

Archival Linked (Open) Data: Empfehlungen für bestehende Metadaten und Massnahmen für die Zukunft am Fallbeispiel des Schweizerischen Sozialarchivs

Würtz, Fabian

Archival Linked (Open) Data: Empfehlungen für bestehende Metadaten und Massnahmen für die Zukunft am Fallbeispiel des Schweizerischen Sozialarchivs

Informationswissenschaft: Theorie, Methode und Praxis, vol. 6, núm. 1, 2020

Universität Bern, Suiza

Disponible en: <https://www.redalyc.org/articulo.oa?id=664372382016>



Esta obra está bajo una Licencia Creative Commons Atribución 4.0 Internacional.

Archival Linked (Open) Data: Empfehlungen für bestehende Metadaten und Massnahmen für die Zukunft am Fallbeispiel des Schweizerischen Sozialarchivs

Fabian Würtz

Einleitung

Informationswissenschaft: Theorie,
Methode und Praxis, vol. 6, núm. 1, 2020

Universität Bern, Suiza

Redalyc: [https://www.redalyc.org/
articulo.oa?id=664372382016](https://www.redalyc.org/articulo.oa?id=664372382016)

Eine Kernaufgabe der Archive ist die Erschliessung des Archivguts. Dank ihr wird das Archivgut besser verständlich und auffindbar. Erschlossenes Archivgut macht Entstehungsprozesse durchschaubar und stellt Transparenz her. Bisher wurden Archivbestände meist als hierarchische und isolierte Einheiten verzeichnet. Die zunehmende Digitalisierung, neue Fachbereiche wie die Digital Humanities oder Entwicklungen wie das semantische Web bzw. Linked Open Data haben jedoch neue Ideen in die Archivwelt getragen.

Einer der deutlichsten Vorboten dieser neuen Welt ist *Records in Context* (RiC). Der neue Verzeichnungsstandard des wichtigen *International Council on Archives* (ICA) ist konzeptionell auf Linked Open Data und das Semantic Web ausgerichtet. Doch was bedeutet es für die Archive, wenn aus den bisher isolierten Beständen verlinkte und maschinenlesbare Netzwerke entstehen sollen? Wie sollen archivarische Metadaten und Datenmodelle in Linked Open Data aussehen und welche Qualitätsansprüche sollen diese neu berücksichtigen?

Oder konkret gefragt: Welche Massnahmen sollen bezüglich bestehender Metadaten und welche bezüglich zukünftiger Bestände getroffen werden?

Um diese Frage zu beantworten, geht die vorliegende Arbeit wie folgt vor: Zunächst werden das Konzept und die Technologien, die Linked Open Data zugrunde liegen, anhand der aktuellen wissenschaftlichen Literatur vorgestellt. Im zweiten Kapitel geht die Arbeit der Frage nach, welches die Merkmale für die Datenqualität bei Linked Open Data sind. Anschliessend wird der momentane Stand von Linked Open Data im Archivbereich beleuchtet. Dabei sollen auch bereits existierende Anwendungen vorgestellt und analysiert werden. Aufbauend auf den resultierenden Erkenntnissen wird im vierten Kapitel anhand des Fallbeispiels der Metadaten des Schweizerischen Sozialarchivs eine Linked-Open-Data-Modellierung erstellt. Dabei soll untersucht werden, wie sich die gesammelten Qualitätsmerkmale auf die Praxis übertragen

lassen. Die Arbeit schliesst mit einigen allgemeinen Empfehlung für die Archive.

Die Relevanz des Themas ergibt sich aus dem neuen Verzeichnungsstandard RiC und dem digitalen Wandel, dem viele Archive unterworfen sind. Wenn sich der Standard durchsetzt, könnte dies in vielerlei Hinsicht einen Paradigmenwechsel in der Archivwelt bedeuten. Umso wichtiger ist es, dass sich die Archivar*innen mit dem Thema auseinandersetzen und über Chancen und Herausforderungen diskutieren. Das Ziel dieser Arbeit ist es, einen Beitrag zu dieser Diskussion zu leisten.

Die aktuelle wissenschaftliche Literatur zu dem Thema kann grundlegend in drei Bereiche eingeteilt werden: (1) Zum Thema Linked (Open) Data existiert eine grosse Fülle von sowohl allgemeinen ¹ als auch spezifischen Werken, von denen die meisten in englischer Sprache verfasst sind. Speziell zu nennen sind die diversen Artikel und Dokumentationen der W3C. (2) Daneben existieren mehrere Bücher und Artikel zum Thema Linked Open Data im Kulturerbe-Bereich. ² Hervorzuheben ist dabei insbesondere das Buch «(Open) Linked Data in Bibliotheken» ³. (3) Über Linked Open Data in Archiven gibt bisher eher wenig Literatur, die sich vor allem aus Artikeln in Fachzeitschriften und Blogbeiträgen zusammensetzt. Erwähnenswert ist hier- bei die Arbido-Ausgabe 2013/4 mit dem Thema «Linked open data, big data, alles vernetzt». ⁴ Sie enthält zahlreiche Artikel in deutscher und französischer Sprache zu dem Themengebiet.

Linked Open Data

Dieses Kapitel soll eine Einführung in das Thema Linked Open Data bieten. Dazu werden zunächst die geschichtlichen Wurzeln und das grundlegende Konzept des Semantic Webs sowie der Bereiche Linked Data und Open Data vorgestellt. Anschlies- send werden vier wichtige Komponenten des Bereichs erläutert. Dabei handelt es sich um URIs, RDF, Vokabulare und Ontologien sowie SPARQL.

Semantic Web & Linked Data

Die Idee des Semantic Web ergab sich aus dem Problem, dass Informationen im Internet für Maschinen oft schwer verständlich sind ⁵. Enthält eine Website beispiels- weise im Rahmen von natürlicher Sprache das Wort «Ente», kann ein Mensch meist problemlos aus dem Kontext heraus erkennen, ob das Tier, der Autotyp, der Fluss oder die niederländische Judoka gemeint ist. Für ein Computer-Programm stellt diese Unterscheidung jedoch eine Herausforderung dar.

Um dieses Problem anzugehen, veröffentlichte das *World Wide Web Consor- tium (W3C)* ⁶ das Konzept des Semantic Web ⁷. Die treibende Kraft hierbei war dessen Vorsitzender Tim Berners-Lee. Bei diesem Konzept ging es um formulierte Grundsätze zu einer Verbesserung

der Lesbarkeit von Informationen im Internet für Maschinen. Darüber hinaus beinhaltete es eine weitreichende Zukunftsvision: Dank der maschinenlesbaren Informationen sollten sogenannte Semantic-Web-Agenten (Bots) dazu befähigt werden, Webnutzende bei ihrer Suche im Internet intelligent zu unterstützen. Die Agenten sollten zum Beispiel dazu in der Lage sein, bei der Suche nach einem passenden Arzt in einem nahen Krankenhaus zu helfen ⁸.

