



Bundesarchiv, 11. September 2020

Dienstleistung «Linked Data Service LINDAS»

Übersicht Angebot und Kosten

Inhaltsverzeichnis

1	Das BAR und Linked Data	4
2	Rechtliche Grundlage und Rahmenbedingungen	4
3	Was ist Linked Data (LD)?	4
4	Datenmodell	5
5	Linked Data Service LINDAS	6
5.1	Datenkonvertierung	6
5.1.1	Konfiguration der Konvertierungen	7
5.1.2	Zuordnungsformat	7
5.1.2.1	Data Cube Curator	8
5.1.2.2	Schema Manager	8
5.2	Speicher: Triplestore Infrastruktur	9
5.3	Zugang: Triplestore-Schnittstellen	10
5.3.1	Daten-Dereferenzierung	10
5.3.2	Offene Schnittstelle SPARQL	11
5.3.3	LINDAS.admin.ch	11
6	Das Ökosystem von LINDAS	12
6.1	Visualize	12
7	Dienstleistung des Bundesarchivs	12
7.1	Eckwerte des Angebots	12
7.2	Triplestore Infrastruktur – Service Levels	13
7.2.1	Produktionsumgebung	13
7.2.2	Test- und Integrationsumgebung	13
8	Mitwirkungspflichten des Datenlieferanten	13
9	Kosten	13
10	Vorgehen	14
11	Kontakt	14
12	Anhänge	14

Glossar

Abkürzung	Definition
AIS	Archivinformationssystem
BAFU	Bundesamt für Umwelt
BAR	Schweizerisches Bundesarchiv
BBI	Bundesblatt
BGA	Bundesgesetz über die Archivierung
CSV	comma-separated values
FHG	Finanzhaushaltgesetz
FTP	File Transfer Protocol
IAFP	Voranschlag mit integriertem Aufgaben- und Finanzplan
LD	Linked Data
LINDAS	Linked Data Service
RDF	Resource Description Framework
SBB	Schweizerische Bundesbahnen
SPARQL	Protocol And RDF Query Language
URI	Uniform Resource Identifier
W3C	World Wide Web Consortium
XML	eXtensible Markup Language
URL	Uniform Resource Locator
W3C	World Wide Web Consortium

1 Das BAR und Linked Data

Linked Data ist eine Technologie, die die Maschinenlesbarkeit der Daten erhöht und damit die Wiederverwendung derselben stark vereinfacht. Linked Data baut dabei auf bestehenden Web Technologien auf. Linked Data ermöglicht, Daten über technische und organisatorische Grenzen hinweg zu verknüpfen und Probleme anzugehen, die mit bestehenden Technologien nur schwer oder mit deutlich mehr Aufwand umsetzbar sind. Das graphbasierte Datenmodell RDF benötigt initial eine gewisse Einarbeitungszeit, sowohl auf Seiten des Datenhalters wie des Datennutzers. Danach kann aber deutlich mehr Wissen wiederverwendet werden, als dies typischerweise in geschlossenen Systemen der Fall ist.

Die vom BAR aufgebaute Linked Data Infrastruktur LINDAS (Linked Data Service) steht im Zeichen der Modernisierung des Bundesarchivs. Die Linked Data Technologie eignet sich nicht nur besonders gut für die Langzeitaufbewahrung von Datenbanken, sondern könnte mittelfristig auch für das Archivinformationssystem AIS des Bundesarchivs eingesetzt werden. Zudem eignet sich Linked Data auch als Publikationskanal von Daten, die auf der Basis von analogen Inhalten aufgebaut wurden (z.B. Stapfer Enquête, Staatsrechnungen, Amtliche Sammlung des Bundesrechts). Die Linked Data Technologie ist somit ein wichtiger Baustein, damit das BAR seinen Grundauftrag auch künftig erfüllen kann.

2 Rechtliche Grundlage und Rahmenbedingungen

Die Tätigkeit des Schweizerischen Bundesarchivs (BAR) stützt sich auf das Bundesgesetz über die Archivierung vom 26. Juni 1998 (BGA, SR 152.1). Das BAR sorgt für die Archivierung rechtlich, politisch, wirtschaftlich, historisch, sozial oder kulturell wertvoller Unterlagen des Bundes (Art. 2 BGA). Es sorgt überdies auch für die sichere und sachgemässe Aufbewahrung, Erschliessung und Vermittlung des Archivguts und beteiligt sich an dessen Auswertung (Art. 17 Abs. 3 BGA und BBI 1997 964). LINDAS leistet dazu einen Beitrag.

