



Offene Daten im öffentlichen Sektor: Online-Karte und Monitoringbericht (Open Data Monitor Public Sector)

Antrag auf eine Zuwendung

1. Projektdarstellung

1.1 Ziel des Vorhabens

Ziel ist es, einen möglichst vollständigen laufenden online Überblick über alle Angebote offener Daten von möglichst allen Stellen im öffentlichen Sektor in Deutschland (11.191 Gemeinden, mehrere Dutzend Landesbehörden pro Bundesland, Bundesbehörden und öffentliche Unternehmen auf allen drei Ebenen) mit einer noch näher zu bestimmenden thematischen Zuordnung und Beschreibungstiefe sowie einen jährlichen Statusbericht zu erhalten.

Damit soll u.a. der Stand der Umsetzung des Zieles größerer Transparenz durch offene Daten im Regierungsprogramm „Transparente und vernetzte Verwaltung“ transparent gemacht und eine verlässliche Datenbasis für internationale Vergleiche geschaffen werden.¹

Adressaten sind in erster Linie Studierende, Wissenschaftler und Journalisten. Es ist anzunehmen, dass das Interesse an Rohdaten bei Online-Journalisten in Zukunft noch steigen wird. Auch in Deutschland zeichnet sich eine Spezialisierung durch sogenannte Datenjournalisten ab, die wiederum dazu beitragen werden, dass die Kommunen mehr Daten bereitstellen werden.

1.2 Ausgangspunkte und Weiterentwicklungen

Das Vorhaben ist eine Weiterentwicklung des von der Open Knowledge Foundation durchgeführten Open Government Data Census und eine Ergänzung des Bund-Länder Portals GovData.

a) Open Government Data Census

¹ Die Bundesregierung weist selbst auf die Herausforderungen in einem internationalen Wettbewerb auf diesem Gebiet im Regierungsprogramm hin: „Dabei ist nicht zu übersehen, dass auch die Verwaltung in einem internationalen Wettbewerb steht. In diesem Zusammenhang ist es bedeutsam, dass die USA, Großbritannien und andere Staaten über Open Data und Open Government nicht nur diskutieren, sondern tatsächlich strukturiertes Wissen der Verwaltung für jedermann öffnen – auch mit dem Ziel, Innovationsprozesse mit allen gesellschaftlichen Gruppen anzustoßen.“ (S. 10f.).

Die Open Knowledge Foundation erstellt in bestimmten Abständen einen weltweiten Open Government Data Census. Zur Zeit wird für 60 Länder und für 34 Städte angezeigt, ob es generell Daten im Internet und speziell offene Daten zu zehn Themengebieten wie Verkehr, Haushalt und Finanzen, Wahlergebnisse, nationale Karten u.a.m. gibt (<http://census.okfn.org/country/>). Die Vorgabe der zehn Themen erscheint derzeit etwas willkürlich. Die Datenbasis in Form nicht überprüfter Angaben von „Experten aus der Community“ ist wenig verlässlich. Sie resultiert aus einem Mangel an Ressourcen für verlässlichere, aber aufwendigere Vorgehensweisen. Hier setzt das Vorhaben an, in dem es eine aussagekräftigere thematische Gliederung, eine differenziertere Erfassung der verschiedenen Verwaltungsebenen und verlässlichere Methoden der Datenermittlung entwickeln und erproben will.

Wie im Open Government Data und City Open Data Census sollen die identifizierten Angebote auch graphisch auf einer Karte, hier einer Deutschlandkarte, dargestellt werden, insgesamt oder nach Themengebieten gefiltert. (Vorbild soll hier die Karte der Bürgerhaushalte sein <http://www.buergerhaushalt.de/de/map>).

