



Stellungnahme

des Wikimedia Deutschland, Gesellschaft zur Förderung Freien Wissens e. V. zu Use Cases für das Dateninstitut

Berlin, den 14.01.2023

Wikimedia Deutschland setzt sich für Chancengleichheit beim Zugang zu Wissen und Bildung ein und unterstützt die vielen tausend Ehrenamtlichen, die z. B. in Wikidata und der Wikipedia tagtäglich ihr Wissen mit allen Menschen teilen.

Wir begrüßen daher die Bestrebungen der Bundesregierung, das Wissen um eine zukunftssichere und nachhaltige Datenhaltung zu mehrten und allen Gliederungen des Bundes zugänglich zu machen. Wir erhoffen uns, dass die öffentliche Hand damit besser in die Lage versetzt wird, Informationen sowohl intern besser zu nutzen als auch mehr Informationen gemäß des Grundsatzes „Öffentliches Geld – Öffentliches Gut!“ der Allgemeinheit als Allmende zur Verfügung zu stellen.

Bei den bisher vorgestellten Use Cases für das geplante Dateninstitut fehlt uns der Einbezug der vielfältigen bereits gemachten Erfahrungen des Digitalen Ehrenamts beim Umgang mit Daten. Das zeigt sich besonders bei Use Case aa) „Kommunale Mobilitätsdatenplattform“. Statt ein weiteres Pilotprojekt mit den gleichen Ergebnissen anzuschleppen, **sollte das Dateninstitut die bereits gewonnenen Erkenntnisse – u.a. zu immer wieder auftretenden Hürden – aufbereiten und aus ihnen Handlungsempfehlungen entwickeln, wie diese Hürden systematisch abgebaut werden können.**

Die Use Cases bb) und cc) kranken vor allem daran, dass sie komplexere gesellschaftliche Herausforderungen in ein Datenproblem zu transformieren versuchen. So scheitern referenzbasierte politische Entscheidungen zwar auch an Datenverfügbarkeit, doch primär am gesellschaftlichen Konsens darüber, inwiefern der Staat Umverteilung betreiben soll und, im konkreten Fall, ob bestimmte Gruppen besonders vor Preiserhöhungen geschützt werden sollen. Wir stehen zudem den immer wieder in den vorgestellten Use Cases anklingenden Bestrebungen zur Vermarktung von Fakteninformationen und der damit notwendigerweise verbundenen künstlichen Herstellung von Exklusivität von Daten durch technische Mittel kritisch gegenüber.

Wir **empfehlen stattdessen einen Fokus auf Linked Data und Linked Open Data** bereits bei der Datenorganisation und -Speicherung. In der Demonstration der praktischen Anwendung von Datenarchitekturen mit einem strategischen Blick auf

**Wikimedia Deutschland –
Gesellschaft zur Förderung
Freien Wissens e. V.**

Tempelhofer Ufer 23 – 24
10963 Berlin

Postfach 61 03 49
10925 Berlin

Tel.: +49 (0) 30 219 158 26 - 0
Fax: +49 (0) 30 219 158 26 - 9
wikimedia.de
politik@wikimedia.de

Spendenkonto:
IBAN: DE05 1002 0500 0003
2873 00
BIC: BFSWDE33BER

Geschäftsführender Vorstand:
Dr. Christian Humborg

Eingetragen beim Amtsgericht
Berlin-Charlottenburg unter der
Nummer 23855 B

Linked-Data-Prinzipien vor allem für die öffentliche Hand sehen wir den wichtigsten Beitrag, den das Dateninstitut derzeit leisten kann.

Zu Use Case aa) „Kommunale Mobilitätsdatenplattform“

Der Text zum Use Case aa) „Kommunale Mobilitätsdatenplattform“ bezieht sich auf die existierende freie/Open-Source-Software „digitransit“, die ursprünglich auf Basis des ebenfalls freien Routenplaners OpenTripPlanners in Helsinki entwickelt und 2017¹ durch Freiwillige des digitalen Ehrenamts auf Deutsch übersetzt wurde.

