Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Bundeskanzlei BK

Transformation und Interoperabilität

Daniel Keller 18. März 2025

Abschlussbericht
Prototyp Linked Data (natürliche Person)

Datenökosystem Schweiz

Aktenzeichen:

1 Ausgangslage

Verschiedenste Gesetze schreiben vor, dass verschiedenen Informationen zu Menschen (natürlichen Personen) durch verschiedene Verwaltungsstellen geführt werden. Eine Zentralisierung solcher Registerdaten bedarf deshalb einer umfangreichen und aufwändigen Anpassung einer Vielzahl von Bundes-, Kantons- und Gemeinde-Gesetzen und Verordnungen. Aus technischer Sicht verwenden heute Gemeinden, Kantone und die Bundesverwaltung unterschiedlichste Lösungen für die Verwaltung ihrer Stammdaten, welche teilweise aus den Registern stammen oder lokal jeweils neu erhoben werden. Diese Systeme beruhen auf verschiedensten technologischen Konzepten. Ein Konsens über eine gemeinsame technische Lösung wäre daher eine grosse Herausforderung. Aktuell wird der Austausch von Daten über den Verein eCH standardisiert. Die Verwendung der Standards ist jedoch mehrheitlich den Beteiligten überlassen.

Anwender und Nutzer von Registerdaten müssen viele verschiedene Schnittstellenintegrationen vornehmen, um die notwendigen Daten zu erhalten. Für jede Schnittstelle benötigen die Anwendungen unterschiedliche Berechtigungsnachweise.

Unsere Vision ist es, mit Hilfe der Digitalisierung die öffentliche Verwaltung der Schweiz von Morgen zu gestalten. Der Austausch zwischen Behörden und der Schweizer Bevölkerung soll vereinfacht werden. Daten sind der Grundbaustein für digitale Behördenleistungen. Dafür treiben wir die Bereitstellung von verteilt (dezentral) geführten Registerdaten über föderal aufgestellte Organisationen (Bund, Kantone und Gemeinden) voran. So können die Registerdaten unter Berücksichtigung des Datenschutzes sowie geltender Gesetze genutzt und Prozesse zuverlässiger automatisiert werden. Heute sind die geführten Register unzureichend miteinander verbunden und es ist für nutzende Behörden aufwändig und teuer, sich die benötigten und verbindlichen Angaben zu Menschen zusammenzusammeln. Die technisch einfachste Lösung wäre die Zentralisierung der Daten. Dafür braucht es die Zustimmung des Bundes, der Kantone und der Gemeinden sowie umfangreiche Gesetzesanpassungen bis hin zu einer Verfassungsänderung. Genau diese «einfache» Lösung ist deshalb in der Praxis mit einem hohen Risiko des Scheiterns bereits auf Stufe der Verfassungs- und Gesetzesanpassungen verbunden.

2 Idee und Ziele

Linked Data ist ein Konzept, welches insbesondere für die öffentliche Verwaltung der Schweiz mit der föderalen Verantwortungsteilung eine grosse Wirkung zeigen kann: Linked Data ermöglicht es, Daten (Informationen zu Menschen) durch verschiedene Behörden dezentral und unabhängig der lokalen IT-Infrastruktur (z. B. SQL, PostgreSQL, etc.) zu führen und diese mittels (gegenseitigen) Verknüpfungen miteinander in Bezug zu bringen. Auf diese Weise können Daten am Ort ihrer Entstehung mit geltender Verantwortungszuteilung verbleiben, währenddem gleichzeitig diese Daten weiteren Behördenstellen für ihre gesetzlichen Aufgabenerfüllung effizient und ressourcenschonend bereitgestellt werden. Eine Gesetzesanpassung kann schlanker erfolgen als bei einem zentralen Register, weil der Ansatz von Linked Data auf Interoperabilität der Daten, d. h. die Kompatibilität unter Daten, und nicht auf Umverteilung der Verantwortlichkeiten aufbaut.

Es soll ein Prototyp auf Basis des Konzepts Linked Data gebaut werden. Verteilt geführte Informationen zu Menschen können damit dezentral verwaltet und gemeinsam genutzt werden.