Nachdem die Vision des Semantic Web sich aus technischen Gründen nur ansatzweise umsetzen liess, verebbte das Interesse ab 2001 zunehmend ⁹. Um dieser Entwicklung entgegenzuwirken, veröffentlichte das W3C 2006 ein neues Konzept, in dem der Begriff Linked Data eingeführt wurde. ¹⁰ Kern des Konzepts waren die durch Berners-Lee formulierten Linked-Data-Prinzipien:

1. *Use URIs as names for things.*
2. *Use HTTP URIs so that people can look up those names.*
3. *When someone looks up a URI, provide useful information, using the standards (RDF*, SPARQL).*
4. *Include links to other URIs, so that they can discover more things.»* ¹¹

Als eine der ersten grossen Linked-Data-Plattformen startete die DBpedia im Januar 2007. ¹² Durch das Extrahieren von Daten aus der Wikipedia konnte sie schnell einen grossen Datensatz aufbauen. Wie klein die Linked-Open-Data-Welt zu diesem Zeitpunkt noch war, zeigt die erste Visualisierung des Projekts Linked Open Data Cloud: Die Website, welche es sich zum Ziel gesetzt hat, das gesamte Linked-Open-Data-Netz zu visualisieren, führte 2007 gerade einmal zwölf Datensätze auf (siehe Abb. 1).

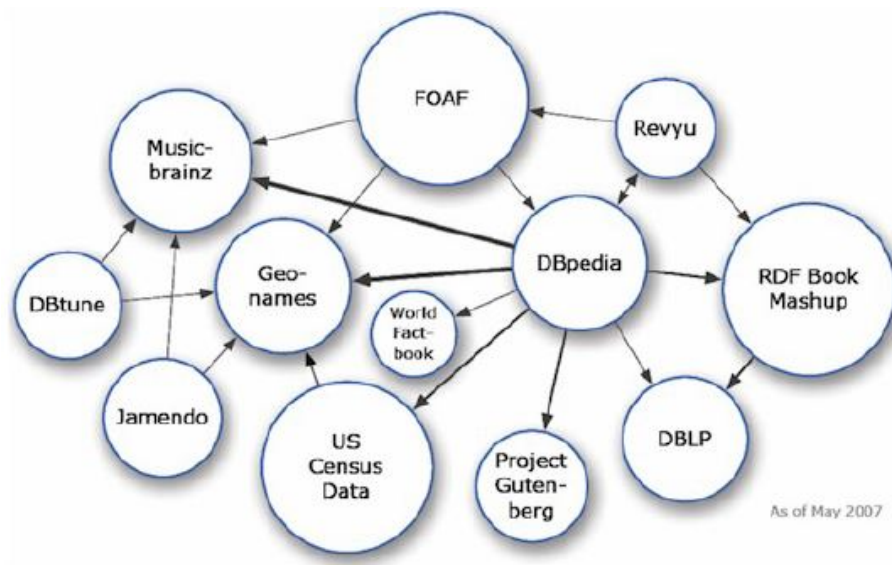


Abb. 1
Erste Version der Linked-Open-Data-Cloud (Mai 2007)

Zurverfügungstellung die Verwendung von Open Data nicht zwingend notwendig. Will man jedoch Daten verschiedener Quellen kombinieren und eine möglichst vielfältige Nutzung gewährleisten, so liegen die Vorteile von Open Data auf der Hand. Tim Berners-Lee begann erstmals 2009, für die Verbindung von Open Data und Linked Data zu werben.¹⁷ Ein Jahr später ergänzte er die Linked-Data-Prinzipien um ein Fünf-Sterne-Schema für Linked Open Data:

«★ Available on the web (whatever format) but with an open license, to be Open Data

★★ Available as machine-readable structured data (e.g. excel instead of image scan of a table)

★★★ As (2) plus non-proprietary format (e.g. CSV instead of excel)

★★★★ All the above, plus: Use open standards from W3C (RDF and SPARQL) to identify things, so that people can point at your stuff

★★★★★ All the above, plus: Link your data to other people's data to provide context»¹⁸

HTTP und Uniform Resource Identifier (URI)

Das *Hypertext Transfer Protocol (HTTP)* stellt das Protokoll zur Datenübertragung dar, auf dem das World Wide Web basiert. Auch Linked Open Data verwendet dieses Protokoll. Es kann somit als Erweiterung des World Wide Web verstanden werden.¹⁹ Eine *Uniform Resource Identifier (URI)* ist ein eindeutiger Identifikator einer abstrakten oder physischen Ressource. Sie wurde ursprünglich von Tim-Berns Lee 1994 im RFC 1630 definiert.²⁰ Die heute gültige Definition findet sich im RFC 3986²¹. Wie in Abb. 3 zu sehen ist, besteht eine URI aus einem Schema, einer Autorität, einem

Pfad, einer Query und einem Fragment:

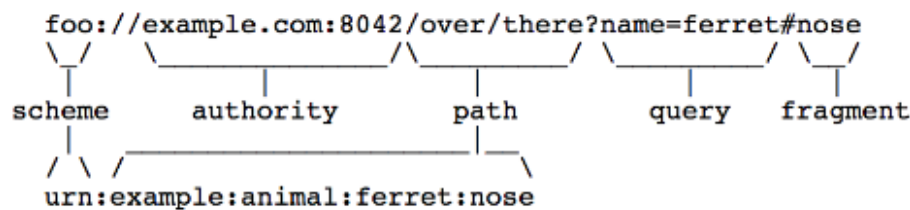


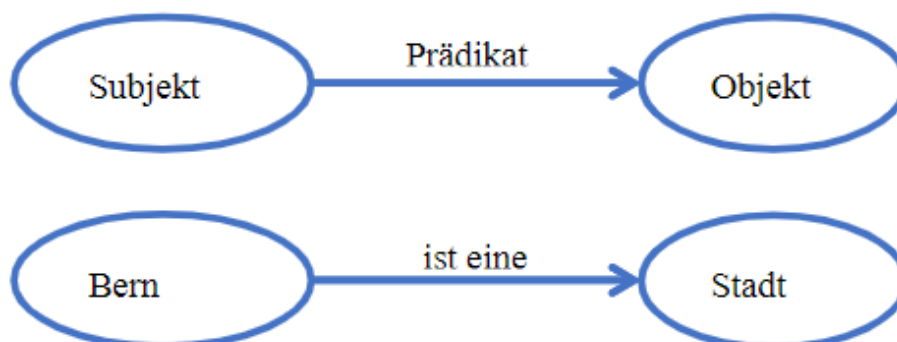
Abb. 3
Beispiel einer URI²²

URIs sind vor allem als Adressen für Internet-Seite bekannt. Sie werden aber auch für zahlreiche andere Anwendungsfälle benutzt, zum Beispiel für E-Mail-Adressen (<mailto:fred@example.com>) oder als Digital-Object-Identifizier (<doi:10.1000/182>).

RDF

Der zentrale Baustein von Linked Open Data ist das *Resource Description Framework (RDF)*. Es handelt sich dabei um ein Modell zur

Formulierung von logischen Aussagen über beliebige Zusammenhänge. Die Aussagen werden durch ein Subjekt, ein Prädikat und ein Objekt ausgedrückt, weswegen man auch von semantischen Tripeln spricht. Die Aussage, dass Bern eine Stadt ist, kann zum Beispiel wie folgt ausgedrückt werden:



Tripel sind vollkommen flexibel; jedes beliebige Subjekt kann mit einer frei wählbaren Relation mit jedem Objekt verbunden werden. Durch das einfache Verlinken der Daten erreicht das Modell somit eine sehr hohe Flexibilität.