LINDAS kann technisch problemlos für unterschiedlichste Daten genutzt werden. Deshalb stellt das BAR seinen Service auch Dritten zur Verfügung. Gewerbliche Leistungen von Verwaltungseinheiten an Dritte (d.h. an verwaltungsexterne Akteure) bedürfen gemäss Art. 41 des Bundesgesetzes über den eidgenössischen Finanzhaushalt vom 7. Oktober 2005 (FHG, SR 611.0) einer formell-gesetzlichen Grundlage. Für das Bundesarchiv findet sich in Art. 18 Abs. 1 BGA die Ermächtigung in seinem Aufgabenbereich einzelne besondere Dienstleistungen auf privatrechtlicher Basis an Dritte zu erbringen. Die gesetzliche Bestimmung knüpft diese Befugnis an das Vorhandensein einer entsprechenden Position im Voranschlag mit I-AFP.

Öffentliche Gemeinwesen und bundesnahe Betriebe können den Linked Data Service LINDAS nutzen. Darunter fallen bundesinterne Verwaltungseinheiten, Kantone, Gemeinden und z.B. auch die SBB.

3 Was ist Linked Data (LD)?

Oft liegen unterschiedlichen Daten unterschiedliche Datenmodelle zugrunde: die Datenstruktur ist nicht gleich, Definitionen weichen voneinander ab und auch die Formate unterscheiden sich. Auch müssen sie zur Weiterverarbeitung manuell mit spezifischen Anwendungen geöffnet werden. Oft sind diese Inhalte für Computer nicht in einer verständlichen Logik verfügbar und nur für Menschen interpretierbar.

Einen direkteren Zugriff auf Daten bieten Schnittstellen. Damit können Daten, die in Datenbanken enthalten sind, direkt abgefragt werden. Für jede Datenbank braucht es aber eine eigene Schnittstelle und jede dieser Schnittstellen ist wie eine Tür, die ein eigenes Schloss besitzt. Jede dieser Türen ist unabhängig voneinander gebaut, so dass jedes Schloss einzigartig ist. Mit anderen Worten, jede Schnittstelle ist anders, sie muss untersucht und ihre Verwendung programmiert werden. Insbesondere bei einer Kombination mehrerer Datenbanken bereitet das Schwierigkeiten.

Linked Data oder auch das «Web der Daten» genannt, ist eine technologische Lösung, die die oben beschriebenen Probleme löst. Sie ermöglicht, Inhalte von Datenbanken aus Fachanwendungen so zu veröffentlichen, dass diese einfach wieder- und weiterverwendet sowie miteinander verknüpft werden können. Eigene Daten können mit fremden verlinkt werden, weil jedes Datenelement eine eindeutige Kennung (seinen Uniform Resource Identifier URI) hat. Dadurch, dass diese strukturierten Daten verfügbar und wiederverwendbar gemacht werden, können die Silostrukturen der Fachanwendungen teilweise aufgelöst werden. Mit LINDAS und seiner SPARQL-Abfrageschnittstelle ist es für eine Fachanwendung möglich, offene Daten aus anderen Fachanwendungen zu verwenden, ohne diese Daten lokal kopieren zu müssen. Aktualisiert eine Organisation einen Datensatz wird die aktualisierte Version den Drittsystemen direkt verfügbar und muss nicht mehr von Informatikspezialisten in die Applikation eingespielen werden.

Zusammenfassend, wenn die Daten gut aufbereitet und mit der richtigen Interpretation (Semantik) zur Verfügung gestellt werden, erleichtert das Drittnutzern die Verwendung dieser Daten erheblich. Statt diese Arbeit an die Nutzer der Daten zu delegieren und damit zu vervielfachen, wird sie einmalig und korrekt beim Datenhalter selber gemacht.

4 Datenmodell

Die bei LINDAS angewandte Modellierung von Daten basiert auf einem W3C-Standard, der als RDF (Resource Description Framework) bezeichnet wird. Dieses Modell legt fest, dass ein Informationselement in Form von Triples beschrieben wird (siehe Abbildung 1): <Subjekt> <Prädikat> <Objekt> oder <Subjekt> <Eigenschaft> <Wert>, wobei die <Eigenschaft> ein besonderer Fall des <Prädikats> ist. Das <Subjekt> und das <Objekt> sind Einheiten, die durch URIs (Uniform Resource Identifiers) identifiziert werden. Jede Einheit ist eine computerbasierte Darstellung eines abstrakten Konzepts oder eines realen Objekts, und alle Triples, die das gleiche <Subjekt> haben, stellen eine Beschreibung dieser Einheit dar.

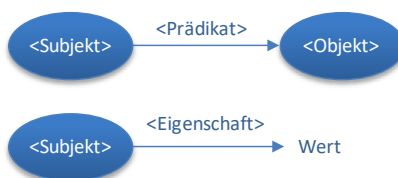


Abbildung 1: Die beiden Triples-Formen, die in einem Triplestore nebeneinander bestehen.

