LINKED DATA SCHWEIZ — KONFERENZ

I14Y × LINDAS
Die Governance wird zum Graphen

Mehrwert für das Schweizer Linked-Data-Ökosystem

Bundesamt für Statistik BFS
Fabrice Gouzi

09.11.2026 · Linked Data Day 2026 Berne

Programm: drei Akte

01 **Zusammenarbeit LINDAS × I14Y-metadata.swiss**
Pipeline, GitHub, Identifikation

02 **Mehrwert auf LINDAS**
Governance, Standard, gemeinsame Dimensionen, Beziehungen

03 **Zukunftsperspektiven**
Semantische Harmonisierung: z. B. Kanton / Geschlecht

I14Y: der semantische Vertrag, vor jeglichen Daten

I14Y definiert, was ein Code, ein Konzept, eine Struktur bedeuten — die Governance findet hier statt, auch im Dienst von LINDAS.

Konzepte & Codelisten

Versionierte, mehrsprachige Definitionen

SHACL-Strukturen

Eine Shape pro strukturiertem Datensatz

Publikationsstatus

Wer exportiert werden darf, und wann

Quelle: i14y.admin.ch — öffentlicher Katalog

1 · ZUSAMMENARBEIT LINDAS × I14Y – metadata.swiss

Zwei Systeme, eine Pipeline dazwischen

I14Y
Privater Graph — Quelle der Wahrheit

→

Tägliche Pipeline
2 Python-Exporter

→

LINDAS
Benannter Graph fso/i14y

concept2sharedDimension · dataset2lindas

24h
Standard-Aktualisierungsfenster

96h
am Montag, um das Wochenende abzudecken

I14Y_MODIFIED_LOOKBACK_HOURS — die Pipeline exportiert nur, was sich geändert hat

Quelle: github.com/metadata-swiss/LINDAS

Open Source: Die Pipeline liegt auf GitHub

Öffentliches Repository

github.com/metadata-swiss/LINDAS

Code der Exporter, Workflows, SPARQL-Referenzabfragen.

2 GitHub Actions

Konzepte: concept2sharedDimension
Datensätze: dataset2lindas

Montag bis Freitag. [Cron-Zeit zu bestätigen]

Schutzmechanismen

24-h-Fenster (96 h am Montag), Zählerkontrolle, Löschung dessen, was in I14Y verschwunden ist.

Alles ist reproduzierbar und überprüfbar: Man kann lesen, forken und Korrekturen vorschlagen.

Eine Idee, ein Bedarf, ein Bug? Reichen Sie eine Anfrage in unseren Issues ein.
github.com/metadata-swiss/feature-requests/issues

Quelle: https://github.com/metadata-swiss/LINDAS

Funktionsweise der Pipeline: Der Publikationsstatus entscheidet

I14Y-Datensatz / -Konzept → publicationLevel = Public ? → **Für die Pipeline berechtigt**

Mit einigen Ausnahmen: Eine Codeliste, deren Versionen alle leer sind, wird beispielsweise nicht exportiert.

Der benannte Graph und das IRI-Präfix

Der benannte Graph

GRAPH <https://lindas.admin.ch/fso/i14y>

Das IRI-Präfix

register.ld.admin.ch/i14y/{typ}/{identifikator}

Der I14Y-Identifikator wird zu einer eindeutigen IRI: Sie identifiziert jedes Objekt in LINDAS zweifelsfrei und hebt es vom Rest des Graphen ab.

1 · ZUSAMMENARBEIT LINDAS × I14Y – metadata.swiss

Diese Pipeline speist aktuell:

2 459

öffentliche Datensätze, in I14Y registriert

628

öffentliche Konzepte und Codelisten

55%

der Datensätze haben eine formale Struktur

Im Rahmen des Aufbaus der ersten Bausteine von metadata.swiss werden auch alle OGD-Datensätze in diese Pipeline integriert --> Horizont Q1 2027

Quelle: öffentliche I14Y-API (i14y.admin.ch)

① Governance: wer, welche Version, wann

Eigentümer

Jedes Konzept und jeder Datensatz hat einen Herausgeber (Publisher). Untersuchte Konzepte: **CH1**, Bundesamt für Statistik.

Version

Expliziter Lebenszyklus: Ein versioniertes Konzept ersetzt das vorherige. Die Versionen werden mit **version.link** verwaltet.

legalForm 1.2.0 · Standard
legalForm 1.1.0 · Superseded

Letzte Änderung

Datiert und damit in SPARQL abfragbar.

sex 1.0.0 · 2026-06-24
nogaCode 2.0.1 · 2026-05-07

In LINDAS ist die Governance kein PDF mehr, sondern ein Tripel, das man filtern, verknüpfen und überwachen kann.