Das übergeordnete Ziel der Prototypen-Durchführung ist der Erkenntnisgewinn zum Einsatz von Linked Data für die Registerdaten der natürlichen Person und für das Datenökosystem bei der Bundeskanzlei sowie bei den Verwaltungsmitarbeitern aller föderaler Ebenen. Die Hauptfragestellung war: «Wie kann Linked Data in einem Bereich mit verteilter und sicherheitssensitiver Datenhaltung eingesetzt werden?» Dazu wurden fünf Ziele formuliert:

Ziel 1	Der Prototyp soll **Erkenntnisse** zum Umgang mit geschützten und offenen Daten bei der Verwendung von Linked Data aufzeigen. Benötigt es eine mehrstufige Behandlung aufgrund detaillierterer Informationen oder reicht bei offenen Daten eine direkte öffentliche Publikation.
Ziel 2	Der Prototyp soll detailliert aufzeigen, wie **Nutzende die Information** (aus unterschiedlichen Quellen) **finden** und in deren Anwendung einfach und unter Einhaltung der Berechtigungen verwenden können.
Ziel 3	Der Prototyp soll die Nutzung bestehender **Vocabularies und Standards** prüfen. Ein Mapping von eCH-Standards auf EU-Core Vocabularies und allenfalls weitere bereits beschriebene Semantiken soll erstellt werden.
Ziel 4	Ein **erster Entwurf** des Datenmodells eines ausgewählten Use Cases zur natürlichen Person wird implementiert.
Ziel 5	Der Prototyp soll Erkenntnisse über benötigte **technische Lösungskomponenten** aufzeigen und Empfehlung geeigneter (und gewarteter) Open-Source Software-Produkte geben.

Tabelle 1: Ziele Prototyp

3 Organisation

3.1 Kernteam

Die Erarbeitung des Prototyps wurde von einem kleinen Team mit Spezialisten von Fachhochschulen zu Linked Data und IT Security sowie mit fachlichen Vertretern durchgeführt (siehe Abbildung 1). Das Vorhaben wurde durch die Bundeskanzlei beauftragt und gesteuert.

- BFH (Technik und Sicherheit) **[Pascal Mainini / Annett Laube]**
- HES-SO (Data Architect) **[Fabian Cretton]**
- Einwohnderdienste **[Carmela Schürmann und weitere]**, ZAS (Datenlieferanten) **[Lars Steffen]**
- Kanton Zürich (Datennutzer) **[Juan Pablo Lovato]**
- BK (Architektur) **[Daniel Keller]**
- BK (Projektbegleitung, SPOC) **[Jean-Philippe Roulet]**

Abbildung 1: Kernteam

Die Problemstellung wurde in Arbeitspakete aufgeteilt und die Verantwortungen zu den Lieferergebnissen im Kernteam aufgeteilt. Leider konnten der Kanton Zürich und die ZAS aus Zeitgründen dem Kernteam nicht über die gesamte Zeit zur Verfügung stehen. So hat der Verband der Schweizerischen Einwohnerdienste (VSED) sowie die zahlreichen Teilnehmer des Sounding Board das fachliche Know-How (Datenlieferanten und Nutzerseite) vertreten.