Das Subjekt und das Prädikat sind immer Ressourcen. Eine Ressource ist etwas, worüber eine Aussage (in Form von Tripeln) gemacht werden kann, z.B. eine Person, ein Bauwerk, ein Buch. Ressourcen werden mithilfe einer URI als Identifikatoren referenziert.

Ein semantisches Tripel kann wie folgt aussehen:

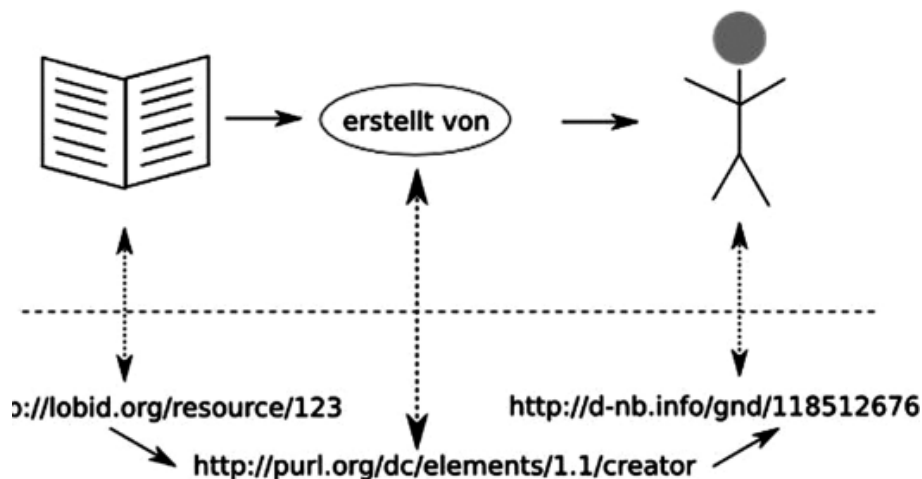


Abb. 4

HTTP-URIs als Identifikatoren raumzeitlicher Dinge

Objekte können, wie in Abb. 4 dargestellt, auch eine Ressource oder ein Literal (Frei-text) sein. Wie in RDF 1 zu sehen ist, lassen sich Literale mit Typen- (Zeile 4) und Sprachenbezeichnungen (Zeile 5) versehen. Mit Literalen können Eigenschaften wie der Vorname einer Person ausgedrückt werden. Seit der Version 1.1 bietet RDF mit «Named Graphs» zusätzlich die Optionen, ein viertes Element zu den Tripeln

hinzufügen. Wie in Kapitel 0 gezeigt wird, kann somit zum Beispiel die Provenienz eines Tripels ausgedrückt werden.

Für RDF ist keine textuelle Darstellung festgeschrieben. Wie in Tabelle 1 zu sehen ist, existieren aber noch eine Reihe weiterer Möglichkeiten. Die meisten davon besitzen erweiterte Formate zur Darstellung von Named Graphs.

Tabelle 1
Formate zur textuellen Darstellung für Linked Open Data

Format	Äquivalent für Named Graphs
RDF/XML	TriX
Turtle	TriG
N-Triples	N-Quads
Notation3	
JSON-LD	

Wie in den beiden Beispielen RDF 1 (Turtle) und RDF 2 (TriG) zu sehen ist, wird im Rahmen dieser Arbeit für Beispiele das Format Turtle verwendet, beziehungsweise TriG für Beispiele mit Named Graphs.

RDF 1

Auszug des Eintrags von Albert Einstein in der DBpedia.²³

1	@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
2	
3	
4	<http://dbpedia.org/ontology/bo:birthDate> "1879-03-
5	14"^^xsd:date ;
5	<http://xmlns.com/foaf/0.1/givenName> "Einstein"@en .

RDF 2

Auszug des Eintrags von Albert Einstein in der DBpedia mit einem Named Graph.

1	@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
2	
3	<http://example.com/graph1> {
4	<http://dbpedia.org/resource/Albert_Einstein>
5	<http://dbpedia.org/ontology/bo:birthDate> "1879-03-
6	14"^^xsd:date ;
6	<http://xmlns.com/foaf/0.1/givenName> "Einstein"@en . }

Vokabulare und Ontologien

Wie im letzten Unterkapitel bereits angedeutet, lassen sich mit RDF Aussagen einfach formulieren:

RDF 3

Verlinkung eines Titels

1	<code><http://example.com/id/1 > <http://example.com/id/hasTitle> "Titel" .</code>
---	--

Es fehlt jedoch im RDF-Modell die Möglichkeit, Terme wie zum Beispiel `hasTitle` zu formalisieren. Diese zusätzliche Funktion bieten Vokabulare und Ontologien an.

RDF 4

Nutzung des «DCMI Metadata Terms»-Vokabulars

1	<code><http://example.com/id/1> <http://purl.org/dc/terms/title> "Titel" .</code>
---	---

Im Sinne einer übersichtlichen Schreibweise können diese in der Turtle-Schreibweise mit Präfixen versehen werden:

RDF 5

Nutzung des «DCMI Metadata Terms»-Vokabulars unter Verwendung eines Präfixes.

1	<code>@prefix dcterms: <http://purl.org/dc/terms/> .</code>
2	
3	<code><http://example.com/id/1> dcterms:title "Titel" .</code>

Vokabulare und Ontologien können durch RDF-Schemata beschrieben werden.

RDF 6

Auszug des «DCMI Metadata Terms»-Vokabulars²⁴

```

1 @prefix dcterms: <http://purl.org/dc/terms/> .
2 @prefix rdf: -rdf-syntax-ns#> .
3 @prefix skos: <http://www.w3.org/2004/02/skos/core#> .
4 @prefix rdfs: -schema#> .
5
6
7 dcterms:modified "2012-06-
8 14"^^<http://www.w3.org/2001/XMLSchema#date> ;
9 dcterms:publisher <http://purl.org/dc/aboutdcmi#DCMI> ;
10 dcterms:title "DCMI Metadata Terms - other"@en .
11
12 dcterms:title
13 dcterms:hasVersion -002> ;
14 dcterms:issued "2008-01-
15 14"^^<http://www.w3.org/2001/XMLSchema#date> ;
16 dcterms:modified "2010-10-
17 11"^^<http://www.w3.org/2001/XMLSchema#date> ;
18 a rdf:Property ;
19 rdfs:comment "A name given to the resource."@en ;
20 rdfs:isDefinedBy <http://purl.org/dc/terms/> ;
21 rdfs:label "Title"@en ;
22 rdfs:range rdfs:Literal ;
23 rdfs:subPropertyOf <http://purl.org/dc/elements/1.1/title> .

```

Eine klare Trennlinie der beiden Begrifflichkeiten Vokabular und Ontologie existiert nicht. In der Praxis geht der Trend dahin, das Wort Ontologie für komplexere und formale Zusammenstellungen zu verwenden.²⁵

Das Erstellen oder die Veröffentlichung eines Vokabulars oder einer Ontologie steht jedem offen, es gibt hierfür keine zentrale, kontrollierende Instanz. Allerdings existieren Empfehlungen für die Erstellung von Vokabularen und Ontologien, auf die in Kapitel 3 näher eingegangen wird.