Neben dem permanent aktualisierten Online Angebot soll einmal pro Jahr ein Statusbericht erstellt werden, der die Anzahl der bereitstellenden Stellen zur Gesamtzahl der verschiedenen Arten von öffentlichen Stellen in Relation setzt und damit die Berechnung eines Index der Offenheit des öffentlichen Sektors insgesamt und nach Verwaltungsebenen und Themengebieten differenziert ermöglicht.

b) GovData

GovData ist ein Bund-Länder-Portal, das im Oktober 2013 über 4.300 Datensätze von einer nicht genannten Anzahl von Stellen auf Bundes- und Länderebene sowie einigen Kommunen zum unmittelbaren Download bereitstellt. Metadaten, die eine Auswahl unterstützen, betreffen das Themengebiet, Lizenzbedingungen, technische Formate u.a.m. Das Angebot entsteht aufgrund von individuellen Meldungen und Registrierungen der datenhaltenden Stellen. Es ist keine aktive Aufforderung von Landesbehörden und Kommunen vorgesehen, um die Vollständigkeit der Angebote zu erhöhen. Daher gibt es dafür auch keine entsprechenden Prozesse.

Das Vorhaben will genau solche Prozesse entwickeln und erproben: Ausgehend von vollständigen Listen öffentlicher Stellen (Gemeindeverzeichnis, Listen der Landesbehörden pro Bundesland) soll zum einen eine Suchmaschine zum (teil-)automatisierten Auffinden von Datenangeboten auf den Webseiten der öffentlichen Stellen entwickelt und erprobt werden und ergänzend ein E-Mail-Formular für eine halbjährliche Erhebung verschickt werden. Identifizierte Angebote sollen mit den für eine erste Einordnung minimal erforderlichen Deskriptoren beschrieben werden. Dies ist neben dem Namen eine Zuordnung zu einer mindestens einstufigen, maximal zweistufigen Gliederung der Themengebiete sowie die URL der jeweiligen Fundstelle auf den Seiten der Bereitsteller.

Im Gegensatz zu GovData sollen von der zu entwickelnden Plattform keine Daten heruntergeladen werden. Interessenten sollen lediglich feststellen, wo es (offene) Daten zu einem bestimmten Thema gibt und sich alle weiteren Informatio-

nen über die angegebenen Links erschließen. Durch diese Minimierung der Angaben, die geringe Tiefe, erscheint es möglich, eine große Breite des Verzeichnisses zu erreichen. Daher konkurriert dieses Angebot nicht mit GovData, sondern ergänzt es um die Hinweise auf Datensätze, die dort (noch) nicht registriert sind und erspart Interessenten die Mühe, in die vielen einzelnen Webseiten selbst zu suchen, wenn sie z.B. Daten auf kommunaler oder Länderebene vergleichen wollen.

1.3. Nutzer und Nutzungsszenarien

Als Hauptnutzer werden hier weniger App Entwickler gesehen, sondern Studierende, Wissenschaftler und (Online-/Daten-)Journalisten. Beispiele für erwartete Nutzungen, an denen auch die thematische Gliederung ausgerichtet werden soll sind:

Studierende suchen für ihre Bachelor-, Master oder Doktorarbeit

- im Fach VWL Daten über die Entwicklung kommunaler Investitionen im Verhältnis zu Steuereinnahmen,
- im Fach Verkehrswissenschaft oder Stadtplanung Daten über kommunale Lärmaktionspläne,
- im Fach Ernährungs- und Gesundheitswissenschaften Prüfergebnisse lokaler Lebensmittelhygienekontrollen,
- in den Politikwissenschaften Daten zur Entwicklung der Wahlbeteiligung bei Kommunalwahlen in den alten und neuen Bundesländern,
- in den Rechtswissenschaften Verträge über die Vergabe von Entsorgungsleistungen,

(Online-)Journalisten suchen

- im Zusammenhang mit Recherchen zum Rückkauf privatisierter Versorgungsbetriebe Daten über die unterschiedlichen Rechtsformen und Eigentumsanteile kommunaler Energieversorger,
- im Zusammenhang mit Lebensmittelskandalen Daten der Gesundheitsämter,
- nach Kommunalwahlen Daten über Wahlergebnisse nach Stimmbezirken und deren Bevölkerungs- und Einkommensstruktur.