Seither wurde digitransit im Rahmen von Förderprojekten unter anderem in den Kommunen [Ulm](#) und [Herrenberg](#) (als „stadtnavi“) sowie [in Brandenburg](#) (als „bbnavi“) praktisch umgesetzt, auch um die notwendigen Datenflüsse zu erkunden. Dabei entstand [Dokumentation](#) für die praktische Umsetzung von digitransit-Installationen. Schon vorher existierten praktische Demonstratoren, wie regelmäßige Mitfahrgelegenheiten oder später auch „[Mitfahrbänke](#)“ in diese Routenplanung für intermodales Tür-zu-Tür-Routing über den reinen ÖPNV hinaus integriert werden können. Die in der Kurzbeschreibung des Use Case umrissenen Ziele zur Datendarstellung oder Routenplanung existieren bereits allesamt als fester Teil von digitransit oder wurden in den beschriebenen Projekten umgesetzt bzw. dokumentiert.

Der praktische lokale Betrieb der in Helsinki bereits seit Jahren funktionierenden digitransit-Lösung und die dafür nötige Datenintegration stieß jedoch sowohl in den bisher geförderten Projekten als auch bei den ehrenamtlichen Bestrebungen auf wiederkehrende strukturelle Hürden. Einen Teil davon haben die Ehrenamtlichen bereits [in einer Liste von Mustern](#) (siehe Einleitung) festgehalten. Viele dieser Probleme haben ihren Ursprung darin, dass die Bereitstellung der notwendigen Daten vor allem durch die öffentliche Hand und die Verkehrsverbände bis heute vielfach als Mehraufwände verursachender vermeintlicher Zusatzdienst an Dritten verstanden wird. Stattdessen müsste vor allem die öffentliche Hand bei der internen Datenhaltung und -verwendung auf eine möglichst hohe Wiederverwendbarkeit setzen und auch die eigene Datennutzung auf diese Prozesse aufbauen. Beispielsweise wären Kommunen bereits heute in der Lage, Informationen zu verkehrsrechtlichen Anordnungen rund um Baustellen im Stadtgebiet automatisiert aus dem zugehörigen Fachverfahren zu exportieren. Diese Informationen kann nicht nur zur Planung der notwendigen Verkehrszeichen verwendet werden, sondern auch automatisiert für Routenplanung oder die Anordnung von Abstellrestriktionen für Sharingfahrzeuge verwendet werden. Solange aber bereits behördenintern die Möglichkeiten der Fachverfahren gar nicht verwendet werden, weil die notwendigen Voraussetzungen für die interne Wiederverwendung der Informationen gar nicht gegeben sind, scheitert auch die externe Verwendbarkeit.

Eine erneute digitransit-Instanz wird lediglich erneut zu diesen Erkenntnissen kommen und kaum zusätzliche generieren. Teilweise liegt das Problem auch in den rechtlichen Rahmenbedingungen für das Teilen und Bereitstellen offener Daten. So ist es in den USA bereits seit Jahren gelebte Praxis, dass aufgrund kommunalen Satzungsrechts

¹ In der verschickten Stellungnahme stand hier 2018. Richtig ist das Jahr 2017.

Sharinganbieter zur Bereitstellung der notwendigen Daten [in längst standardisierten Formaten verpflichtet werden](#). In Deutschland ist noch nicht abschließend geklärt, ob die Kommunen die Kompetenzen für solche Regelungen haben. Sowohl Initiativen auf Länderebenen als auch das anstehende Mobilitätsdatengesetz streben derzeit an, die rechtlichen Voraussetzungen für eine bessere Verfügbarkeit dieser Daten zu schaffen. Der Input des Dateninstituts käme hierfür jedoch viel zu spät, wenn es wie im Use Case umrissen erst erneut die bereits gemachten Erfahrungen von neuem machen will.

Zudem besteht bereits Konsens, dass ein – wie im Use Case als Ziel ausgewiesener – eigenständiger Betrieb durch eine Kommune aufgrund von Skalierungseffekten nicht sinnvoll erscheint, sondern auf überregionaler Ebene beispielsweise durch die Aufgabenträger oder im Verbund von Gebietskörperschaften erfolgen sollte. Der Aufbau von Daten- und IT-Architekturkompetenzen auf kommunaler Ebene ist zwar dringend notwendig. Ihn mit einer Kommune anhand des Beispiels digitransit durchzuspielen, ohne die bisherigen Hürden *vorher* analysiert zu haben und die Beteiligten auf sie und die dafür nötigen Infrastrukturinvestitionen gezielt vorzubereiten, erscheint jedoch geradezu grausam gegenüber den ausführenden Personen.