Die Wichtigkeit der Verwendung von Vokabularen und Ontologien für Datensets liegt darin begründet, dass Beziehungen auf diese Weise stärker formalisiert und dokumentiert werden können. Durch die Verwendung häufiger Vokabulare und Ontologien ist ein Datenset zudem für die Benutzer schneller verständlich und es kann einfacher mit anderen Datensets kombiniert werden. Dies ist wichtig für die Interpretierbarkeit der Daten und stellt somit einen essenziellen Aspekt der Datenqualität dar. Das Konzept von Linked Open Data nimmt keine Priorisierung von Vokabularen oder Ontologien vor. Dennoch gibt es eine Reihe von Vokabularen oder Ontologien, die aufgrund ihres Themengebiets und ihrer Bekanntheit häufig verwendet werden. In der Literatur zu Linked Open Data finden sich entsprechende Zusammenstellungen.²⁶ Dabei handelt es sich weniger um eine feste Liste als vielmehr um Erfahrungswerte, die je nach Autor voneinander abweichen.

Eines der meistgenannten Vokabulare ist das «Dublin Core Metadata Initiative (DCMI) Metadata»-Vokabular. Es definiert häufig verwendete Metadaten-Attribute wie title, creator, date und subject.

Ebenfalls regelmässig aufgeführt wird das «Friend- of-a-Friend (FOAF)»-Vokabular, mit welchem sich Personen und deren Beziehungen zu anderen Personen und Objekten beschreiben lassen. Für die Darstellung von Provenienz wird häufig die «The PROV Ontology (PROV)»-Ontologie angewandt.

Ein Verzeichnis der in dieser Arbeit verwendeten Vokabulare und Ontologien sowie ihrer verwendeten Präfixe und Namespaces findet sich auf S. 392.

SPARQL

SPARQL ist eine Graphen-basierte Abfragesprache für RDF. Das rekursive Akronym steht für SPARQL Protocol And RDF Query Language.²⁷ Die Erstveröffentlichung erfolgte 2008 durch das W3C.²⁸ Fünf Jahre später folgte die Version 1.1.²⁹

Die Sprache dient der Abfrage von Tripel-Stores, wobei es sich um Datenbanken zur Speicherung von semantischen Tripeln handelt. Eine SPARQL-Abfrage in Wikidata, die die Namen und geographischen Koordinaten (wdt:P625) aller Flughäfen (wd:Q1248784) in Belgien (wd:Q31) anzeigt, kann zum Beispiel wie folgt aussehen:

SPARQL 1

Abfrage aller Flughäfen in Belgien in Wikidata.³⁰

```
1 SELECT DISTINCT ?airport ?airportLabel ?coord
2 WHERE {
3   ?airport wdt:P31 wd:Q1248784 ;
4   ?range wd:Q31;
5   wdt:P625 ?coord.
6   SERVICE wikibase:label {
7     bd:serviceParam wikibase:language "[AUTO_LANGUAGE],en".
8   }
9 }
```

Die Abfragen können beliebig komplex sein. So kann zum Beispiel in Wikidata eine SPARQL-Abfrage gestellt werden, die alle Wiener Komponisten und ihre Kompositionen nach Tonart auflistet. Voraussetzung für eine SPARQL-Abfrage ist das Vorhandensein der Daten sowie deren standardisierte Erfassung.

Qualitätsmerkmale von Linked Open Data

Soweit es dem Autor dieser Arbeit bekannt ist, katalogisiert in der heutigen Praxis kein Archiv sein Archivgut direkt als Linked Data. Möchte ein Archiv seine katalogisierten Daten (Metadaten) als Linked Open Data veröffentlichen, muss es diese folglich zunächst umformen. Wie bereits in Kapitel 1.1 gezeigt, müssen dafür primär die Daten in semantische Tripel umgewandelt werden. Damit die migrierten Metadaten später allerdings sinnvoll als Linked Data genutzt

werden können, gilt es, diverse Punkte zu beachten. Hierzu soll zunächst allgemeinen auf den Begriff der Datenqualität und seine Bedeutung eingegangen werden. Danach folgt auf Basis der Best-Practise-Empfehlungen der W3C ³¹ eine Vorstellung und Diskussion der wichtigsten Punkte.

Definition der Datenqualität und ihre Bedeutung für Linked Open Data

Eine in der Literatur häufig verwendete Definition der Datenqualität ist diejenige von Wang und Strong. Sie teilen die Datenqualität in vier Aspekte ein: Zugänglichkeit, Interpretierbarkeit, Relevanz und Akkuratess ³². Diese Merkmale sind für die Archive nicht neu; auch in den bisherigen Archivsystemen sind alle vier Aspekte von Bedeutung. Dennoch gibt es einen wesentlichen Unterschied in deren Bezugssystem. Während die Archivplattformen diese Aspekte bislang lediglich für Menschen erfüllen mussten, müssen diese bei Linked Open Data nun auch für Maschinen sichergestellt werden. Will man also eine gute Datenqualität für Linked Open Data erreichen, so muss man den Begriff nicht neu erfinden, jedoch ihn aus der Sicht der Maschinen- Lesbarkeit denken.

Auswahl des Datensets, der Lizenzierung und der Sperrfristen

Die Frage nach dem Datenset ist für Archive leicht zu beantworten: Grundsätzlich sind alle Metadaten zum Archivgut interessant. Möchte ein Archiv jedoch nicht nur eine interne Linked-Data-Anwendung erstellen, sondern seine Metadaten als Linked Open Data veröffentlichen, müssen hierfür die rechtliche Situation sowie die Sperrfristen ³³ der Metadaten beachtet werden:

Um einen Linked-Open-Data-Dienst zu betreiben, müssen die Daten als Open Data vorliegen. Dabei sollte bedacht werden, dass Metadaten urheberrechtlich geschützt sein können. Im schweizerischen Recht hängt dies von der Frage ab, ob die jeweiligen Informationen als eine geistig-schöpferische Leistung mit individuellem Charakter angesehen werden können und somit Werkcharakter aufweisen (Art. 2 Abs. 1 URG). ³⁴

Auch wenn der Werkcharakter bei den Metadaten in Archiven angezweifelt werden kann, sollte es für Datenanbietende das Ziel sein, den Benutzenden eine möglichst hohe Rechtssicherheit anzubieten. Dies kann erreicht werden, indem alle Metadaten mit Lizenzen versehen werden. Aus Sicht der Benutzer und im Sinne der Kombinierbarkeit mit externen Datensets ist die Verwendung von häufig verwendeten und möglichst offenen Lizenzen zu bevorzugen. Diese Kriterien erfüllen zum Beispiel die Creative-Commons-Lizenzen CC0 oder CC-BY. Liegen verschiedene Lizenzen vor, zum Beispiel durch den Einbezug von Fremddaten, sollte dies klar ausgezeichnet werden. Die Wahl einer möglichst freien Lizenz entspricht auch dem Datenqualitätsmerkmal der Zugänglichkeit.

Des Weiteren müssen bei der Wahl des Datensets auch Sperrfristen berücksichtigt werden, sofern diese die Metadaten betreffen. Da dies allerdings auch schon bei den heutigen Archivinformationssystemen berücksichtigt werden sollte, sollte diese Anforderung für Archive kein wesentliches Problem darstellen.