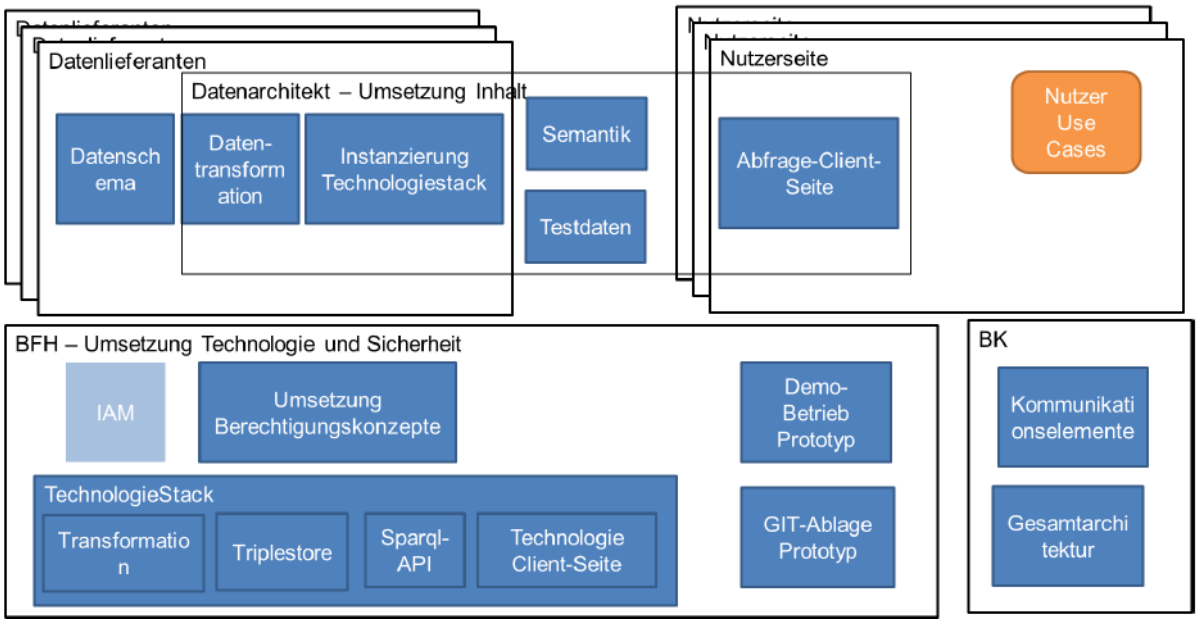

Abbildung 2: Lieferobjekte

Die Berner Fachhochschule (BFH) konnte für die Technik und Sicherheits-Aspekte gewonnen werden. Die Fachhochschule Westschweiz (HES-SO) hat die Datenarchitektur und die Dateninhalte erarbeitet.

Das Zusammenspiel der beiden Hochschulen wurde an regelmässigen Abgleichen im Kernteam koordiniert, wobei die Fachvertretung und die Bundeskanzlei aufkommende Fragen beantworten konnten. Die Zusammenarbeit der Hochschulen hat gut funktioniert und die Lieferergebnisse wurden gut aufeinander abgestimmt.

3.2 Sounding Boards für die Kommunikation gegen Aussen

Mit dem Sounding Board Format war es uns ein grosses Anliegen den Sitzungsteilnehmenden unsere Erkenntnisse mit dem Ansatz Linked Data anhand unseres Fortschrittes, Fragen und Schwierigkeiten zu berichten und zu teilen. Dabei entstand für uns die Gelegenheit die Themen/Hindernisse im Zusammenhang mit den Teilnehmern zu diskutieren und deren Hinweise in den Prototyp miteinfliessen zu lassen. Das Sounding Board war offen für alle Interessierten aus Schweizer Verwaltungen und weiteren verwaltungsnahen Institutionen. Eingeladen wurde:

- Vertreter des Nationalem Koordinationsgremium Datenmanagement
- Mitglieder der eCH Fachgruppe Meldewesen
- Linked Data Interessierte aus der Verwaltung

Die Eingeladenen wurden aufgefordert die Einladung an weitere interessierte Teilnehmer weiterzuleiten.

Ziele des Sounding Boards:

- Lernen, lernen, lernen und das Gelernte teilen
- Illustrieren der Ergebnisse, Fortschritte mit dem Prototyp
- Aufzeigen von Erkenntnissen und/oder Grenzen, Barrieren
- Diskussionen im Umgang mit Linked Data
- Äussere Einflüsse ansprechen und aufzeigen

- Die Fragen, Hürden und Anforderungen der Teilnehmer kennen zu lernen und die eine oder andere Antwort bereits zu erarbeiten.

Es wurden insgesamt sechs Sounding Board Anlässe durchgeführt (siehe Tabelle 2):

Datum	Inhaltsschwerpunkte
15.02.2024	Kick-off, Vorstellen des Prototyps und Einführung in Linked Data
22.03.2024	Vertiefung Datenmodell Linked Data, Vergleich mit EU Vocabulary, Verwendung AHV-Nummer in Testdatensätzen, Vorstellen und Input der Teilnehmer zu Use Case
25.04.2024	Vertiefung Datenmodell Linked Data am konkreten Beispiel, Überlegungen zu gewählter Semantik, Technologie-Evaluation, Alternative Architekturvarianten
13.06.2024	Abfrage-Möglichkeiten, Föderierte Abfragen, Dereferenzieren und Einordnung, wann welche Möglichkeit von Vorteil ist. Präsentation Lösung Triple-Store /(Linked Data Datenbank) mit Testdatensatz. Präsentation des Proxy-Konzepts
22.08.2024	Betrachtung der Auswirkungen auf die fachlichen Abläufe und entsprechenden Veränderungen von Prozessen, Rollenteilung um Sicherheit zu Organisieren
24.10.2024	Abschluss-Präsentation mit Rückblick auf Weg, Aufzeigen erreichter Lieferobjekte, Update zu Datentransformationen, Integration Proxy mit Triple-Store, Präsentation der Resultate, Fazit mit Herausforderungen, Nutzen und Info über die Verwendung der Resultate im Datenökosystems und über angedachte nächste Schritte für die Architektur der Registerdaten natürliche Person.

Tabelle 2: Sounding Boards

4 Vorgehen und erarbeitete Resultate

Die Problemstellung wurde mit einer iterativen Arbeitsweise angegangen. Ausgerichtet an den Zielen haben sich die notwendigen Schritte ergeben. An den Koordinations-Meetings des Kernteams wurden die verschiedenen Teilnehmer aufeinander abgebstimmt.