Eine andere Art zur Interpretation des Begriffs «Triples» besteht darin, jedes dieser Triples als Bogen zwischen den zwei Knoten eines Graphs zu betrachten, wobei die Knoten durch die <Subjekte> und <Objekte> gebildet werden. Häufig spricht man auch von Knowledge Graphs, um die mit RDF generierten Datenstrukturen zu bezeichnen (siehe Abbildung 2).

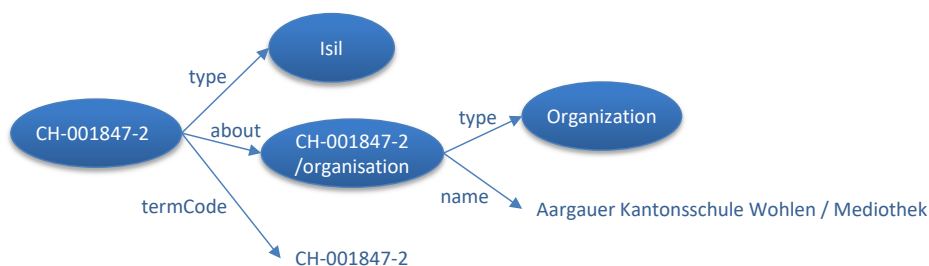


Abbildung 2: Vereinfachte Illustration eines Teils des Graphs «[Aargauer Kantonsschule Wohlen / Mediothek](#)» aus der Liste der in der Schweiz zugewiesenen ISIL-Codes.

5 Linked Data Service LINDAS

Das Bundesarchiv stellt mit LINDAS eine Basisinfrastruktur für das automatisierte Erstellen und Bereitstellen von offenen und nicht-sensiblen Daten als Linked Data zur Verfügung. Die Plattform reduziert den Aufwand für neue Datenhalter, in die Technologie einzusteigen und unterstützt interessierte Behörden mit Hilfe einer Reihe von Werkzeugen, ihre Daten ebenfalls als Linked Data zur Verfügung zu stellen. Für Datennutzende ermöglicht LINDAS einen vereinfachten Zugang zu offenen Daten.

LINDAS besteht aus den folgenden Elementen:

- **Datenkonvertierung:** Infrastruktur zur Konfiguration und Ausführung der Datenkonvertierung in das RDF-Format.
- **Speicher:** Triplestore Infrastruktur, um Daten im RDF-Format zu speichern und zugänglich zu machen.
- **Zugang:** Schnittstellen, um auf die im Triplestore gespeicherten Daten zuzugreifen.

LINDAS ist Teil eines Ökosystems. Dadurch wird die Verwendung von offenen W3C-Standards und die Entwicklung von Werkzeugen, die nicht direkt Teil des LINDAS-Angebots sind, begünstigt.

5.1 Datenkonvertierung

Der Datenkonvertierungsprozess wird als Konvertierungs-«Pipeline» bezeichnet. Diese Pipelines werden entweder automatisch (z.B. zu einer festgelegten Zeit) oder manuell in einer Infrastruktur ausgeführt, die sich nahe bei den Triplestores befindet.

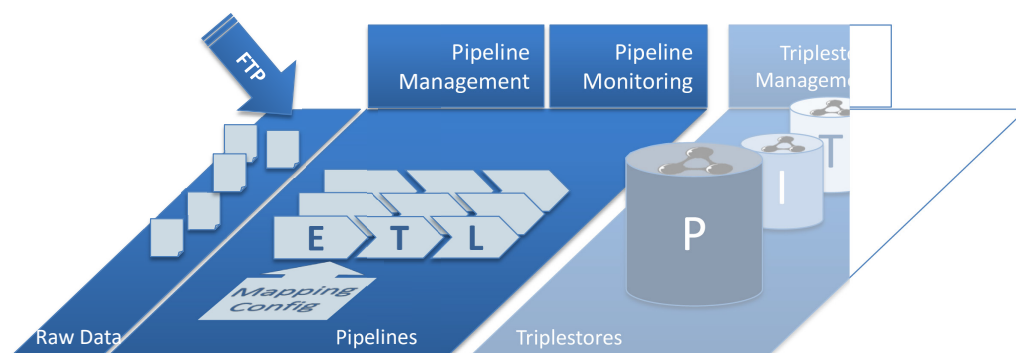


Abbildung 3: Überblick über die Elemente, die zur Einspeisung neuer Daten in die verschiedenen Triplestores dienen.