Erstellung eines Datenmodells mit Vokabularen und Ontologien

Bei Linked Open Data spielt beim Aufbau des Datenmodells die Auswahl der verwendeten Vokabulare oder Ontologien eine zentrale Rolle. Eine gute Auswahl trägt zur Verständlichkeit des Datenmodells sowie zur Interoperabilität zu anderen Quellen bei. Doch wie sollte eine solche Wahl vorgenommen werden? Zunächst stellt sich hierbei die grundlegende Frage, ob auf Bestehendem aufgebaut oder ein neues Vokabular beziehungsweise eine neue Ontologie erstellt werden soll. Die W3C hält dazu fest: «It is best practice to use or extend an existing vocabulary before creating a new vocabulary.»³⁵

Wie kann nun unter der Vielfalt der Vokabulare und Ontologien die richtige gefunden werden? Allein die Website lov.okfn.org führt über 648 Linked-Data-Vokabulare auf. Diese Menge macht die Wahl des bestgeeignetsten Vokabulars oder der optimalen Ontologie schwer. Um dieses Problem anzugehen, schlagen Heath und Bizer vier Faktoren vor, nach denen die Auswahl erfolgen soll:³⁶

1. Die breite Verwendung und die Erlernbarkeit eines Vokabulars.
2. Die aktive Pflege und Bewirtschaftung eines Vokabulars.
3. Der Grad der Daten, die ein Vokabular abdeckt.
4. Die Expressivität gegenüber den Daten. Diese sollte weder zu hoch noch zu tief sein.

Die W3C führt eine Liste mit einem Muss- und sieben Soll-Kriterien aus:³⁷

Ein Vokabular ...

1. muss dokumentiert sein,
2. soll selbsterklärend sein,
3. soll in mehreren Sprachen beschrieben sein,
4. soll von anderen Datensets verwendet werden,
5. soll für eine lange Zeit verfügbar sein,
6. soll von einer vertrauenswürdigen Gruppe oder Organisation publiziert sein,
7. soll persistente URLs verwenden,
8. soll eine Versionierung haben.

Dokumentation des Datenmodells

Sowohl für interne Zwecke als auch im Sinne der Benutzer sollte das Datenmodell möglichst ausführlich dokumentiert sein. Dank der Verwendung von geeigneten Vokabularen und Ontologien ist dies bei Linked Data schon zu einem gewissen Teil gegeben. Bei der Verwendung mehrerer Vokabulare und Ontologien sollte das Gesamt- datenmodell zusätzlich gut dokumentiert sein. Falls Regeln zu einer bestimmten Verzeichnungsart einzelner Literale oder Ressourcen existieren, sollten diese ebenfalls festgehalten werden.

Festlegung des Schemas der URIs

Für die Ausgestaltung von URIs schlägt die W3C verschiedene Prinzipien vor. Zunächst sollen URIs das Protokoll HTTP verwenden.

³⁸ HTTPS soll grundsätzlich nicht verwendet werden, sofern aber Client und Server HTTPS unterstützen, soll die Antwort als HTTPS erfolgen.

³⁹ URIs sollen über WWW abrufbar sein. URIs sollen je Anfragetyp verschiedene Repräsentation anbieten, von denen mindestens eine maschinenlesbar ist. So soll zum Beispiel ein Web-Browser eine HTML-Seite anzeigen und eine Semantic-Web-Anwendung RDF erhalten.

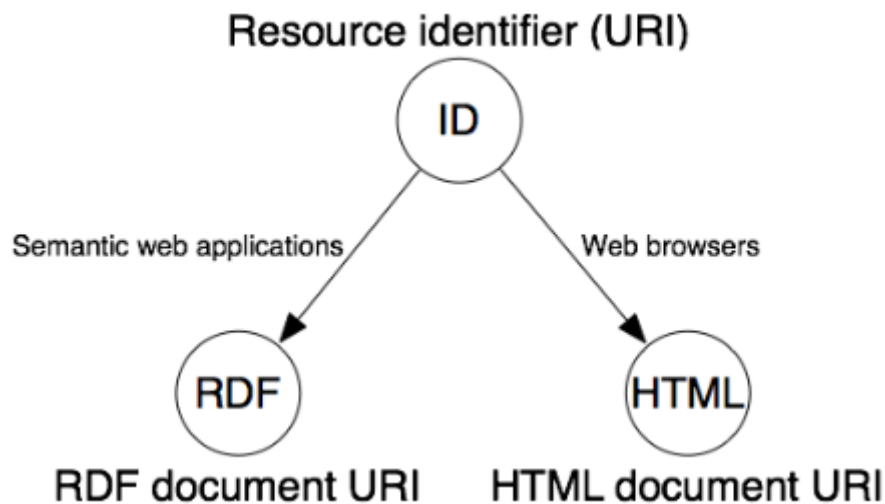


Abb. 5

URI mit mehreren Repräsentationen⁴⁰

Technisch kann dies durch die im HTTP-Protokoll eingebaute Content-Negotiation realisiert werden. Die Content-Negotiation ermöglicht es, bei einer Anfrage anzugeben, welche Art von Antwort man bevorzugt. Gibt ein Client zum Beispiel den Mime-Type `application/rdf+xml` an, weiss der Server, dass der Client gerne einen RDF/XML-Repräsentation der Ressource hätte. Dies kann dann, wie in Abb. 6 zu sehen, durch eine HTTP-Weiterleitung geschehen.⁴¹

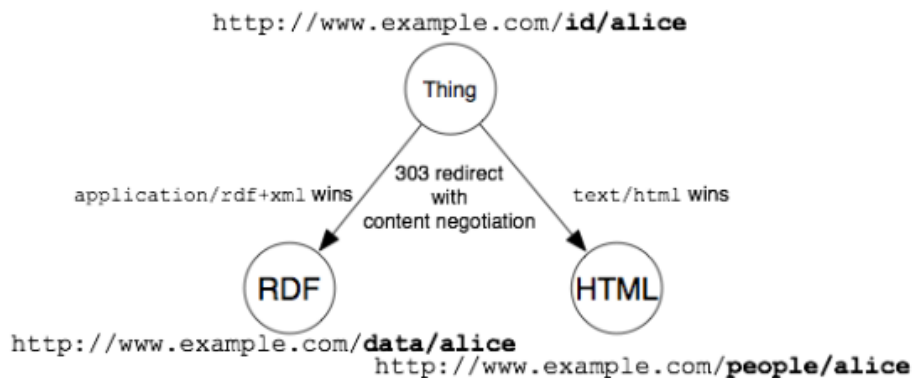


Abb. 6

Content-Negotiation mit einer Weiterleitung⁴²

URIs sollten auch permanent sein und keine Teile enthalten, die sich ändern können, beispielsweise Session-IDs.⁴³ Berners-Lee schlägt generell vor, URIs mit möglichst wenig semantischem Inhalt zu bilden.

⁴⁴ Vor allem, wenn Bezeichnungen sich ändern können, sind generische Identifikatoren zu bevorzugen. Statt `http://example/people/Barack_Obama_1991` sollte eher `http://example/people/2347` verwendet werden.

