fall relevant – auch öffentlich finanzierte Kultureinrichtungen (vgl. z.B. das Projekt „Coding Da Vinci“)¹⁰.

Zusätzlich zu diesen allgemeinen Gründen, die für ein Engagement des ZDF im Bereich Open Data sprechen, gibt es auch ganz konkrete Gründe die für eine Bereitstellung von

⁵ Vgl. <https://www.wikidata.org> [30.04.2017]

⁶ <https://bbcarchdev.github.io/res/>

⁷ <https://medium.com/@digidickinson/why-the-bbc-should-make-its-recipes-database-opendata-4885e7175af2>

⁸ Vgl. <http://data.deutschebahn.com> [30.04.2017]

⁹ Vgl. <https://www.bsr.de/opendata/> [30.04.2017]

¹⁰ Vgl. <https://codingdavinci.de/daten/> [30.04.2017]

Daten im Rahmen einer ZDF-Open-Data-Strategie sprechen:

- Möglichkeit für und Bedarf von Dritten, die Serviceangebote für BeitragszahlerInnen anbieten möchten (z.B. Anbieter von Programmen wie Mediathekview¹¹ oder das Kodi-Plugin)¹².
- Verbesserte Präsenz und Korrektheit von Daten Sendungen und Sender betreffend auf Drittplattformen wie beispielsweise Wikidata.
- Basis für neue Formen von konstruktiver Mitwirkung von Beitragszahlenden jenseits klassisch-defizitorientierter Formate wie Programmbeschwerden (vgl. z.B. Initiativen wie rundfunk-mitbestimmen.de).¹³

¹¹ Vgl. <https://github.com/mediathekview/MediathekView> [27.04.2017]

¹² Vgl. <https://github.com/prof-membrane/repository.membrane> [27.04.2017]

¹³ Vgl. <https://netzpolitik.org/2017/neues-aus-dem-fernsehrat-6-ein-gespraech-ueber-rundfunk-mitbestimmen-de/> [30.04.2017]

Welche Datenbestände des ZDF sollten als Open Data verfügbar gemacht werden?

Die folgende Auflistung erhebt keinen Anspruch auf Vollständigkeit, sondern soll einen Eindruck von möglichen Datenbeständen verschaffen, die in einem Open-Data-Portal des ZDF zugänglich gemacht werden könnten:

Sendungsbezogene Daten	<p>Titel, Beschreibung, Genre, Format (z.B. Kinofilm oder Nachrichtensendung), Thema, Medium(z.B. Radio oder Internet), zugehörige Rundfunkanstalt bzw. Sender (Funk, ZDF, 1LIVE), gewünschte Zielgruppe, geographischer Bezug (z.B. Lokalzeit Bergisches Land oder terrestrische Reichweite von Radiosendungen), beteiligte Personen (z.B. Moderatoren, Schauspieler, Rollen), Website-URL, Bild-URL, Video-URL, Vorschaubild, Trailer, übliche Sendezeit, Anzahl Episoden pro Serie, Episoden einer Serie als Entitäten, Start- und Endpunkt der Sendezeit pro Episode, Beschreibung einer Episode, durchschnittliche Einschaltquote bzw. Radioquote, Einschaltquote pro Episode über die Sendezeit, Einschaltquote in relativer Reichweite und absoluten Zuschauerzahlen, Einschaltquote nach Zuschauergruppen, Online- und Mediatheknutzung, Anzahl Sendeminuten pro Monat, durchschnittliche Produktionskosten pro Sendeminute.</p>
Senderbezogene Daten	<p>Darstellung der Mittelverwendung:</p> <ul style="list-style-type: none"> - je Genre (z.B. Sport, Show, Spielfilm, Serie usw.) und Sendeplatz (z.B. Vorabendserie) – unterteilt nach Produktionskosten und /oder Lizenzkosten - je Sendungsart (z.B. „Heute Journal“, „Maybritt Illner“, „Neo Magazin Royal“) - für Eigenproduktionen, Auftragsproduktionen, Koproduktionen und Lizenzen je Genre - für Einzahlungen an die Filmförderinstitutionen - für Beraterverträge - für die Altersversorgung - für feste, fest-freie sowie freie Mitarbeiter - die an Tochterunternehmen sowie unabhängige Unternehmen abfließen, gesamt sowie je Genre <p>Auskunft über abgeschlossene Produktions- bzw. Lizenzverträge:</p> <ul style="list-style-type: none"> - a. Vertragspartner (Produzent bzw. Lizenzgeber) - b. Titel, Länge - c. Vertragssumme - d. Vertragslaufzeit - e. Gesamtbudget der Produktion (bei Auftrags- und Koproduktionen) - f. evtl. übertragener Rechteleumfang - g. Verantwortliche Redaktion bzw. Redakteur/in <p>KEF-Daten</p> <p>Allgemeine statistisch-demografische Daten (z.B. Gleichstellungs- und Diversity-Berichte, Jahresabschlüsse, Mitarbeiterbestand- und Entwicklung etc.)</p>

Tabelle 2: Exemplarische Auflistung von möglichen Datenbeständen eines Portals für Open ZDF Data

Die Datensätze sollten dabei immer auch in freien und maschinenlesbaren Formaten zur Verfügung gestellt werden (txt, csv, xml, json, etc.) im Rahmen eines zentralen und frei

zugänglichen Datenkatalogs zugänglich gemacht werden.

Hinsichtlich der Frage zur Lizenzierung der Daten, so ist die Open-Data-Idee grundsätzlich nicht an eine spezifische Lizenz gebunden und es wäre denkbar, dass das ZDF eine eigene Open Data Lizenz entwickelt, wie dies beispielsweise mit der Datenlizenz Deutschland für das Bundesportal GovData geschehen ist. Wie auch in Wien eingesetzt empfiehlt sich jedoch die Lizenz Creative Commons Namensnennung (CC BY), da diese dem internationalen Standard zu offenen Daten entspricht und somit die maximale Kombinierbarkeit der Daten gewährleistet.¹⁴

¹⁴ <http://opendefinition.org/licenses/>

Fazit und Handlungsempfehlungen

Die eingehende Auseinandersetzung öffentlich-rechtlicher Rundfunkanstalten mit dem Thema Open Data ist überfällig, es gilt die Innovations- und Profilierungspotentiale von Open Data zu nutzen. Hinzu kommt, dass eine konsequente Open-Data-Strategie (z.B. im Kontext von Sendungsdaten) auch ein weiteres Differenzierungskriterium gegenüber privaten Anbietern darstellt.

Für eine differenzierte Auseinandersetzung mit der Materie wären deshalb folgende nächsten Schritte zu empfehlen:

- Konsultation mit ExpertInnen im Bereich Open Data sowie von potentiellen Open-Data-AnwenderInnen (z.B. Mediathekview-Entwickler).
- Erhebung und Kategorisierung von Datenbeständen hinsichtlich ihrer Nutzbarkeit im Rahmen eines Open-Data-Portals.
- Ggf. Kooperation mit anderen öffentlich-rechtlichen Sendern hinsichtlich geplanter Open-Data-Richtlinien und -Angeboten.
- Ggf. Einrichtung eines Prototypen für ein Open Data Portal in Rücksprache mit Open Data Experten und Anwendern (so geschehen beispielsweise im Rahmen einer Vorstudie der Senatsverwaltung zu Open Data in Berlin).