Die erarbeiteten Resultate und die zugehörige weiterführende Dokumentation wurden im Internet publiziert:

- https://github.com/swiss/ld-prototype-proxy
- https://github.com/swiss/ld-prototype-data

Die BFH wie auch die HES-SO haben einen technischen Abschlussbericht zu deren Beiträge zum Prototyp verfasst:

- https://www.bfh.ch/.documents/ris/2024-385.206.909/BFHID-648595275-1/Zugriffskontrolle-Linked-Master-Data.pdf
- https://github.com/swiss/ld-prototype-data/blob/main/Rapport_HES_SO_VS_Prototyp_Linked_Data_v1.0.pdf

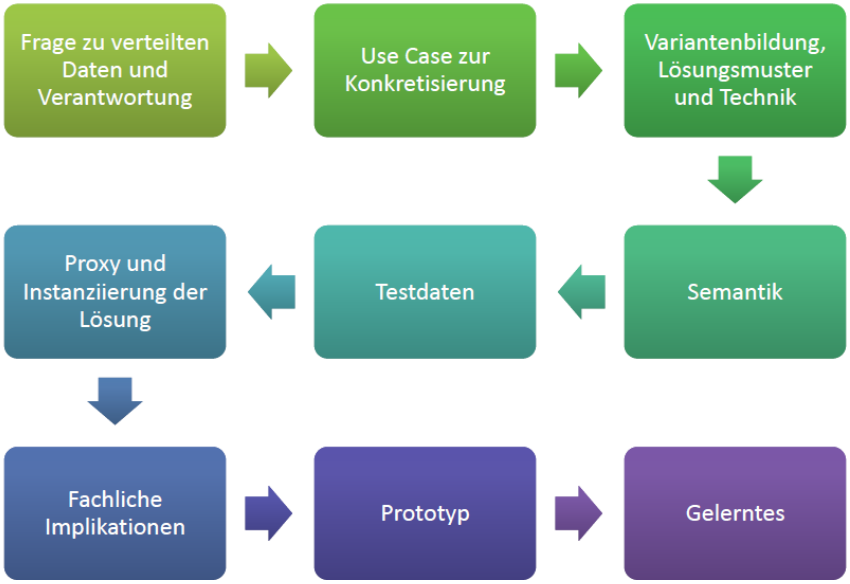

Abbildung 3: Vorgehen

Um die Problemstellung zu verstehen, wurde nach der erarbeiteten Fragestellung erst ein Use Case zur Konkretisierung gesucht. Es wurden drei Use Cases identifiziert und grob beschrieben. Die Use Cases waren «Erhebung Radio- und Fernsehgebühren», «Steuerveranlagung von Einzelunternehmen» und «Zugriff Feuerwehr». Um den Fokus in der beschränkten Zeit zu behalten, wurde die Umsetzung auf den Use Case der «Erhebung Radio- und Fernsehgebühren» beschränkt.

Auf der technischen Seite wurde erst evaluiert, welche Software für die Herausforderung gewählt werden soll. Die Idee, für alle dezentral geführten Register einen fertigen Open-Source Software-Stack (abgestimmte Kombination aus unterschiedlichen Programmen) zur Verfügung zu stellen, wurde nicht weiterverfolgt, da es keine standardisierten Sicherheitsmechanismen für den Schutz der Daten in den Triple-Stores (Linked Data Datenbanken) gab. Auf eine herstellerspezifische Zugriffsschutz-Lösung wurde bewusst verzichtet, um die Unabhängigkeit der Lösung nicht zu gefährden. Dementsprechend ist die Wahl der zugrundeliegenden Datenbank für die Zielerreichung nicht relevant.

Es wurde darauf hingewiesen, dass die herkömmlichen Daten nicht zwingend in Linked Data Triples umgewandelt werden müssen. Diese könnten auch in den bestehenden Datenbankformaten verbleiben und mittels «virtuellem Graph» (einer simulierten Linked Data Ablage) verfügbar gemacht werden kann. Eine Software nimmt dafür die Linked Data Abfragen (SPARQL-Abfragen) entgegen und wandelt dieses in herkömmliche Abfragen (zum Beispiel SQL) um. Im Prototyp wurde bewusst auf die Verwendung solcher Umwandlungen verzichtet, da wir die Erfahrung der Datenspeicherung in Triple-Stores machen wollten.

Mit dem Use Case konnte mit dem Ausarbeiten der Semantik begonnen werden. Die Prüfung der EU-Vocabularies[1] und der Vergleich zu den Standards von eCH haben ergeben, dass auf die EU-Vocabularies aufgebaut werden kann.