Die konstituierenden Elemente einer Pipeline sind:

- Eine Datenquelle, die sich aus einer oder mehreren Dateien (CSV, XML oder andere)

zusammensetzt, die in einem dafür bestimmten Speicherplatz abgelegt sind (Raw Data). Diese Dateien werden über FTP in diesen Speicherort importiert.

- Das Kernstück einer Pipeline besteht aus einer Reihe von *Extract-Transform-Load*-Operationen. Diese Pipeline ist entweder massgeschneidert, weil sie auf einen sehr besonderen Fall ausgerichtet ist, oder sie ist allgemein (im Folgenden als «generische Pipeline» bezeichnet) und verwendet eine Zuordnungsdatei (Mapping Config) zur Definition der charakteristischen Aspekte der Konvertierung. Wir fördern diesen zweiten Ansatz, um die Kosten für die Entwicklung der Datenkonvertierungen zu senken. Die Erstellung einer Mapping Config wird im Abschnitt 5.1.1 erläutert.

Mit einem Pipeline-Management-Tool können neue Pipelines instanziiert und bestimmte Merkmale bestehender Pipelines geändert werden. Mit dem Monitoring-Tool kann der Ausführungsstatus der Pipelines jederzeit rasch abgefragt werden. Kommt es bei der Ausführung einer Pipeline zu Problemen, so werden den Zuständigen dieser Umgebung (BAR) und den Datenlieferanten, denen die problematischen Pipelines gehören, E-Mails geschickt.

5.1.1 Konfiguration der Konvertierungen

Die Konvertierung von Daten in RDF kann deutlich vereinfacht werden, wenn eine RDF-Ontologie besteht und diese dem Datenmodell ähnlich ist, das in der Applikation verwendet wird, aus der die Daten extrahiert werden. Diese Konvertierung kann auf mehreren Standard-Ontologien¹ wie [schema](#), [skos](#), [dctterms](#), [prov](#) etc. basieren. In diesem Fall muss lediglich angegeben werden, welche Entsprechungen zwischen den ursprünglichen Daten und den in RDF konvertierten Daten erstellt werden müssen (siehe Kapitel 5.1.2.1).

Muss eine Ontologie formalisiert werden, weil nichts genau mit den Merkmalen eines Datensatzes übereinstimmt, muss diese – zusätzlich zur *Definition der Entsprechungen* – beschrieben werden (siehe Kapitel 5.1.2.2).

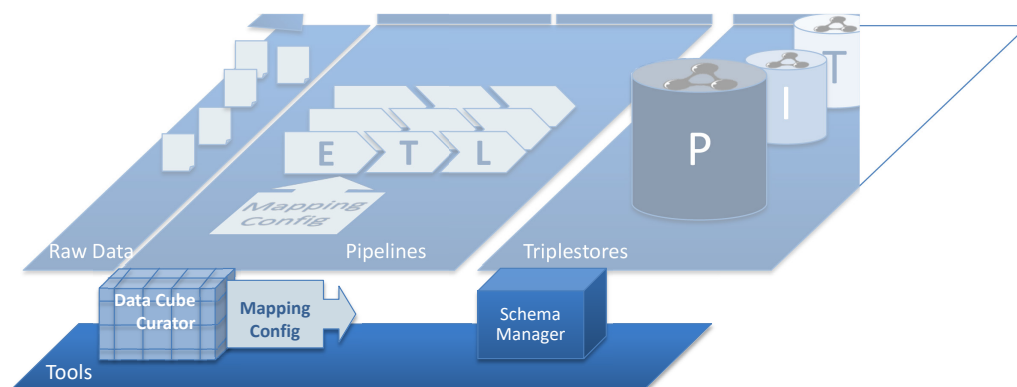


Abbildung 4: Liste der verfügbaren Web-Tools zur Erleichterung der Arbeit zur Spezifizierung der Datenkonvertierung.

5.1.2 Zuordnungsformat

In einer Zuordnungsdatei (Mapping Config) werden die Modalitäten zur Konvertierung der in Form von Tabellen aus den Datenbanken exportierten Daten in RDF definiert. Es existieren mehrere Formate für die Definition solcher Zuordnungen: [R2RML](#), [RML](#), [CARMML](#) oder [CSV on the Web](#). In der generischen Pipeline, die wir zur Verfügung stellen, wird das letzte Format verwendet.

¹ Eine sehr umfassende Liste von Datenmodellen wird auf der Website <https://lov.linkeddata.es/dataset/lov/> geführt.