Théreaux empfiehlt zudem, die URIs kurz zu gestalten und eine Richtlinie für ihre Beschriftung festzulegen (alles klein oder erster Buchstabe gross).⁴⁵ Eine solche Richtlinie bietet zum Beispiel das *UK Cabinet Office* an.⁴⁶ Darin wird empfohlen, Kleinschreibung und Einzahl zu verwenden sowie Wörter durch Trennstriche zu verbinden.⁴⁷

Sehr radikal schlägt Berners-Lee vor, URIs vollkommen von Klassen und Themen zu lösen.⁴⁸ Diesen Ansatz verfolgt zum Beispiel Wikidata mit URIs wie `https://www.wikidata.org/wiki/Q76` für Barak Obama und der fast identischen URI `https://www.wikidata.org/wiki/Q2` für die Erde. Die Loslösung von Klassen und natürlicher Sprache ist einerseits sehr elegant und macht die URIs sehr flexibel, andererseits ist sie für Menschen unübersichtlich und erschwert dadurch das Schreiben von SPARQL-Anfragen und das Verständnis der Datenstruktur.

Unabhängig von der Gestaltung der URI hält die W3C fest, dass Eigenschaften nicht aus der URI einer Ressource abgeleitet werden sollen.

⁴⁹ Ein Identifikator oder der Name einer Person sollten deswegen, selbst wenn sie in der URI unverändert vorkommen, nochmals in der Ressource aufgeführt werden.

Verlinkung, Normalisierung und Anreicherungen

Damit Linked Open Data seine vollen Stärken ausspielen kann, muss ein Datensatz möglichst gut verlinkt sein. Im Gegensatz zu relationalen Datenbanken liegt dabei der Fokus nicht nur auf der inneren Verlinkung, sondern auch auf derjenigen zu externen Ressourcen. Dabei gilt es vor allem, die Verwendung von Literalen zu minimieren. Wenn zum Beispiel eine Person mit einer Ressource verlinkt wird, sollte dies nicht nur über

Fliesstext geschehen, sondern über eine weitere Ressource. Noch besser wäre eine Verlinkung zu einer externen, häufig verwendeten Ressource. Im Fall einer Person könnte dies die URI eines Normdatensatzes von VIAF oder der GND sein. Die Verlinkung kann, wie in RDF 7 zu sehen, direkt oder über ein Zwischenressource erfolgen.

RDF 7

Beispiel für die Verlinkung einer Person (Albert Einstein) mit der Normdatenbank VIAF.

1	@prefix dcterms: <http://purl.org/dc/terms/> .
2	
3	<http://example/object/1> dcterms:creator <http://viaf.org/viaf/75121530> .
1	@prefix dcterms: <http://purl.org/dc/terms/> .
2	@prefix owl: <http://www.w3.org/2002/07/owl#> .
3	
4	<http://example/object/1> dcterms:creator <http://example/person/3> .
5	
6	<http://example/person/3> owl:sameAs <http://viaf.org/viaf/75121530> .

Metadaten-Provenienz und Referenzen

Das Wort Provenienz (aus dem lateinischen *provenire*, «herkommen») bezeichnet die Herkunft einer Person oder Sache. In Archiven wird der Begriff vor allem für Informationen über die Organisationen oder die Personen verwendet, die im Rahmen ihrer Tätigkeit das Archivgut hergestellt haben. Diese Angaben stellen oft eine wichtige

Grundlage für die Beurteilung der Authentizität und die Einordnung von Archivalien dar.



Titel Robert Grimm hält eine Rede, vermutlich bei einer 1.-Mai-Kundgebung in Zürich, ca. 1920

Signatur F 5069-Fa-031
Bestand F 5069 comedia Zürich
Periode 1911-1930
Person Grimm, Robert (1881-1985)

Wie bei dem Archivale selbst, stellt sich die Frage nach der Herkunft natürlich auch bei deren Metadaten. Wer hat beispielsweise die Metadaten der Archivale F 5069-Fa- 031 verfasst und woher ist bekannt, dass es sich bei dem Redner um Robert Grimm handelt? Standen diese Informationen auf der Rückseite der Fotografie? Hat der/die katalogisierende Archivar*in Grimm erkannt? Ergaben sich die Informationen aus dem Kontext? Oder handelt es sich vielleicht um eine maschinelle Bilderkennung?

Derartige Informationen wurden bis zu einem gewissen Grad auch schon in klassischen Archivsystemen festgehalten. Mit der Zunahme von Metadaten sowie der Möglichkeit von Fremddatenanreicherungen und systemübergreifenden Abfragen bei Linked Open Data gewinnt die Thematik jedoch zusätzlich an Komplexität und Bedeutung. Leider ist die Darstellung von Informationen zur Provenienz in Linked Data bislang uneinheitlich gelöst.⁵⁰ Im Folgenden sollen deshalb verschiedene Möglichkeiten vorgestellt und daraus eine mögliche Lösung abgeleitet werden.

Eine sehr einfache Möglichkeit bietet die Verwendung von dcterms. Dieses Vokabular beinhaltet einen Satz grundlegender Prädikaten zur Darstellung von Provenienz:

RDF 8

Provenienz-Modellierung mit dcterms

1	@prefix dcterms: <http://purl.org/dc/terms/> .
2	@prefix dc: <http://purl.org/dc/terms/> .
3	@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
4	@prefix ex: /> .
5	
6	ex:1
7	dc:identifier "F 5069-Fa-031"^^xsd:string;
8	dc:creator "Erika Mustermann"^^xsd:string;
9	dc:created "2018-04-15T14:00:00+02:00"^^xsd:dateTime .

Ein mächtigeres, aber auch komplexeres Modell bietet die Provenienz-Ontologie (PROV-O) der W3C Provenance Working Group. Wie in Abb. 7 zu sehen, besteht die Ontologie aus drei Grundklassen:⁵¹ (1) Entitäten sind physische, digitale, konzeptuelle oder andere Dinge mit festen Aspekten. (2) Aktivitäten sind Handlungen, die über einen gewissen Zeitraum hinweg geschehen und mit einer Entität interagieren. (3) Agenten sind Personen oder Dinge, die die Verantwortung für die Aktivitäten tragen.