[1] Core Vocabularies - EU Vocabularies - Publications Office of the EU (https://op.europa.eu/en/web/eu-vocabularies/corevocs)

Da das Bundesamt für Statistik (BFS) selbst die Daten des Gebäude- und Wohnungsregister in ein Linked Data Format überführen wollte, wurde im Use Case der Radio- und Fernsehgebühren von der Modellierung dieser Daten abgesehen. Zu diesem Thema fand zudem ein Austausch mit den Mitarbeitern des BFS statt.

Es konnte kein Testdatensatz für Daten der natürlichen Person bei den Personen-Registern ausfindig gemacht werden. So wurde auf der erarbeiteten Semantik ein neuer Testdatensatz aus fiktiven Personen in Anlehnung an die Figuren aus Disney's Micky Mouse erstellt. Für die Verwendung zu Testzwecken der AHV-Nummer[2] hat das UPI-Register der Zentralen Ausgleichsstelle (ZAS) einen kleinen Nummernbereich reserviert, welche verwendet wurden. Der Testdatensatz ist auf Github[3] verfügbar und kann weiter ausgebaut werden. Sollte ein grösserer Testdatensatz erstellt werden, kann der in der ISO-Norm 3166[4] vorgesehene Test-Ländercode anstelle der ersten 3 Ziffern der AHV-Nummer zur eindeutigen Identifikation verwendet werden.

Nachdem klar war, dass der Zugriffsschutz in keiner der Software-Lösungen gemacht wird, musste die Konzeption als eigenständige Lösung angegangen werden. Die Zugriffe über «Named Graph» zu schützen, wurde aufgrund der grossen Datenmenge der zu erwartenden produktiven Daten und vielschichtigen Berechtigungsanforderungen verworfen. Die Komplexität wäre zu gross geworden. Es wurden zwei Möglichkeiten geprüft:

- Eine Konfigurations-Übersetzung (Config Engine), welche eine immer gleich beschriebene und standardisierte Zugriffsregeln (Access Control Lists ACL) in die entsprechende Konfiguration des verwendeten Triple-Stores verwandelt. Für jede Triple-Datenbank müsste eine entsprechende Config Engine erstellt werden.
- Ein Proxy, der die Abfragen (SPARQL) entgegennimmt, anhand von Zugriffsregeln (ACL) filtert und die Daten gefiltert vom Triple-Store (der Datenbank) abholt. Diese Daten liefert der Proxy danach dem Datennutzer zurück.

Die Entscheidung für den Proxy wurde gefällt, da diese Lösung für alle möglichen SPARQL-Interfaces funktioniert (und keinen zusätzlichen Programmieraufwand für die Triple-Stores benötigt). Mit dem Proxy können auch «virtuelle Graphen» und weitere (externe) Datenquellen abgefragt werden. In der Sicherheitsarchitektur bietet der Proxy den Vorteil, in einer anderen Netzzone laufen zu können. Die Proxy-Variante würde für alle Use Cases funktionieren. Der Proxy hat den Vorteil, dass dieser mit weiteren Funktionalitäten (z.B. fachliches Logging, Accounting, IAM-Integration etc…) erweitert werden kann. Ein Nachteil ist, dass föderierte Abfragen (also Abfragen über mehrere Triple-Stores hinweg) nicht einfach zu implementieren sind.

In beiden Fällen müssen die Zugriffsregeln (Access Control Lists ACL) unter den Teilnehmern harmonisiert werden. Ansonsten kennen die Datenanbieter die Rollen der Datennutzenden nicht und können entsprechend deren Datenschutzanforderungen nicht erfüllen.

Eine weitere Herausforderung ist die konsistente Beantwortung von Anfragen. Wenn ein Datennutzer berechtigt ist, auf Informationen zuzugreifen, sollte dieser von den verschiedenen Stellen die gleichen Informationen zurückerhalten. Dies funktioniert nur, wenn die Zugriffsrechte und die geführten Informationen harmonisiert sind.

[2] AHV-Nummer: https://www.zas.admin.ch/zas/de/home/partenaires-et-institutions-/navs13.html

[3] ld-prototype-data/UC-Serafe at main · swiss/ld-prototype-data: https://github.com/swiss/ld-prototype-data/tree/main/UC-Serafe

[4] ISO - ISO 3166 — Country Codes https://www.iso.org/iso-3166-country-codes.html

Diese Erkenntnisse zählen bereits zu den fachlichen Implikationen. Nebst der gemeinsamen Sicherstellung der Zugriffsrechte müssen auch Prozesse angepasst werden. Damit Linked Data funktioniert, benötigt es das gezielte Beschreiben von Beziehungen bei den beteiligten Registern. Damit eine Person gefunden werden kann, weiss in unserem Beispiel das UPI-Register, in welchem Einwohnerregister eine natürliche Person das Meldeverhältnis hat. Erfolgt jetzt ein Umzug in eine andere Gemeinde, muss im UPI-Register die neue Gemeinde ebenfalls hinterlegt werden. Erst mit diesem Schritt kann eine Person gefunden werden. Beim bisherigen Ansatz geschieht das nicht, was dazu führt, dass die Person nicht einfach gefunden werden kann.

