

Linked Data

Datenzentriertes Arbeiten mit Knowledge Graphen

Daten sind die Basis informierter Verwaltungsentscheidungen. Um diese Daten nachhaltig zu bewirtschaften und zu nutzen, ist es nötig, dass sich die Verwaltung von einem spezifisch problembasierten Vorgehen hin zu einer datenzentrierten Vorgehensweise weiterentwickelt.

Wenn Daten im Hinblick auf die Beantwortung einer spezifischen Frage respektive zur Lösung eines konkreten Problems gesammelt und aufbereitet werden, wird das als problemspezifisches Vorgehen bezeichnet. Ein solches Vorgehen führt direkt zur Lösung des Problems, verhindert aber typischerweise die weitere Verwendung dieser Daten zur Lösung anderer Probleme, da der Kontext der Daten in der Regel fehlt. Dies ist beispielsweise der Fall, wenn ein Datensatz zu den politischen Gemeinden der Schweiz nur aus einer Liste von Gemeindefürnamen besteht. Viel aussagekräftiger wäre dagegen, wenn ausserdem die Zugehörigkeit der Gemeinden zu Bezirken und Kantonen und die historisierte Entwicklung der Gemeinden durch Fusionen und Kantonswechsel aufgeführt sind. Solche mit Kontext angereicherte Daten werden als «Knowledge Graph» bezeichnet. In einem Knowledge Graph werden definierte Verbindungen zwischen einzelnen Datenpunkten erstellt, was zu einem dichten Netz von Datenpunkten mit dem dazugehörigen Kontext führt. Ein datenzentriertes Arbeiten zeichnet sich nun dadurch aus, dass Daten nicht problemspezifisch, sondern möglichst kontextreich in Knowledge Graphen erfasst und verarbeitet werden. Dies ermöglicht eine universelle Weiterverwendung der Daten (auch für Hackathons, S. 44, oder für Visualisierungen, S. 42).

Linked Data ermöglicht ein iteratives Vorgehen

Es scheint offensichtlich, dass ein vollständiger und allumfassender Knowledge Graph realistischerweise kein vollständig zu erreichendes Ziel ist. Darum sollten die Anstrengungen der Verwaltung darauf hinzielen, kleine Schritte weg von einem problemzentrierten hin zu einem datenzentrierten Vorgehen zu machen. Dafür eignet sich der Einsatz von Linked Data ganz besonders. Linked Data ermöglicht es, den grossen Knowledge Graphen in kleinen Schritten von unterschiedlichen Seiten her beginnend aufzubauen und diese Verästelungen durch Hinzufügen neuer Verbindungen langsam zusammenwachsen zu lassen. Linked Data erlaubt aufgrund der Flexibilität auch spätere, mit Sicherheit notwendige Anpassungen an den Strukturen der Daten vorzunehmen, ohne den Knowledge Graphen komplett neu aufbauen zu müssen. Sinnvollerweise werden zuerst Daten als Knowledge Graph bereitgestellt, die in sehr vielen anderen Datensätzen eine Rolle spielen – wie beispielsweise die oben erwähnten Gemeinden. Solche Daten werden auch als Basisdaten, Konzepte oder «Shared Dimensions» bezeichnet.

Data Stewardship für übergreifende Koordination im Knowledge Graphen

Um dieses gemeinsame Vorgehen anzustossen und zu koordinieren, braucht es den Aufbau von Data Stewardship innerhalb der Verwaltung. Die Aufgabe von Data Stewards (vgl. «Data Governance und Open Government Data», S. 38) ist es, die Brücke zu schla-

gen weg von isolierten, problemspezifischen Datensilos hin zu gemeinsam genutzten, datenzentrierten Knowledge Graphen. Somit ist Data Stewardship eine Querschnittsfunktion, die immer übergreifend tätig ist und neben technischen Aspekten sehr viel organisationales Wissen beinhaltet. Data Stewards arbeiten über Sektions-, Amts- und Departementsgrenzen hinweg und ermöglichen das schrittweise Zusammenwachsen der bisher isolierten Datenbestände. Dabei sind Data Stewards in Zusammenarbeit mit den Data Owners für die Einhaltung der datenschutzrechtlichen Grundsätze besorgt, die durch die Verlinkung von Daten an zusätzlicher Dringlichkeit und Komplexität gewinnen. Data Stewardship erfordert eine zentrale Finanzierung, weil die bisherigen Data Owner von Amtes wegen wenig Ressourcen haben, ihre Daten für unspezifische Zwecke aufzubereiten. Knowledge Graphen erfordern eine Vorinvestition, ermöglichen aber später eine Kosten- und Aufwandsreduktion, weil nicht mehr ständig neue isolierte Datenbestände aufgebaut und gepflegt werden müssen.

Unsere Empfehlungen



1. Data Stewardship aufbauen

Eine zentrale Finanzierung von organisationsübergreifenden Anstrengungen für eine datenzentrierte Vorgehensweise fördert Data Stewardship.

2. Iteratives Vorgehen in kleinen Schritten umsetzen

Eingeschlossene Informationen in Datensilos sollten Schritt für Schritt durch Knowledge Graphen abgelöst werden um deren Potenzial zur Sekundärnutzung freizusetzen.

3. Fokus auf zentrale Datensätze legen

Besonders für Daten, die vielfach verwendet werden (Stammdaten), kann sich Linked Data lohnen.

Mehr Informationen



Kontaktmöglichkeiten und weitere Informationen zu Linked Open Government Data: bfh.ch/ipst/linked-data

Kontakt



Dr. Benedikt Hitz-Gamper
Wissenschaftlicher Mitarbeiter

benedikt.hitz@bfh.ch
T +41 31 848 50 22



Tobias Böni
Wissenschaftlicher Mitarbeiter

tobias.boeni@bfh.ch
T +41 31 848 61 54