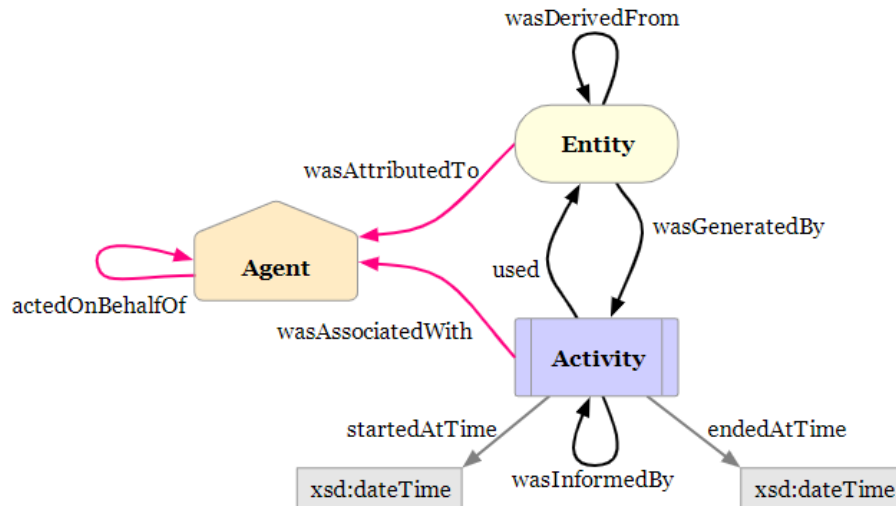


Abb. 7
Grundklassen von PROV-O

RDF 9 Provenienz-Modellierung mit PROV-O

1	@prefix dct: <http://purl.org/dc/terms/> .
2	@prefix dc: <http://purl.org/dc/terms/> .
3	@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
4	@prefix foaf: <http://xmlns.com/foaf/0.1/> .
5	@prefix ex: /> .
6	@prefix prov: <http://www.w3.org/ns/prov#> .
7	@prefix agent: <http://example.com/agent/> .
8	@prefix activity: <http://example.com/activity/> .
9	
10	ex:1
11	a prov:entity;
12	dc:identifier "F 5069-Fa-031"^^xsd:string;
13	prov:wasGeneratedBy activity:1;
14	prov:wasAttributedTo agent:1 .
15	
16	activity:1
17	a prov:activity;
18	prov:startedAtTime
19	"2018-04-15T13:00:00+02:00"^^xsd:dateTime;
20	prov:endedAtTime
21	"2018-04-15T14:00:00+02:00"^^xsd:dateTime;
22	prov:used ex:1;
23	prov:wasAssociatedWith agent:1 .
24	
25	agent:1
26	a prov:Agent, prov:Person;
27	foaf:name "Erika Mustermann"^^xsd:string .

Nicht ideal ist in den beiden aufgezeigten RDF-Beispielen die Vermischung von Metadaten des zu beschreibenden Objekts und Metadaten über die Metadaten. Eckert spricht in diesem Zusammenhang auch von «Metametadaten»⁵². So ist in den Beispielen unklar, ob der «creator» bzw. der «agent» der/die Ersteller*in des Objekts oder von dessen Metadaten ist. Um dieses Problem zu lösen, müssen

Metadaten und Metametadaten voneinander getrennt werden. Ein möglicher Ansatz besteht darin, einen Verweis der Ressource anzufügen:

RDF 10 Provenienz-Modellierung mit PROV-O

1	@prefix dc: <http://purl.org/dc/terms/> .
2	@prefix rdfs: -schema#> .
3	@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
4	@prefix ex: /> .
5	@prefix metadata: <http://example.com/metadata/> .
6	
7	ex:1
8	dc:identifier "F 5069-Fa-031"^^xsd:string;
9	rdfs:seeAlso metadata:1 .
10	
11	metadata:1
12	dc:creator "Erika Mustermann"^^xsd:string;
13	dc:created "2018-04-15T14:00:00+02:00"^^xsd:dateTime .

General haben die bisher besprochen Ansätzen jedoch einen grossen Nachteil: Sie beschreiben stets die Provenienz der gesamten Ressource, einzelne Tripel lassen sich somit nicht getrennt ausweisen. Dies kann problematisch sein, wenn zum Beispiel mehrere Mitarbeiter*innen an der Ressource gearbeitet haben oder wenn einzelne Eigenschaften aus einer Fremddatenanreicherung stammen.

Eine mögliche Lösung für dieses Problem ist die in RDF vorgesehene Reification. Dieser Mechanismus erlaubt durch die Erstellung eines «Statements» unter erneuter Angabe von Subjekt, Prädikat und Objekt Aussagen über einzelne Tripel:

RDF 11

Provenienz-Modellierung mit PROV-O

```

1 @prefix rdf: -schema#> .
2 @prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
3 @prefix dc: <http://purl.org/dc/terms/> .
4 @prefix ex: <http://example.com/id/ex/> .
5
6 ex:1
7   dc:identifler "F 5069-Fa-031"^^xsd:string;
8   dc:title "Robert Grimm hält eine Rede ..."^^xsd:string .
9
10 ex:stmt1
11   rdf:type rdf:Statement;
12   rdf:subject ex:1;
13   rdf:predicate dc:identifler;
14   rdf:object "F 5069-Fa-031"^^xsd:string;
15   dc:creator "Erika Mustermann"^^xsd:string;
16   dc:created "2018-04-15T14:00:00+02:00"^^xsd:dateTime .
17
18 ex:stmt2
19   rdf:type rdf:Statement;
20   rdf:subject ex:1;
21   rdf:predicate dc:title;
22   rdf:object "Robert Grimm hält eine Rede ..."^^xsd:string;
23   dc:creator "Erika Mustermann"^^xsd:string;
24   dc:created "2018-06-15T14:00:00+02:00"^^xsd:dateTime .

```

Trotz der allgemeinen Anwendbarkeit der Reification wird diese in der Praxis kaum verwendet. Der Grund dafür liegt in der grossen Anzahl von Tripeln, die durch die Wiederholung entstehen.

Ein weitaus eleganterer Weg steht seit RDF 11 zur Verfügung. Mit Named Graphs lassen sich die entsprechenden Angaben kompakter und ohne Duplizierungen modellieren:

RDF 12

Provenienz-Modellierung mit «Named Graphs»

```

1 @prefix rdf: -schema#> .
2 @prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
3 @prefix dc: <http://purl.org/dc/terms/> .
4 @prefix ex: /> .
5
6 ex:graph1 {
7   ex:1 dc:identifler "F 5069-Fa-031"^^xsd:string .
8
9 ex:graph2 {
10  ex:1 dc:title "Robert Grimm hält eine Rede ..."^^xsd:string .
11
12 ex:graph3 {
13   ex:graph1
14   dc:creator "Erika Mustermann"^^xsd:string;
15   dc:created "2018-04-15T14:00:00+02:00"^^xsd:dateTime .
16
17 ex:graph2
18   dc:creator "Erika Mustermann"^^xsd:string;
19   dc:created "2018-06-15T14:00:00+02:00"^^xsd:dateTime .
20 }

```


Eine weitere offene Frage betrifft die Darstellung von Referenzen, wenn man zum Beispiel angeben will, woher die Geburtsdaten von Robert Grimm stammen. Sehr gut lassen sich solche Informationen in eigenen Ressourcen in der Kombination mit Named Graph und PROV-O darstellen.

RDF 13

Modellierung einer Person mit Referenz-Angaben.

```
1 @prefix prov: <http://www.w3.org/ns/prov#> .
2 @prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
3 @prefix dcterms: <http://purl.org/dc/terms/> .
4 @prefix ex: /> .
5 @prefix foaf: <http://xmlns.com/foaf/0.1/> .
6
7 ex:graph1 {
8   ex:1 foaf:name "Robert Grimm"^^xsd:string .
9
10 ex:graph2 {
11   ex:1 foaf:birthday "1881-04-16"^^xsd:date .
12
13 ex:graph3 {
14   ex:graph1
15   prov:wasGeneratedBy ex:activity1;
16   prov:wasAttributedTo ex:agent1 .
17
18 ex:graph2
19   prov:wasGeneratedBy ex:activity1;
20   prov:wasAttributedTo ex:agent1;
21   prov:wasDerivedFrom: ex: reference1;
22   prov:wasDerivedFrom: ex: reference2.
23
24 ex:reference1
25   dcterms:references: <http://d-nb.info/gnd/11869779X>;
26   prov:wasGeneratedBy: ex:activity1;
27   prov:wasAttributedTo ex:agent1 .
28
29 ex: reference1
30   dcterms:references: <http://www.hls-dhs-
31     dss.ch/textes/d/D4516.php>;
32   prov:wasGeneratedBy: ex:activity1;
33   prov:wasAttributedTo ex:agent1 .
34
35 ex:activity1
36   a prov:activity;
37   prov:startedAtTime "2018-04-
38     15T13:00:00+02:00"^^xsd:dateTime;
39   prov:endedAtTime
40     "2018-04-15T14:00:00+02:00"^^xsd:dateTime;
41   prov:used ex:1;
42   prov:used ex: reference1;
43   prov:used ex: reference2;
44   prov:wasAssociatedWith ex:agent1 . }
45
46 ex:graph4 {
47   ex:agent1
48   a prov:Agent, prov:Person;
49   foaf:name "Erika Mustermann"^^xsd:string . }
```

Zusammenfassend kann festgehalten werden, dass Linked Data zwar keine vordefinierte Lösung anbietet, mit der durch den Autor dieser Arbeit entwickelten Kombination aus Named Graphs und PROV-O

lassen sich aber selbst komplexe Bedürfnisse, wie sie in den Archiven vorhanden sind, erfüllen.

