SocietyByte
BFH-Magazin für die Humane Digitale Transformation

Sicher vernetzte Personendaten: Neue Ansätze mit Linked Data

Von Pascal Mainini (BFH Technik und Informatik) | 0 Kommentare

Die Berner Fachhochschule entwickelt gemeinsam mit der Bundeskanzlei neue Wege, um sensible Personendaten sicher zu vernetzen. Das Projekt nutzt die Linked Data-Technologie und schafft zusätzliche Schutzmechanismen für dezentral gespeicherte Datenbestände. Damit könnten künftig beispielsweise Einwohnerregister verschiedener Behörden effizienter zusammenarbeiten.

Einführung

In der Schweiz werden Daten zu natürlichen Personen dezentral geführt und unterliegen verschiedenen Zuständigkeiten (Stammdaten). Ein Teil der Personendaten findet sich beispielsweise in den Einwohnerregistern der Kantone und Gemeinden, ein anderer (z.B. die AHV-Nummer) werden in Systemen auf Bundesebene verwaltet. Die Nutzung von Informationen aus verschiedenen Registern ist daher oft umständlich und aufwändig. Linked Data-Technologie könnte hier eine Lösung bieten; diese lässt sich jedoch nicht bedenkenlos für schützenswerte Daten, wie z.B. Personendaten, einsetzen, sondern bedarf weiteren Schutzmechanismen.

Was ist Linked (Open) Data?

Linked Data ist eine Technologie, die Informationen miteinander verknüpft und in Beziehung setzt – ähnlich einem Netzwerk. Es basiert auf Standards wie RDF (Resource Description Framework) und SPARQL (eine Abfragesprache für RDF), welche vom W3C publiziert wurden.[1] [#_ftn1] [2] [#_ftn2] Ziel ist es, Datensätze dezentral zu speichern und interoperabel zu machen.

Am bekanntesten ist Linked Open Data (LOD), wo Daten öffentlich zugänglich gemacht werden, um sie frei nutzen zu können. Ein Beispiel für LOD in der Schweiz ist der Dienst LINDAS (Linked Data Service) des Schweizerischen Bundesarchivs, über welchen z.B. Daten zur Waldbrandgefahr oder zu den Strompreisen abgerufen werden können.[3] [#_ftn3] [4] [#_ftn4] [5] [#_ftn5] Dieses Modell eignet sich jedoch nicht für schützenswerte Informationen wie Personendaten, da es keine inhärenten Mechanismen zur Zugriffskontrolle und somit zum Schutz der Daten bietet.

Vorteile von Linked Data für Stammdaten

Der Einsatz von Linked Data hat viel Potential, z.B. in der Verwaltung von Personendaten. Der vereinfachte Datenaustausch zwischen verschiedenen Behörden könnte aufwändige und individuelle Schnittstellen ersetzen oder für mehr Ausfallsicherheit sorgen.

Ein Beispiel ist die Erhebung von Radio- und Fernsehgebühren: hierzu müssen Daten aus verschiedenen Registern zusammengeführt werden. Da die Abgaben pro Haushalt anfallen, ist sowohl eine Anfrage beim sog. UPI-Register (Unique Person Identifier, konkret die AHV-Nummer) als auch bei den einzelnen Einwohnerregistern der Gemeinden notwendig, da nur diese das Meldeverhältnis einer Person kennen.

Dezentrale Personenstammdaten mit Linked Data

Das gleichnamige Projekt [https://www.bfh.ch/de/forschung/forschungsprojekte/2024-385-206-909/] der Berner Fachhochschule und der Bundeskanzlei hat die notwendigen Grundlagen erarbeitet, um Use Cases wie diesen mittels Linked Data zu ermöglichen.[6] [#_ftn6] Es wird das «O» gestrichen – die Daten sind nicht mehr offen, sondern die Linked Data-Technologie wird für schützenswerte Daten eingesetzt.

Nach einer initialen Anforderungsanalyse und der Erarbeitung von verschiedenen Use Cases wurde der Markt der Triplestores (Datenbanksystem für Linked Data) untersucht. Der Fokus lag dabei auf den Möglichkeiten zur Zugriffskontrolle und Rechtevergabe, welche für den Schutz der Daten notwendig sind. Es wurden sowohl kommerzielle Lösungen wie auch Open Source Software verglichen. Das vorgefundene Spektrum reichte von nicht vorgesehen bis zu sehr umfassenden Möglichkeiten. Es wurden jedoch keine einheitlichen Standards identifiziert, was ein spezifisches Vorgehen je nach eingesetztem Triplestore erfordert. Dies ist in Bezug auf eine gesamtschweizerische Architektur suboptimal.

Architekturvarianten

Aufbauend auf den so weit gewonnenen Erkenntnissen wurden verschiedene Architekturvarianten ausgearbeitet:

1. **Standardisierung eines Triplestores**: Ein einheitlicher Triplestore für alle Beteiligten würde einen Ansatz zur Zugriffskontrolle standardisieren. Dieser Ansatz ist jedoch zu wenig flexibel.
2. **Config-Engine**: In dieser Variante würde eine eigene Sprache für die Zugriffskontrolle definiert. Die Zugriffsregeln würden dann gemäss dieser

beschrieben und mittels einer Software für spezifische Triplestores übersetzt. Diese Lösung wäre denkbar, erfordert jedoch einen hohen Grad an Standardisierung.