Empfehlung: Fokus auf Linked (Open) Data und Wiederverwendbarkeit

Das größte zu hebende Potenzial, dem sich das Dateninstitut widmen sollte, ist die Aufbereitung und Zugänglichmachung von Linked Data und den dafür notwendigen IT- und Datenarchitekturentscheidungen – sowohl für die öffentliche Hand als auch für viele andere Bereiche, einschließlich der Wirtschaft. Wikimedia nimmt mit dem auf Linked-Open-Data-Prinzipien beruhenden Projekt [Wikidata](#) bereits seit über 10 Jahren eine Vorreiterrolle bei der praktischen Anwendung dieser Konzepte ein. Die Software [Wikibase](#), auf der Wikidata beruht, wird bereits von vielen Akteuren in der Wissenschaft, im Bibliothekswesen und für Projekte der Europäischen Union eingesetzt. Die Europäische Union setzt strategisch großes Gewicht auf die Wiederverwendbarkeit von Informationen, was sich in den einschlägigen Verordnungen klar widerspiegelt.

Die Organisation von Informationen in sogenannten Knowledge Graphs nach Linked-Data-Prinzipien ist eine essentielle Voraussetzung zur Umsetzung dieser Wiederverwendbarkeit. Knowledge Graphs ermöglichen die in den vorgeschlagenen Use Cases immer wieder geforderte Verknüpfung von Informationen aus verschiedenen Quellen, um daraus größere Zusammenhänge erkennbar zu machen und zu analysieren. Sie basieren auf etablierten Standards und ermöglichen die Verwendung von Informationen in anderen Kontexten. Wikidata ist als Wissensgraph beispielsweise in der Lage, [das Durchschnittsalter aller Mitglieder des 20. Deutschen Bundestags](#) tagesaktuell auszugeben – und wer dem eigenen Sprachassistenzsystem diese Frage stellt, bekommt die Antwort in der Regel ebenfalls aus Wikidata. Das Linked-Data-Prinzip erlaubt auch eine Abkehr vom architekturell längst überholten Ansatz, Datensätze erst auf zentralen Plattformen (Hubs, Datenräumen, etc) zusammenführen zu müssen, um sie in gemeinsamen Zusammenhang bringen zu können. Stattdessen kann und soll mit diesem Prinzip eine Vielzahl dezentraler Datenendpunkte abgefragt und ihre Daten gemeinsam in Bezug gebracht werden – genau

so, wie auch das World Wide Web als Sammlung dezentraler Informationsquellen anstatt zentraler Kataloge gedacht war, nur eben nun automatisiert maschinenlesbar.

Die Erfahrung der Ehrenamtscommunities rund um u.A. Wikidata oder Code for Germany zeigen, dass sich vor allem die öffentliche Hand bislang schwer damit tut, die Prinzipien von Linked Data in ihrer alltäglichen Arbeit zu operationalisieren. Derweil würde gerade die öffentliche Hand am Meisten davon profitieren, wenn sie beispielsweise behördenintern statistische Informationen über Stadtquartiere automatisiert mit Raum- und Mobilitätsdaten zusammenführen könnte, um Auswirkungen verkehrsplanerischer Vorhaben zu analysieren. Die hierfür notwendigen Daten sind in der Regel bereits vorhanden, müssen bislang aber mit großem Aufwand oder durch externe Dienstleister händisch zusammengeführt und ausgewertet werden. Eine Datenorganisation nach Linked-Data-Prinzipien würde diese Zusammenführung und Auswertung automatisiert und durch die Verwaltung selbst ermöglichen.

Dass der Teil der Informationen, die unbedenklich als Open Data veröffentlicht werden können, als Ergebnis dieser IT- und Datenarchitektur quasi en passant und automatisch als Nebenprodukt herausfallen, ist ein weiteres gewichtiges Argument.

Auch hier empfehlen wir den Austausch mit den Beteiligten des Digitalen Ehrenamts, die sich häufig seit vielen Jahren auf hohem Niveau mit diesem Thema beschäftigen. Es ist der Natur dieses Ehrenamts geschuldet, dass dieses meist nicht formell organisierte Engagement nicht so umfassend dokumentiert wird. Wie wir bereits [in unserem gemeinsamen Politikbrief mit der Deutschen Stiftung für Ehrenamt und Engagement](#) dargestellt haben, ist es umso wichtiger, die Erfahrungen der Ehrenamtlichen systematisch zu dokumentieren und für Digitalisierungsvorhaben nutzbar zu machen.