Quelle: I14Y, öffentliche Konzepte (647 zum Zeitpunkt der Extraktion)

②Standard: I14Y sagt, welcher

180 Konzepte

mit Status **Standard** (172) oder **PreferredStandard** (8) von 647 öffentlichen Konzepten.

243 Konzepte

deklarieren ein **conformsTo** auf eine Norm oder Rechtsgrundlage: 145 nennen eCH, 4 nennen ISO.

Geschlecht · sex 1.0.0

Standard
eCH-0044 · Katalog der Personenmerkmale

legalForm 1.2.0

Standard
eCH-0108

NOGA 2025 · nogaCode 2.0.1

PreferredStandard
eCH-0108 · NACE

Schweizer Kantone

Recorded
kein conformsTo

Quelle: I14Y, Analyse von 647 öffentlichen Konzepten

Was sich leicht abfragen lässt

Alle publizierten Konzepte auflisten

Alle Codelisten-Identitäten in einem Durchgang

Einen Datensatz + seine Struktur rekonstruieren

Der vollständige Teilgraph, ohne externen Links zu folgen

Die Historie eines Codes verfolgen

Alle exportierten Versionen im Zeitverlauf

Nach Status filtern

Standard, Qualified, PreferredStandard…

Referenzabfragen im Anhang

Ein I14Y-Konzept wird zur «Shared Dimension»

In Cube Creator ist eine Spalte kein freier Text mehr, sondern eine wiederverwendbare URI zwischen Datensätzen.

I14Y-Konzept
z. B. legalForm

→

Shared Dimension
Cube Creator

Die vollständige Kette, vom Register bis zur Visualisierung

I14Y-Konzept → **Shared Dimension** → **Cube (Cube Creator)** → **visualize.admin.ch**

Der Exporter heisst concept2sharedDimension — kein Zufall bei der Namensgebung

③Eine gemeinsame Dimension wiederverwenden: NOGA in visualize

In I14Y

nogaCode 2.0.1
PreferredStandard

In LINDAS

Das Konzept wird als **gemeinsame Dimension** (Cube Creator) exportiert.

In visualize

Cubes, die die Dimension verwenden, teilen dieselben Bezeichnungen und dieselbe Hierarchie.