Weiter wurde geprüft, wie Datentransformationen gemacht werden können. Transformationen ab den bei eCH üblichen XML-Formaten zeigten sich generell als eher herausfordernd. Keines der Open Source verfügbaren Tools konnte auf Anhieb eingesetzt werden. Mit kleinen Korrekturen am Datenformat konnte eines der Tools eingesetzt werden. Bei anderen Tools wurde Kontakt mit den Entwicklern aufgenommen und teils gar Korrekturen ausgelöst. Die Korrekturen wurden im Rahmen des Prototyps nicht mehr getestet. Transformationen ab CSV-Formaten funktionierten auf Anhieb. Es empfiehlt sich daher die Datenstrukturen in möglichst unkomplizierten Formaten bereit zu stellen.

Zur Demonstration des Prototyps wurde ein kleines Video[5] mit Untertiteln erstellt. Die Resultate der Fachhochschulen wurden in separaten Berichten[6] dokumentiert. Alle entwickelten Artefakte (Proxy, Testdatensatz, Code zur Instanziierung der Instanzen, Proxy etc…) wurden auf Github unter einer Open Source Lizenz publiziert und dokumentiert. Alle Sounding Board Inhalte wurden auf der Webseite der Bundeskanzlei geteilt.

5 Erkenntnisse

Der Einsatz von Linked Data ist eine valable Möglichkeit, um verteilte Daten zugänglich zu machen.

Es ist dabei nicht so wichtig, in welcher Datenbank-Technologie die Daten gespeichert werden. Wichtig ist, dass diese im RDF-Format beschrieben und miteinander über eindeutige Identifikatoren in Beziehung gebracht werden. Die Daten müssen über eine SPARQL- oder REST-Schnittstelle abgefragt werden können und dementsprechend bei den Abfragen transformiert werden.

Der Integrations-Aufwand in Fachanwendungen kann durch die Standardisierung auf REST- und SPARQL-Schnittstellen gegenüber den bisherigen SOAP- oder XML-Schnittstellen reduziert werden. In jedem Fall muss die konsumierende Anwendung die Datenformate verarbeiten und lokal speichern können. Entsprechende Transformation in das eigene Datenmodel muss bei allen Integrations-Varianten getan werden und dürfte bei Ressourcen / Objekt-Orientierten Formaten einfacher sein als bei Befehlsorientierten Formaten.

Der Linked Data Ansatz erhöht die Flexibilität zur Beantwortung neuer Fragestellungen. Sofern die Beziehungen zwischen den Objekten möglichst beidseits geführt werden, können einfache Perspektiven-Wechsel auf die Daten erfolgen. So kann aus dem Datenkonstrukt unseres Beispiels die Frage der Gebührenpflichtigen in einer Wohnung ermittelt werden. Gleichzeitig könnten aber auch alle Einwohnerinnen und Einwohner eines bestimmten Gebietes ermittelt werden.

---

[5] https://www.bk.admin.ch/dam/bk/de/multimedia/dti/Demo%20Prototyp%20Linked%20Data%20nat%C3%BCrliche%20Person.mp4.download.mp4/demo_prototyp_linkeddata.mp4

[6] Zugriffskontrolle Linked Master Data: https://www.bfh.ch/.documents/ris/2024-385.206.909/BFHID-648595275-1/Zugriffskontrolle-Linked-Master-Data.pdf

Bericht HES-SO: https://github.com/swiss/ld-prototype-data/blob/main/Rapport_HES_SO_VS_Prototyp_Linked_Data_v1.0.pdf

Verknüpfte Daten ermöglichen den Austausch und die Integration von Daten zwischen verschiedenen Verwaltungen aller föderaler Ebenen unter Verwendung gemeinsamer Standards und Formate. Dies erleichtert die Zusammenarbeit zwischen den Ämtern. Die Verwendung eines auf dem RDF-Modell basierenden Datenformats erleichtert die gemeinsame Nutzung von Daten, indem durch die dokumentierten Beziehungen einfacher neue Fragestellungen beantwortet werden können.

Somit können die Daten leichter eingesehen und wiederverwendet werden. Die Einführung der Linked Data Konzepte kann dazu beitragen, die Qualität der öffentlichen Dienstleistungen zu verbessern.