Die Interessenten suchen diese Daten für möglichst viele Kommunen, um Vergleiche anstellen zu können. Das geplante Portal liefert dafür keine verlässliche endgültigen Listen, sondern Links zu den Kommunen, wo sich eine genauere Suche lohnt, gewissermaßen eine Long List, die dann näher geprüft werden muss. Diese erspart jedoch die eigene Suche auf den Seiten der über 11.000 Gemeinden.

2. Arbeitspakete und Meilensteine

Die Entwicklung einer solchen Plattform ist nicht trivial. Es hat seine Gründe, dass GovData und der Zensus Schwächen haben. Diese liegen vor allem darin begründet, dass Qualität einen hohen Aufwand erfordert. Maßgeblich für eine

erfolgreiche Nutzung ist eine möglichst hohe Vollständigkeit und Korrektheit der bereitgestellten Daten. Daraus ergibt sich eine doppelte Herausforderung. Zum einen muss ein Erhebungsverfahren entwickelt werden, das eine hohe Akzeptanz bei den Daten haltenden Stellen findet und gleichzeitig mit einem vertretbaren Aufwand durchgeführt werden kann. Dies wird nicht in einem Vorgang gelingen, sondern erfordert Nacherhebungen. Zum anderen muss eine Qualitätssicherung der gemeldeten Daten sowie der Klassifizierung vorgenommen werden. Denn nicht alle Daten, die in einer Umfrage gemeldet werden, erfüllen die Anforderungen an Offenheit und die Zuordnung zu Themengebieten durch die Daten haltenden Stellen muss überprüft werden.

Wenn eine möglichst vollständige und qualitätsgesicherte Datenbasis geschaffen worden ist, muss in einem zweiten Schritt deren Nutzbarkeit optimiert werden. Dies erfolgt durch eine evolutionäre Systementwicklung mit Expertenreviews und Nutzertests in Bezug auf die inhaltliche Klassifizierung und die Dialogführung.

Da keineswegs sicher ist, dass der Aufbau einer möglichst vollständigen und qualitativ hochwertigen Datenbasis in einem definierten Zeitraum gelingt, sondern es durchaus zu Verzögerungen kommen kann, soll das Projekt in zwei Phasen gegliedert werden, und zunächst eine Förderung der ersten Phase der Systementwicklung beantragt werden, die mit einem Prototyp abgeschlossen wird. Wenn sich der erfolgreiche Abschluss dieser ersten Phase abzeichnet, soll die Förderung der zweiten Phase, der Erprobung und Nutzungsoptimierung beantragt werden. An diese Erprobungsphase schließt sich dann drittens die Phase des Regelbetriebs an, für die keine Förderung mehr beantragt werden wird.

Im Einzelnen beinhalten die beiden Phasen folgende Arbeitspakete

1. Entwicklung eines Prototypen der Plattform:

- Aufbau der Datenbasis mit Suchmaschine,
- Datenmodell
- Dialog und Präsentation

2. Erprobung und Nutzungsoptimierung

- Expertenreview und Nutzertests
- Überarbeitung des Prototypen
- Erster Monitoringbericht
- Planung für Dauerbetrieb

3. Projektorganisation

Das Vorhaben soll in enger Zusammenarbeit zwischen dem Institut für Informationsmanagement Bremen GmbH (ifib) und der Open Knowledge Foundation Deutschland durchgeführt werden. Organisatorisch liegt die Projektleitung dem ifib. Die ifib GmbH ist Zuwendungsempfänger und gibt einen Teil der Zuwendung an die OKF als Zuwendungsletztempfänger.