Da es sich um ein Text-Format handelt, kann eine solche Datei ohne besondere Hilfe problemlos manuell editiert werden, wobei jedoch das Risiko besteht, dass viele Fehler eingefügt werden, die sich bei der Ausführung der Pipeline, die diese Konfiguration verwendet, nur schwer feststellen lassen. Zur Erleichterung der Erstellung einer Mapping Config wird ein Tool angeboten, das Bestandteil des LINDAS-Angebots (siehe nächster Abschnitt) ist.

5.1.2.1 Data Cube Curator

Einer der häufigsten Anwendungsfälle bei der Benutzung von LINDAS ist das Umwandeln von filebasierten, strukturierten Daten in das RDF-Format und dessen Integration in die Tripstore Infrastruktur. Der Data Cube Curator (siehe Abbildung 5) ermöglicht es,

- Daten aus einer mehrdimensionalen Cube-Struktur in RDF umzuwandeln;
- die Daten, die die Dimensionen beschreiben, und diejenigen, die die Messungen enthalten, zu identifizieren;
- die Daten zu prüfen, ob sie selbst und ihre Struktur der Spezifizierung entsprechen.
- eine Mapping Config zu erstellen, die leicht auf eine generische Pipeline übertragbar ist (siehe Kapitel 5.1).

Der [Data Cube Curator](#)² ist eine Web-Applikation, die für die Strukturierung der RDF-Daten auf der Ontologie [RDF Data Cubes](#) basiert. Diese Applikation erfordert keine vertieften Linked-Data-Kenntnisse. Sie wurde entwickelt, um nicht spezialisierten Personen, die ihre Daten kennen, zu ermöglichen, die Konvertierung dieser Daten in RDF mit einem Minimum an Ausbildung zu konfigurieren.

5.1.2.2 Schema Manager

Bei Linked Data ist es üblich, eine RDF-Einheit mit einer in einer Ontologie definierten Klasse zu verknüpfen. Diese Verknüpfung wird mit dem Prädikat «Typ» ausgedrückt (siehe Abbildung 2). Eine Klasse ist insbesondere durch eine Reihe von Merkmalen (Prädikate und Eigenschaften) definiert, die zur Charakterisierung einer ihr zugeordneten Einheit dienen. Dieses Prinzip ist keine Besonderheit von Linked Data, sondern im Bereich der Datenbanken allgemein üblich. Der Unterschied bei Linked Data besteht darin, dass mit RDF sowohl Daten als auch Ontologien dargestellt werden können.

Die Web-Anwendung [Schema Manager](#) dient genau dazu, Ontologien – man spricht auch von Schemen – für Linked Data in Linked Data zu erstellen. Sie weist folgende Merkmale auf:

- die Erstellung einer Ontologie kann in Zusammenarbeit und transparent erfolgen;
- das veröffentlichte Ergebnis dient direkt als Dokumentation der Ontologie für die Nutzer (siehe Abbildung 6);
- die Ontologien werden automatisch als Datensatz in Linked Data auf LINDAS publiziert.

² Der Entscheid, dieser Applikation einen Namen zuzuweisen, der in Verbindung mit der Verwaltung der Pipelines steht, ist dadurch begründet, dass der Data Cube Curator ein Teil von dem ist, was diese Website ermöglichen wird, nämlich Pipelines zu konfigurieren und zu verwalten.

ISIL

Canonical URL: <https://schema.ld.admin.ch/Isil>

Quick links: [Proposals](#) [Conversation](#)

Description

International Standard Identifier for Libraries

Long Description
 Le code 'International Standard Identifier for Libraries' (ISIL) est défini par la norme ISO 15511 et est unique au niveau international.

Cette classe définit la représentation de l'**ISIL suisse** sous contrôle de la Bibliothèque nationale suisse BN.

Cette classe étend **DefinedTerm** de schema.org par les informations supplémentaires disponibles sur l'**ISIL suisse**.

Equivalent Class

<http://www.w3.org/2002/07/owl#equivalentClass>

- <https://schema.org/DefinedTerm>

Properties

Property	Expected Type	Description
about	Organization	L'organisation à qui le numéro ISIL est attribué.
administrativeHistory	string	Histoire de l'institution
inventoryDescription	string	Description des fonds
libraryMaterials	string	Fonds bibliothécaire

Abbildung 5: Screenshot der Dokumentation in schema.id.admin.ch des ISIL-Datenmodells.

5.2 Speicher: Triplestore Infrastruktur

Linked Data wird im Format RDF gespeichert und die Daten bestehen aus sogenannten Triples. Der Triplestore **P** von LINDAS ist eine hochverfügbare Produktionsumgebung, ausgelegt um fünf Milliarden Triples³ zu speichern (siehe Abbildung 3). Es ist eine flexible Lösung, die mit der Nachfrage wachsen und bei Bedarf skaliert werden kann. LINDAS umfasst auch eine Integrations- und eine Testumgebung (die sogenannten Triplestores **I** und **T**), die zur Erstellung von Use Cases und zur Validierung von neuen technologischen Versionen dienen.