Quelle: I14Y-Konzept nogaCode 2.0.1 · LINDAS · visualize.admin.ch

④Datensätze über die technische Repräsentation verknüpfen

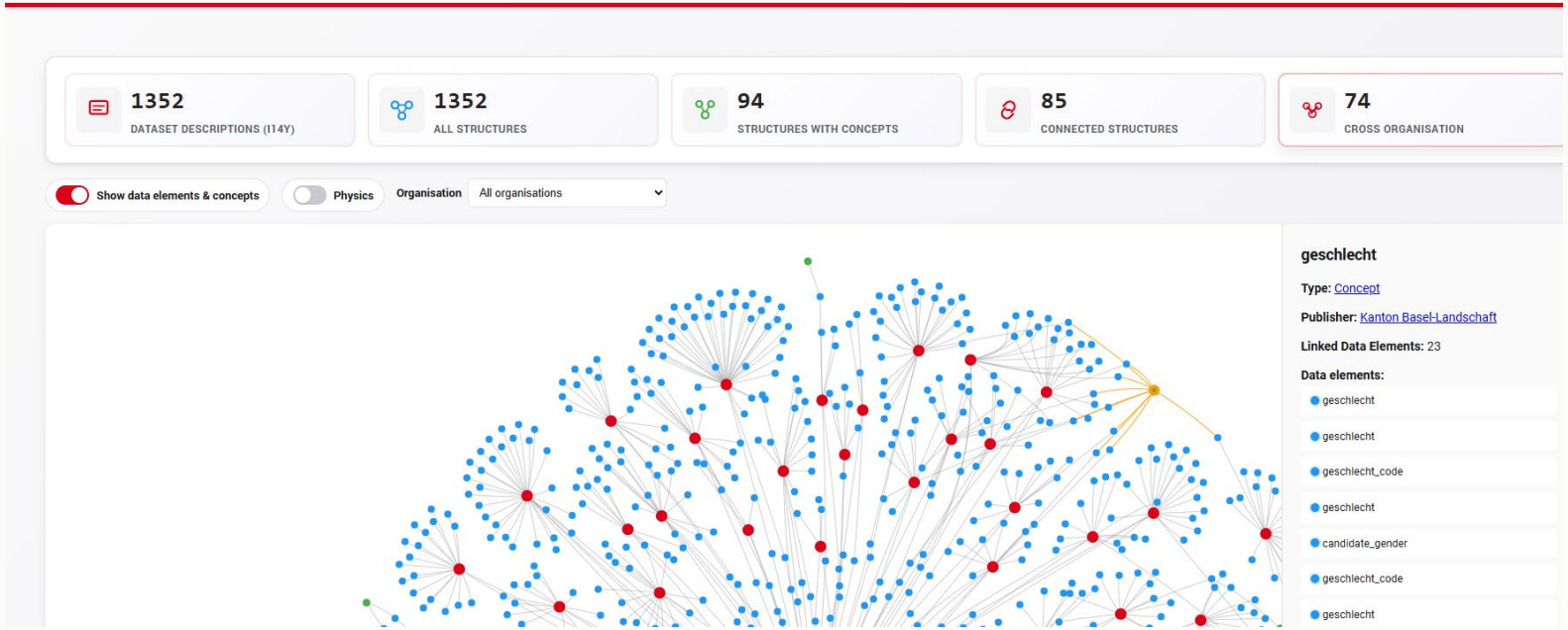
Prinzip

Ein Attribut eines Datensatzes ist mit einer **technischen Repräsentation** verknüpft, die ihrerseits mit dem Konzept verknüpft ist. Zwei Datensätze, die auf dieselbe Repräsentation verweisen, sind miteinander verbunden.

I14Y-Tool

harmgraph.toolbox.i14y.c.bfs.admin.ch

Erkundet die Beziehungen zwischen Datensätzen.



Quelle: I14Y-Tool harmgraph

Feature : semantische Verknüpfungen zu 100 %

Ein Konzept, mehrere technische Repräsentationen

Semantisches Konzept →

- x mögliche Werte · BFS - Codes
- Abkürzungen
- NUTS · Namen · …

Solange die Verknüpfung technisch ↔ semantisch nicht perfekt ist, bleibt die semantische Harmonisierung unerreichbar.

Feature : semantische Verknüpfungen zu 100 %

Fall 1: Der Kanton Zürich wird auf 4 Arten geschrieben

Abkürzung	BFS-Code	NUTS 3	Name (Sozialhilfe)
ZH	1	CH040	Zürich

Datensatz: ts-x-09.03.02.01.03b

Spalte **geo** = NUTS-3-Code

Datensatz: 13-02-01_Sozialhilfequote

Kanton über seinen **Namen**

In I14Y ist «Schweizer Kantone» Recorded, ohne conformsTo, und die Spalte GEO ist als einfache Zeichenkette deklariert.

Quelle: Dateien ts-x-09.03.02.01.03b und 13-02-01_Sozialhilfequote (2023), Konzept HGDE_KT 1.0.0

Feature : semantische Verknüpfungen zu 100 %

Zwei Dateien über den Kanton verknüpfen: 0 von 26

0 / 26

direkter Join NUTS 3 ↔ Name (Jahr 2023)

26 / 26

Join über die Kantonsliste (BFS-Code, NE, CH024, Name)

Die Kantonsliste dient als Pivot, liegt aber ausserhalb der Metadaten: Genau das kann metadata.swiss deklarativ machen.

Quelle: Berechnung auf den beiden Dateien, 26 Kantone

Feature : semantische Verknüpfungen zu 100 %

Fall 2: Das Geschlecht, ein Konzept, zwei Repräsentationen

Basel-Landschaft · Datensatz 10700

Gemeinderatswahlen 2024 : Kandidierendenresultate

geschlecht · xsd:string

Männlich	Weiblich

Basel-Stadt · Datensatz 100405

Kandidierende der Regierungsratswahlen, 2. Wahlgang

Geschlecht · xsd:string

f	m

Für einen Menschen ist der Sinn identisch. Für eine SPARQL-Abfrage ist «Weiblich» nicht «f».

Quelle: Dateien 10700 (Männlich, Weiblich) und 100405 (f beobachtet; m erwartet, im Auszug nicht vorhanden)

Feature : semantische Verknüpfungen zu 100 %

Fall 2: Das Geschlecht, ein Attribut in 23 Basler Datensätzen

23 Datensätze

Bevölkerung, Geburten, Heiraten, Wahlen… alle mit einem Attribut Geschlecht.

Geprüfte Struktur

Grossratswahlen 2024

Geschlecht · xsd:string

Beschreibung: «Amtliches Geschlecht». Kein conformsTo, keine Verknüpfung zum Konzept.

Ein Attribut, 23 Datensätze, ein einziger Sinn — aber keine Deklaration, dass es sich um dasselbe Konzept handelt.