3.1 Antragsteller

Institut für Informationsmanagement Bremen GmbH

Forschungsinstitut an der Universität Bremen

Am Fallturm 1

28359 Bremen

Tel.: +49/(0)421/218-56571

<http://www.ifib.de>

Gerichtsstand: Amtsgericht Bremen, HRB 21271

Geschäftsführer: Prof. Dr. Andreas Breiter

Die ifib GmbH ist eine hundertprozentige Tochter des Verein zur Förderung der wissenschaftlichen Forschung in der Freien Hansestadt Bremen e.V.(VFwF e.V.)

<http://www.vfwf.uni-bremen.de/>

Verantwortlich für dieses Projekt: Prof. Dr. Herbert Kubicek

Referenz: Aufbau und Pflege einer bundesweiten Datenbank der öffentlichen Internetzugangsorte mit über 7.000 Einrichtungen, inzwischen betrieben von der Stiftung Digitale Chancen, die aus diesem Projekt hervorgegangen ist (<http://www.digitale-chancen.de/einsteiger/>)

Kompetenz: Erfahrungen mit der Strukturierung von Internetangeboten (z.B. Informationsportal der Freien Hansestadt Bremen), Begleitung der Umsetzung des Hamburger Transparenzgesetzes sowie Tests von Internetangeboten auf Nutzbarkeit und Nützlichkeit sowie Barrierefreiheit

Perspektive: Das ifib bereitet einen Antrag bei der Deutschen Forschungsgemeinschaft vor mit dem Titel „Der Beitrag von Open Government Data zur Verwaltungstransparenz - Konzeptionelle und methodische Grundlagen zur Bestimmung des Entwicklungsstandes und seiner Einflussgrößen“. Darin soll aufbauend auf den auf der hier beantragten Ersterhebung eine qualitative Analyse der hemmenden und fördernden Faktoren der Bereitstellung offener Daten erfolgen sowie die zweite und dritte Erhebung (Fortschrittskontrolle) finanziert werden.

3.2 Partner

Open Knowledge Foundation Germany e.V.

Gneisenaustraße 52

D-10961 Berlin

Tel.: +49 30 577 036 66 2

<http://www.okfn.de>

Kompetenz: Programmierung der Suchmaschine, Aufbereitung der Daten, Programmierung der Datenbank, Userinterface Design

3.2 Kooperationspartner

Um die Resonanz auf der kommunalen Ebene zu erhöhen, soll das Vorhaben über die kommunalen Spitzenverbände (Deutscher Städtetag, Städte- und Gemeindebund und Landkreistag) auf Bundes- und Landesebene bekannt gemacht werden. Es wird auch eine Unterstützung bei der Umfrage angestrebt. Die drei Verbände sind im Beirat des ifib vertreten. Eine Zusage des Deutschen Städte- und Gemeindebundes liegt bereits vor.

Mit dem Referat O1 im BMI als Betreiber von GovData wurde bereits eine enge Kooperation und Abstimmung vereinbart.

4. Aufwand

Phase 1: Entwicklung eines Prototypen Mai – Dezember 2014			
		ifib	OKF
1	Erfassung der öffentlichen Stellen mit URLs und Mail-Adressen	0,5 PM	0,5 PM
2	Erprobung einer ersten Themengliederung an 100 identifizierten Angeboten auf kommunaler Ebene	1 PM	0,5 PM
3	Iterative Entwicklung einer Suchmaschine	0,5 PM	2,0 PM
4	Einsatz der Suchmaschine und manuelle Bearbeitung der Treffer		2,0 PM
5	Ergänzende manuelle Suche und Abgleich (Qualitätssicherung)	1,0 PM	
6	Abwägung ein- oder zweistufige thematische Gliederung in Abhängigkeit vom Zuordnungsaufwand	0,5 PM	0,5 PM
7	Umfrage bei Kommunen und Landesbehörden (inkl. Nachfassen)	1,0 PM	
8	Prüfen und Einpflegen der gemeldeten Datensätze	1,0 PM	
9	Prototyp der Datenbank mit Kartendarstellung und thematischer Filterung	0,5 PM	1,5 PM
10	Projektkoordination	1,0 PM	
	SUMME 1	7,0 PM	7,0 PM