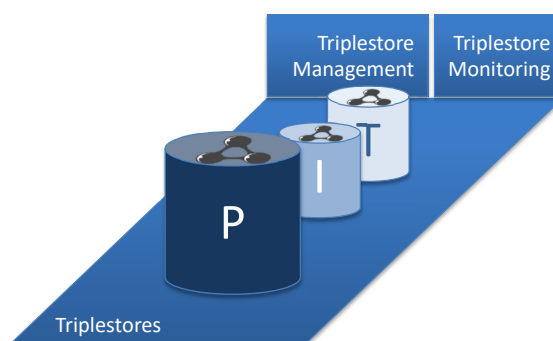


Abbildung 6: Liste der verschiedenen Triplestores von LINDAS, die mit ihren Management- und Monitoring-Tools ausserhalb der Bundesinfrastruktur gehostet werden.

Der Triplestore kann von der öffentlichen Verwaltung und bundesnahen Organisationen für Tests und produktive Anwendungen mit LD genutzt werden. Je mehr Daten organisationsübergreifend zur Verfügung gestellt werden, desto besser lassen sich das Potential und die

³ Zur Schätzung der Grösse einer relationalen Datenbank in Triples muss pro Konvertierungstabelle die Zahl der Zeilen multipliziert mit der Zahl der zu übertragenden Spalten berechnet werden. Die Summe all dieser Tabellen liefert eine gute Schätzung.

Verwendungsmöglichkeiten dieser Technologie erschliessen.

5.3 Zugang: Triplestore-Schnittstellen

Bei den verschiedenen Triplestores (P, I und T) handelt es sich um direkt über das Internet zugängliche Datenbanken. Durch die Identifizierung jeder Einheit durch einen URI, der die gleiche Syntax wie die URL einer Webseite verwendet, wird die Konsultation dieser Daten über das Internet sehr einfach. Diese Schnittstelle, die auf die in RDF veröffentlichten Daten standardisiert ist, setzt sich aus drei Elementen zusammen (siehe Abbildung 7).

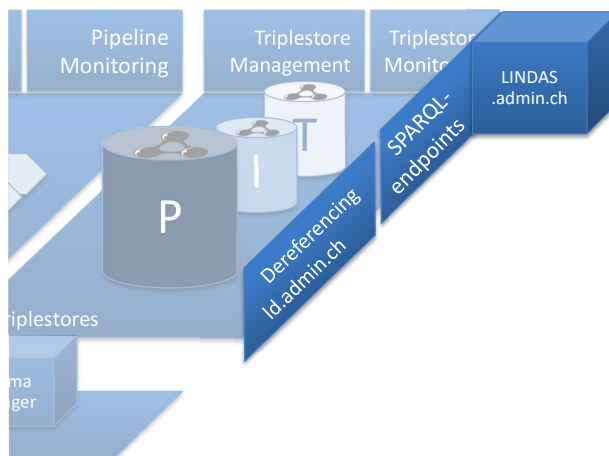


Abbildung 7: Liste der Applikationen und Schnittstellen, die für die Umwandlung von RDF-Daten in Web Daten zur Verfügung stehen.

5.3.1 Daten-Dereferenzierung

Die Infrastruktur LINDAS beinhaltet ein Modul zur Dereferenzierung der RDF-Daten mit URIs, die als Basis «Id.admin.ch» nutzen. Mit dieser Dereferenzierungsfunktion kann jede Einheit, die diese Domäne nutzt, in Form einer dynamischen Webseite mit allen Prädikaten und Werten dargestellt werden, die im Triplestore für diese Einheit enthalten sind.

The screenshot shows a web browser window with the address bar containing the URL `https://culture.ld.admin.ch/isil/CH-001847-2`. The page title is 'Aargauer Kantonsschule Wohlen / Mediothek'. Below the title, there is a URL `https://culture.ld.admin.ch/isil/CH-001847-2` and a list of properties and values:

type	DefinedTerm
type	Isil
about	organization
dateModified	2019-02-06
inDefinedTermSet	isil
name	Aargauer Kantonsschule Wohlen / Mediothek
replacer	https://culture.ld.admin.ch/isil/
termCode	CH-001847-2
url	https://www.helveticarchives.ch/detail.aspx?ID=883304 (anyURI)
administrativeHistory	Die Mediothek der Kantonsschule Wohlen ist seit 1987 im

Abbildung 8: Darstellung in Form einer dynamischen Webseite eines Teils des Graphs der in LINDAS erfassten Daten für den ISIL-Code CH-001847-2 (siehe auch Abbildung 2).