3. **SPARQL-Proxy**: *Diese Variante wurde als Prototyp im Projekt umgesetzt. Siehe nächster Abschnitt.*
4. **Datenmodellierung / Separation:** Mit entsprechender Modellierung der Daten kann der Zugriff ebenfalls eingeschränkt werden. Dies erfordert jedoch eine Anpassung der Daten und ist fehleranfällig.
5. **Verschlüsselung der Daten:** Als letzte untersuchte Möglichkeit könnten Daten auch mittels Verschlüsselung geschützt werden. Es ergeben sich hier ähnliche Probleme wie bei der Datenmodellierung, sowie zusätzliche z.B. durch die nötige Verwaltung der Schlüssel.

Prototyp: SPARQL-Proxy

Der SPARQL-Proxy ist ein Konzept, das einen Vermittler auf Ebene der SPARQL-Abfrage vorsieht. Dieser fungiert als «intelligenter Filter»: Zwischengeschaltet zwischen Triplestore und anfragender Stelle wird eine eingehende SPARQL-Abfrage durch den Proxy überprüft und gemäss der Berechtigungen der Stelle eingeschränkt. Mit dieser modifizierten Anfrage findet dann die Abfrage beim eigentlichen Triplestore statt. Die ursprüngliche Anfrage wird dann nochmals auf die erhaltenen Antwortdaten angewendet und das Resultat an die anfragende Stelle zurückgegeben. Dieser Schritt ist notwendig, um die grundlegende Struktur der abgefragten Daten zu erhalten.

Ein zentraler Vorteil des Proxys gegenüber den anderen Architekturvarianten liegt in der Flexibilität: komplexe Berechtigungsprüfungen und feingranulare Zugriffskontrollen können umgesetzt werden, ohne Änderungen an den dahinter liegenden Triplestores vornehmen zu müssen. Damit ist der SPARQL-Proxy eine effektive Lösung, um Linked Data zu schützen.

Fazit

Das Projekt hat aufgezeigt, dass sich der Linked Data-Ansatz auch für dezentrale Personenstammdaten in der Schweiz eignet. Damit ist ein wichtiger Grundstein für die weitere Erforschung der Möglichkeiten, die die Technologie bietet, gelegt – besonders im Hinblick der digitalen Transformation in Verwaltungen. Nebst einiger spezifischer Fragen bezüglich der Proxy-Implementierung gibt es zwei wesentliche offene Punkte, die als nächstes zu klären sind:

1. **Die Integration mit bestehenden Identitätsmanagement-Systemen (IAM):** Zum Beispiel die Anbindung an IAM Bund.
2. **Performance und Skalierbarkeit**: Wie performant bleibt das System auch bei bei großen Datenmengen und vielen Abfragen?

Die BFH könnte in einem Folgeprojekt die nächsten Vorrausetzungen für eine besser vernetzte Verwaltung in der Schweiz schaffen.

Quellen

[1] [#_ftnref1] W3C RDF Standards incl. SPARQL [https://www.w3.org/standards/techs/rdf]

[2] [#_ftnref2] https://www.w3.org/ [https://www.w3.org/]

[3] [#_ftnref3] https://lindas.admin.ch/?lang=de [https://lindas.admin.ch/?lang=de]

[4] [#_ftnref4] https://environment.ld.admin.ch/foen/gefahren-waldbrand-warnung/1 [https://environment.ld.admin.ch/foen/gefahren-waldbrand-warnung/1]

[5] [#_ftnref5] https://energy.ld.admin.ch/elcom/electricityprice-swiss [https://energy.ld.admin.ch/elcom/electricityprice-swiss]

[6] [#_ftnref6] https://www.bfh.ch/de/forschung/forschungsprojekte/2024-385-206-909/ [https://www.bfh.ch/de/forschung/forschungsprojekte/2024-385-206-909/]

AUTHOR: PASCAL MAININI

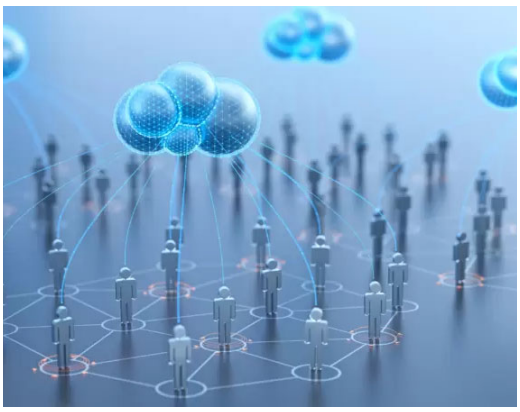
Pascal Mainini ist Tenure Track Dozent am Institute for Cybersecurity and Engineering ICE der Berner Fachhochschule. Als Experte für angewandte Kryptographie, sichere Soft- und Hardware sowie Datenschutz und Privatsphäre setzt er sich dafür ein, die Integrität, Vertraulichkeit und Sicherheit moderner digitaler Systeme zu gewährleisten.

Posts from Pascal Mainini | Website

Create PDF

Ähnliche Beiträge

Datenethik braucht ein Gleichgewicht zwischen Innovation und Verantwortung



Lassen sich Datenschutz und Informationsgehalt vereinbaren?

0

COMMENTS