Daten können automatisch abgefragt werden, unabhängig davon, in welchem Bundesamt, Kanton oder in welcher Gemeinde diese gespeichert sind. Kopien der Daten müssen nicht mehr an mehreren Stellen geführt werden. Die Datennutzer haben zu jeder Zeit direkten Zugriff auf die aktuellen Daten, weil sie diese direkt von der Quelle beziehen.

Die Nutzung verknüpfter offener Daten kann die Innovation fördern, indem Entwickler und Unternehmen neue Dienstleistungen, Anwendungen oder Analysen erstellen können.

Die Möglichkeit, gemeinsame Listen (zum Beispiel zu definierten Werten) dezentral zu erweitern, um gemeinsame Vokabulare für mehrere Zwecke zu erstellen, ist für Projekte von entscheidender Bedeutung und wird im Zusammenhang mit verknüpften Daten möglich.

6 Offene Punkte

Das erarbeitete Konzept zum Prototyp Linked Data[7] zusammen mit dem prototypisch implementierten Proxy zeigt einen möglichen Lösungsweg zur Zugriffskontrolle bei verteilt geführten Daten auf. Im Laufe der Analyse und Implementation sind weitergehende Fragen aufgetaucht, welche im Rahmen des Projekts nicht beantwortet werden konnten und weiterer Abklärungen bedürfen:

- Die (während dem Prototyp abgegrenzte) Anbindung an IAM-Systeme muss ausgestaltet werden. Dabei ist insbesondere zu berücksichtigen, wie Mutationen von Benutzern und Rechten in einer verteilten Bereitstellung gehandhabt werden.
- Federated Queries[8] (Abfragen über verteilte Datenendpunkte) sind aktuell mit dem Proxy-Ansatz nicht möglich. Sie bedürfen je nach Variante einer angepassten Implementation des Proxy oder aber weiterer Standardisierung. Die Abfragen wurden jeweils sequentiell (mit Derefernzieren) verarbeitet[9].
- Die Fehlerbehandlung, welche durch die Einschränkung des Zugriffs auf Linked Data entstehen, sind nicht harmonisiert. Dies führt bei Datennutzern zu unklaren Situationen, da die Antworten der Datenlieferanten interpretiert werden.
- Wie werden Antworten für den Client aufbereitet, wo ein Backend-Triple-Store nicht verfügbar ist oder andere unerwartete Fehler zurückgibt? Wie soll der Client damit umgehen.
- Auswirkungen auf die Performance, sowie Skalierbarkeit wurden explizit nicht untersucht.

Diese offenen Punkte müssen bei einer weitergehenden Umsetzung der Zugriffskontrolle in Linked Data geklärt werden.

[7] Zugriffskontrolle Linked Master Data: https://www.bfh.ch/.documents/ris/2024-385.206.909/BFHID-648595275-1/Zugriffskontrolle-Linked-Master-Data.pdf

[8] SPARQL 1.2 Federated Query: https://www.w3.org/TR/sparql12-federated-query/

[9] Weitere Erläuterungen sind in den Folien des Sounding Board vom Juni 24 enthalten: https://www.bk.admin.ch/dam/bk/de/dokumente/dti/themen/architektur/Stammdatenverwaltung/4.%20%20Sounding%20Board%20Juni%202024.pdf.download.pdf/4.%20%20Sounding%20Board%20Juni%202024.pdf

7 Reflektion

7.1 Fachliche Beurteilung

Die fachliche Vertreterin Carmela Schürmann aus dem Kernteam hat das Vorhaben wie folgt beurteilt:

- Das Vorgehen war sehr effektiv und Zielführend. Die Zusammenarbeit war sehr angenehm, unkompliziert und konstruktiv.
- Die Erkenntnisse während der Erarbeitung waren interessant, auch wenn diese manchmal weit über das fachliche hinaus gingen und für Fachvertreter eher schwierig zu verstehen waren Es lohnt sich, die Erkenntnisse zu teilen und in die weiteren Arbeiten zur Bereitstellung der natürlichen Person einfliessen zu lassen.
- Der VSED sieht eher eine Zentralisierung der Datensammlung zur natürlichen Person. Dies trägt zur Entlastung der Einwohnerdienste, für die Datennutzenden Ämter und die betroffenen Menschen bei. Unklar ist der Einfluss der Lösung auf die Datenqualität und die Prozesse der datenführenden Stellen.

Das Zusammenspiel mit den Quellregistern müsste für dieses Konzept nun überprüft werden. Dafür müssen folgende weiteren Schritte untersucht werden:

- Auswirkungen auf die Prozesse der datenführenden Stellen bei den Einwohnerdiensten
- Auswirkungen und notwendige Anpassungen unterschiedlicher Erfassungsgrundlagen und Datenfeldern.