Diese Funktionalität ermöglicht nicht nur, in diesem Datengraph zu navigieren, sondern garantiert auch die Echtheit dieser Daten, da nur der Eigentümer der Domäne «ld.admin.ch» ein solches Modul mit dem Triplestore verbinden kann, der diese Daten offiziell verwaltet.

5.3.2 Offene Schnittstelle SPARQL

[SPARQL](#) ist eine Abfragesprache für RDF-Daten, die durch [W3C](#) definiert ist. SPARQL-Abfragen werden über einen Zugangspunkt (eine URL), der als SPARQL-Endpoint bezeichnet wird, an einen Triplestore weitergeleitet.

Die SPARQL-Sprache ist so standardisiert, dass bei Abfragen die Triplestores verschiedener Lieferanten durchsucht werden können, ohne dass die Abfrage angepasst werden muss.

5.3.3 LINDAS.admin.ch

Die Website LINDAS umfasst eine Reihe von Informationen über Linked Data und die laufenden und geplanten Entwicklungen rund um LINDAS.

Eine der Seiten (lindas.admin.ch/sparql/) ist direkt mit den Daten des Triplestores P (Produktion) verbunden. Über diese Seite können SPARQL-Abfragen von Daten editiert und getestet werden, die in diesem Triplestore enthalten sind.

6 Das Ökosystem von LINDAS

Mit LINDAS wird die Absicht verfolgt, im Triplestore P die von den Verwaltungsstellen verwalteten Daten zu publizieren. Mit der Wahl von Lösungen, die auf offenen W3C-Standards basieren, wird anderen Akteuren die Gelegenheit gegeben, Tools anzubieten, die das LINDAS-Angebot ergänzen.

Die unten aufgeführte Applikation ist sehr gut mit LINDAS kombinierbar. Deshalb wird sie in diesem Abschnitt aufgeführt. Sie ist jedoch nicht Teil des LINDAS-Angebots, und ihre Nutzung ist in den jährlichen Nutzungskosten für LINDAS nicht enthalten.

6.1 Visualize

Das Bundesamt für Umwelt BAFU verwendet LINDAS, um Umweltdaten für Mensch und Maschine zugänglich zu machen. Daher ist das BAFU ein bevorzugter Partner des Projekts zur Einführung von LINDAS. Es hat die die Web-Applikation [visualize](#) entwickelt.

Mit Hilfe des Tools visualize können die in der Triplestore Infrastruktur abgelegten Daten ohne Linked Data spezifische Kenntnisse abgerufen und in Form von Tabellen, Grafiken oder Karten auf beliebigen Webseiten publiziert werden. Dieses Tool kann von anderen Verwaltungseinheiten genutzt und weiterentwickelt werden.

7 Dienstleistung des Bundesarchivs

7.1 Eckwerte des Angebots

Nutzung von LINDAS, namentlich:

- Publikation von Daten im Triplestore P (Produktion) und Verwendung der Triplestores I und P zu Testzwecken und zur Datenvalidierung.
- Unbegrenzter Zugang zum Data Cube Curator
- Hosting von Daten, die in Raw Data konvertiert werden sollen
- Hosting und Monitoring von Pipelines zur Konvertierung von Daten
- Unbegrenzter Zugang zum Schema Manager
- Dereferenzierung der in LINDAS gehosteten Daten, sofern sie «ld.admin.ch» als Basis nutzen.

Support und Wartung (1st level Support)

- Triplestore Infrastruktur (Test-, Integrations-, Produktionsumgebung)
- Zugang zu den Daten, die sich im Triplestore befinden (über die SPARQL Endpoints) für den Datenlieferanten, weitere Verwaltungseinheiten und die Öffentlichkeit
- Pipeline-Management (insbesondere das Monitoring)

Bis 8 Std. jährlich Beratung und Unterstützung pro Datenlieferant.

Gefäss für Informations- und Erfahrungsaustausch zwischen den Datenlieferanten.

Nur für Bundesverwaltung: Abruf Stunden RDF- und/oder Visualisierungsspezialisten in Absprache mit dem BAR. Die Rahmenverträge sind bis am 31.05.2025 gültig.

7.2 Triplestore Infrastruktur – Service Levels

7.2.1 Produktionsumgebung

Betriebszeit der Lösung

- Die Betriebszeit beträgt 24 Stunden pro Tag an 365 Tagen im Jahr. Ausgenommen sind in Anspruch genommene Wartungsfenster, standardmässig wöchentlich dienstags von 22:00 bis 02:00.