Quelle: I14Y, Datensätze Statistisches Amt Basel-Stadt

Feature : semantische Verknüpfungen zu 100 %

I14Y liefert die Referenz: das Konzept «sex»

Basel-Landschaft (10700) conformsTo deklariert		Konzept I14Y «sex» v1.0.0 Bundesamt für Statistik · Standard		Basel-Stadt (100405) keine Verknüpfung deklariert
Männlich	→	1 · Männlich	←	m (erwartet)
Weiblich	→	2 · Weiblich	←	f
–		3 · Unbestimmt		–

BL schreibt die deutschen Bezeichnungen des Konzepts, nicht seine Codes. BS schreibt Buchstaben, ganz ohne Verknüpfung. Drei Schreibweisen für einen einzigen Sinn.

Feature : semantische Verknüpfungen zu 100 %

Ohne semantische Verknüpfung scheitert der Vergleich stillschweigend

Zeichenkettenvergleich

Weiblich ≠ f

Keine Übereinstimmung. Die beiden Kantone sind nach Geschlecht nicht vergleichbar ohne von Hand geschriebene Zuordnungstabelle.

Join über den Konzeptcode

Weiblich → 2 ← f

Die Kandidaturen beider Kantone werden nach Geschlecht vergleichbar, ohne Ad-hoc-Verarbeitung.

Die Zuordnung Wert → Code wird einmal auf Konzeptebene deklariert, statt von jedem Konsumenten neu geschrieben zu werden.

Feature : semantische Verknüpfungen zu 100 %

Wo stehen wir wirklich?

55%

der 2 459 öffentlichen Datensätze haben eine deklarierte SHACL-Struktur

6 / 22

Eigenschaften des untersuchten Basler Datensatzes verwenden tatsächlich conformsTo

Der Mechanismus existiert und funktioniert. Seine systematische Nutzung muss noch aufgebaut werden — Datensatz für Datensatz, Produzent für Produzent.

Wir haben ein Feature, mit dem sich diese semantische Verknüpfung zwischen einfachen Objekten (100%-ige Übereinstimmungen) herstellen lässt. Es sollte priorisiert werden.

3 · PERSPEKTIVEN

Perspektive: die semantischen Verknüpfungen stärken

Die semantischen Verknüpfungen zwischen den verschiedenen technischen Repräsentationen (Codes, Abkürzungen, NUTS, Namen usw.) und der semantischen Repräsentation, dem Konzept, verbessern.

1 · Deklarieren

Jedes Attribut mit seiner technischen Repräsentation und seinem Konzept verknüpfen.

2 · Verknüpfen

Zuordnungen von Wert zu Wert, zuerst die einfachen Fälle zu 100 %.

3 · Nutzen

Joins und semantische Harmonisierung, danach Reasoning.

SCHLUSSFOLGERUNG

Was man sich merken sollte

I14Y – metadata.swiss ist für die Governance zuständig, LINDAS verbreitet — zwei Verantwortlichkeiten, ein einziges Ökosystem

Die Pipeline ist ein verifizierter Spiegel, kein einfacher Export

Die geteilte Semantik eliminiert das manuelle Mapping — und eröffnet neue Perspektiven im Bereich der semantischen Interoperabilität

ANHANG

SPARQL-Referenzabfragen

Quelle: github.com/metadata-swiss/LINDAS — SPARQL_EXAMPLES.md

Alle Konzeptidentitäten auflisten

```
PREFIX schema: <http://schema.org/>
PREFIX vl: <https://version.link/>

SELECT ?concept ?identifier
WHERE {
  GRAPH <https://lindas.admin.ch/fso/i14y> {
    ?concept a schema:DefinedTermSet, vl:Identity ;
             schema:identifier ?identifier .
  }
}
ORDER BY ?identifier
```

Alle publizierten Datensätze auflisten

```
PREFIX dcat: <http://www.w3.org/ns/dcat#>
PREFIX dct: <http://purl.org/dc/terms/>

SELECT ?dataset ?identifier
WHERE {
  GRAPH <https://lindas.admin.ch/fso/i14y> {
    ?dataset a dcat:Dataset ;
             dct:identifier ?identifier .
    FILTER(STRSTARTS(STR(?dataset),
      "https://register.ld.admin.ch/i14y/dataset/"))
  }
}
ORDER BY ?identifier
```

Die aktuelle Version eines Codes (z. B. NOGA «G»)

```
PREFIX schema: <http://schema.org/>
PREFIX vl: <https://version.link/>

SELECT ?currentCodeVersion
WHERE {
  GRAPH <https://lindas.admin.ch/fso/i14y> {
    ?concept a schema:DefinedTermSet, vl:Identity ;
             schema:identifier "nogaCode" ;
             schema:hasDefinedTerm ?codeIdentity .
    ?codeIdentity a schema:DefinedTerm, vl:Identity ;
             schema:identifier "G" ;
             vl:Version ?currentCodeVersion .
  }
}
```