Dem VSED ist es wichtig, dass diese für die Schweiz grundlegenden Daten adäquat sind. Diese Daten müssen wir im Griff haben.

7.2 Inhaltliche Beurteilung

Für den Prototypen wurden die fünf Ziele gemäss Kapitel 2 definiert. Die Zielerreichung der Ziele wird wie folgt beurteilt.

Ziel 1: Je nach Datensatz können die Daten gemischt geschützt und offen publiziert werden. Allerdings bietet das Vorschalten eines Proxy in beiden Fällen Vorteile. Mit der Implementierung des Zugriffsschutzes auf dem Proxy können die offenen und geschützten Daten gleichbehandelt werden. Offene Daten werden ebenfalls durch den Proxy angeboten und können von den zusätzlichen Funktionen des Proxy, wie der Protokollierung des Zugriffs profitieren. Die Nutzung der Daten funktioniert aber auch, wenn die offenen Daten von anderer Stelle bezogen werden. Für das nutzende System erscheint ein Proxy als normaler Endpunkt.

Fachlich fragt sich, ob die Trennung offener und geschlossener Daten Sinn ergibt. Sehr oft werden wohl gewisse Inhalte eines Datensatzes nur für den verwaltungsinternen Gebrauch freigegeben sein. Hier kann der Proxy für unautorisierte Nutzer die offenen Daten und gleichzeitig bei erfüllter Autorisierung die geschützten Daten bereitstellen.

Ziel 2: Mit der Proxylösung und dem erstellten Datenmodell konnte aufgezeigt werden, dass mit Linked Data einfach die Daten auf der einen Seite einfach auffindbar sind und auf der anderen Seite geschützt werden können.

Ziel 3: Die überprüften Semantiken von eCH passen gut zu den EU Core Vocabularies. Die im Umfeld der natürlichen Person verwendeten Standards lassen sich in Interoperablen Datenformaten beschreiben.

Ziel 4: Das Datenmodell[10] konnte erarbeitet und im Prototyp implementiert werden

Ziel 5: Es konnte keine geeignete Open-Source Lösung gefunden werden, die am Markt verfügbar ist und welche einen anerkannten Standard umsetzt. Die Wahl der Datenbank-Lösung beeinflusst die Verbreitung von Linked Data nicht. Selbst herkömmliche Datenbanken können mit geeigneten Transformationen in das Netzwerk der Linked Data eingebunden werden. Von der Empfehlung von zu verwendender Open-Source Software wurde daher abgesehen. Die entwickelte Lösung hingegen wurde unter einer Open Source Lizenz publiziert.

8 Empfehlungen

Das Thema von Linked Data wird international in Datenräumen breit umgesetzt und sollte vertieft weiter betrachtet werden. Um die Daten schützen zu können, benötigt es abgestimmte Identity und Accessmanagement Umsetzungen. Hierbei wird empfohlen, folgende Punkte genauer zu betrachten.

- Gemeinsames Rollenkonzept definieren oder bestehendes übernehmen (z.B. ISO-5009)
- Fehlerhandling und Fehlermeldungen standardisieren: Unter anderem muss das Fehlerhandling bei unberechtigten Zugriffen einheitlich geregelt werden. Dies kann z.B. mittels fehlender/maskierter Daten in der Antwort, Fehlermeldungen oder anderer Mechanismen erfolgen.

8.1 Datenökosystem

- Darauf achten, dass in den Datenräumen die Beschreibungen in RDF und auf Basis bereits getroffener Definitionen erfolgen. Es hat sich gezeigt, dass insbesondere im europäischen Umfeld stark auf den Ansatz gesetzt wird. Dadurch können unsere Datenräume by Design Interoperabel gestaltet werden.
- Verwendung von gemeinsamen und fachlichen Identifikatoren
- Standard festlegen, um Zugriffsregeln zu beschreiben
- In Datenräumen und über diese hinaus für die Harmonisierung von Informationen sorgen und die Informationen Maschinen und Menschenlesbar zugänglich machen.
- Prototyp des Proxy auf internationaler Ebene teilen und an der Konzeption und Weiterentwicklung des Zugangsschutzes partizipieren.

8.2 Bereitstellung natürliche Person

- Ausarbeiten der benötigten Ideal-Architektur
- Vergleichen der Auswirkungen der unterschiedlichen Architekturen auf die beteiligten Stakeholder (Datenbewirtschafter, Datennutzer und Kunden)

[10] Bericht HES-SO: https://github.com/swiss/ld-prototype-data/blob/main/Rapport_HES_SO_VS_Prototyp_Linked_Data_v1.0.pdf