Verfügbarkeit der Lösung

- Für die Lösung ist während der Betriebszeit eine Verfügbarkeit von 99.9% garantiert.
- Pro Vorfall darf der Betriebsunterbruch nicht länger als 4 Stunden betragen.

Supportzeiten

- Montag bis Freitag, 09.00 - 17.00 Uhr, exkl. Feiertage

7.2.2 Test- und Integrationsumgebung

Betriebszeit der Lösung

- Die Betriebszeit beträgt 12 Stunden von Montag bis Freitag, 07.00 - 19.00 Uhr, 52 Wochen pro Jahr. Ausgenommen sind in Anspruch genommene Wartungsfenster.
- Die restliche Zeit läuft die Test- und Integrationsumgebung ohne Gewähr.

Verfügbarkeit der Lösung

- Für die Lösung ist während der überwachten Betriebszeit eine Verfügbarkeit von 98% garantiert.
- Pro Vorfall darf der Betriebsunterbruch nicht länger als 12 Stunden betragen.

Supportzeiten

- Montag bis Freitag, 09.00 - 17.00 Uhr, exkl. Feiertage

8 Mitwirkungspflichten des Datenlieferanten

Der Datenlieferant verpflichtet sich, seine Datensätze so zu dokumentieren, dass sie auf dem Portal opendata.swiss referenziert werden können. Bei der Wahl von Ontologien für seine Daten bevorzugt er im Rahmen des Möglichen die auf der Website LINDAS.admin.ch aufgeführten Ontologien. Bei allen massgeschneiderten Ontologien wählt er vorzugsweise die Applikation Schema Manager für deren Beschreibung.

9 Kosten

Gemäss Finanzhaushaltgesetz und den entsprechenden Weisungen sind Leistungen an Dritte zu kostendeckenden Preisen zu erbringen, d.h. die Grundlage für die Preiskalkulation bilden die Vollkosten. In Anhang 1 ([Anhang 1 LINDAS - Faktenblatt Kalkulation](#)) ist das Faktenblatt zur Kalkulation der jährlichen Vollkosten von LINDAS zu finden.

Die darauf basierende Kalkulation und Preisfestlegung ist in [Anhang 2 «LINDAS Kalkulation+Preis definitive Variante»](#) zu finden. Sie zeigt jährliche Vollkosten in der Höhe von rund 540'000.00 Fr. (auf der Basis von 5 Mrd. Triples).

Bundesinterne Kunden, welche den Linked Data Service LINDAS nutzen, sollen diese Leistungen mittels Kreditverschiebungen ans BAR begleichen. Bundesexternen Kunden werden die beanspruchten Leistungen mittels Rechnungsstellung durchs BAR verrechnet.

Für die Preisfestlegung werden drei Preisgefässe gebildet. Die Zuordnung zum jeweiligen Gefäss erfolgt auf Basis der voraussichtlichen Anzahl Triples, die ein Datenlieferant im Triplestore speichern wird. Die kalkulatorischen Überlegungen sind im [Anhang 2 «LINDAS Kalkulation+Preis definitive Variante»](#) festgehalten.

10 Vorgehen

Nehmen Sie bei Fragen und für weitere Informationen mit uns Kontakt auf.

Testen Sie Linked Data mit einem Datensatz. Die Testumgebung der Infrastruktur kann kostenlos genutzt werden.

Schliessen Sie eine Vereinbarung mit dem BAR ab und nutzen Sie die Infrastruktur.

Möchten Sie einen Use Case umsetzen, gehen Sie nach folgendem Raster vor:

- Zielsetzung, Kurzbeschreibung, Nutzen, grobes Mengengerüst, Datenformat etc.
- Transformation der Daten nach RDF (Erstellung der Ontologie, Wahl des Vokabulars, Transformation der Daten nach RDF, Verknüpfung mit Referenzdaten des Bundes in RDF (z. B. Swisstopo))
- Automatisierung der Daten-Pipeline
- SPARQL Abfragen (Erstellung von Beispielabfragen in SPARQL, Visualisierung)
- Zeitplan (Datum der Publikation der Daten, Testen vor der Publikation etc.)

Wir beraten Sie gerne bei der Abklärung, ob sie für ihren Use Case Unterstützung von Linked Data und/oder Visualisierungsspezialisten benötigen.

11 Kontakt

Pierre Dumas

Support-Lindas@bar.admin.ch

Tel. +41 58 465 05 60

Schweizerisches Bundesarchiv BAR

Archivstrasse 24, CH-3003 Bern

www.lindas.admin.ch

12 Anhänge

- [Anhang 1: LD-Plattform - Faktenblatt zur Kalkulation](#)
- [Anhang 2: LD-Plattform - Kalkulation und Preisfestlegung definitive Variante](#)
- [Mustervereinbarung](#)