Versionierung

Eine weitere Problemstellung bezüglich der Datenqualität betrifft das Festhalten von Veränderungen an den Metadaten. Führt zum Beispiel ein*e Benutzer*in eine SPARQL-Abfrage über das Durchschnittsalter aller verzeichneten Personen in den Beständen eines Archives durch, so wird diese Abfrage in einem Archiv mit Neuzugängen ein Jahr später kaum dasselbe Resultat liefern. Das Nachvollziehen von Forschungsergebnissen kann dadurch in gewissen Fällen erschwert bis unmöglich gemacht werden. Diese Problematik unterscheidet sich nicht grundlegend von der der klassischen Zitation, sie spitzt sich aber, wie schon im Fall der Provenienz, durch die technischen Möglichkeiten von Linked Open Data weiter zu.

Eine Lösung für dieses Problem stellt das Festhalten sämtlicher Veränderungen (Neuanlegungen, Veränderungen und Löschungen) dar. Jede Änderung muss dabei als eigene Version gespeichert werden. Äquivalent zur Provenienz gibt es bislang keine standardisierte Methode zur Abbildung von Versionen in Linked Open Data. Für die Erfassung eignen sicher aber wiederum sowohl Dublin Core als auch PROV-O. In Dublin Code bietet sich dafür die Eigenschaft «dcterms:isVersionOf» und in PROV-O die Eigenschaft «prov:wasRevisionOf» an. Die jeweilige Versionsnummer kann zum Beispiel der URI angehängt werden. Die Version kann dabei entweder alle Eigenschaften des damaligen Zeitpunkts oder auch nur die Abweichungen beinhalten.⁵³

Aktualität (Synchronität)

Bei Linked-Open-Data-Plattformen, die ihre Daten aus periodischen Exporten beziehen -zum Beispiel aus einem Archivinformationssystem - gilt es, sicherzustellen, dass diese Exporte häufig und fortwährend geschehen. Dass dieses Qualitätsmerkmal durchaus eine Herausforderung darstellen kann, zeigt der in Kapitel 0 vorgestellte UK Archives Hub, dessen Daten letztmals 2013 aktualisiert wurden.⁵⁴ Ebenfalls sollte die Synchronität bei importierten Fremddaten beachtet werden. Werden zum Beispiel aus Performanz-Gründen Daten aus der GND auf den eigenen Server übertragen, sollten diese in nicht allzu langen Zeitintervallen aktualisiert werden.

Dateiformate und Zugänglichkeit

Um die semantischen Tripel lesbar zu machen, sollten diese in einem Linked-Data- Format angeboten werden. Die W3C schlägt hierfür die Dateiformate RDFa, JSON-LD, Turtle, N-Triples und RDF/XML vor.

⁵⁵ Als Zugriffsmöglichkeit empfiehlt sie den direkten Zugriff über die

URI, eine RESTful-API, einen SPARQL-Endpoint und/oder einen File-Download.⁵⁶

Aus Sicht eines einfachen Zugangs für die Benutzenden ist das Anbieten möglichst vieler Zugänge anzustreben.

Linked Open Data in Archiven

Motivation

Seit knapp zehn Jahren beschäftigen sich diverse bibliothekarische Einrichtungen und Verbundskataloge mit Linked Open Data. In der letzten Zeit haben auch immer mehr Archive begonnen, sich mit dem Thema auseinanderzusetzen. Inzwischen verfügen mehrere Institutionen über eigene Portale. So betreiben zum Beispiel die Nationalarchive der Niederlande⁵⁷, der Schweiz⁵⁸, Italiens⁵⁹ und Grossbritanniens⁶⁰ Linked-Open-Data-Anwendungen. Doch nicht nur Nationalarchive, sondern auch Spezialarchive wie das IISH⁶¹ in Amsterdam oder organisationsübergreifende Projekte wie der UK Archives Hub⁶² experimentieren mit der Technologie Linked Data. Weiterhin arbeitet das International Council on Archives (ICA) mit Records in Context (RiC)⁶³ an einem neuen Archivstandard, der auf dem Konzept von Linked Data basiert.

Woher stammt dieses Interesse der sonst eher traditionell arbeitenden Archive an der Technologie? Oder anders gefragt, welche Vorteile versprechen sich die Archive von Linked Data bzw. Linked Open Data? Einige wesentliche Punkte seien hier kurz aufgeführt.

- **Vernetzung:** Linked Open Data gibt den Archiven die Möglichkeit, ihre Bestände nicht mehr als «Silos» zu verstehen, sondern sie durch die Verlinkung mit Datensätzen auf der ganzen Welt in einen weiterführenden Kontext zu setzen.⁶⁴ Ein Beispiel hierfür bieten die Normdatenbanken GND und VIAF der Bibliotheken. Dadurch ergeben sich neue Möglichkeiten der Zusammenarbeit und der Nutzung von Synergien.
- **Digitalisierung:** Durch die fortschreitende Digitalisierung steigt der Prozentsatz digitaler Ablieferungen andauernd. Damit verbunden ist auch ein starker Zuwachs der Metadaten und ihrer Bedeutung. Die Standards der Archive entwickelten sich aber mehrheitlich aus den Erfahrungen der Katalogisierung von Papierakten heraus. Insofern ist es kaum erstaunlich, dass sich die Archivwelt Gedanken über neue und flexiblere Datenmodelle wie RDF macht.
- **Auffindbarkeit:** Für viele Archive spielt die Auffindbarkeit im Web heute eine wichtige Rolle. Sich an Entwicklungen wie dem Semantic Web zu orientieren, ist deswegen nur folgerichtig
- **Maschinenlesbarkeit:** Ebenfalls gewinnt die Maschinenlesbarkeit der Metadaten durch neue Technologien und das Aufkommen

der Digital Humanities an Bedeutung. Technologien wie RDF ermöglichen es Forschern, vollkommen neue Suchanfragen zu stellen und Data Mining zu betreiben.

- Open-Bewegung: In vielen staatlichen Verwaltungen hat die Open-Bewegung zu einem Mentalitätswandel geführt.⁶⁵ Immer mehr Datensätze stehen den Archiven als Open Data zur Verfügung.
- Einfluss verwandter Fachgebiete: Sicher ebenfalls eine Rolle spielt die Tatsache, dass verwandte Bereiche wie Bibliotheken und Museen verstärkt auf Linked Open Data setzen.

Ausblick in die Bibliothekswelt

Entwicklung und heutiger Stand

Schon lange vor den Archiven begannen sich die Bibliotheken mit dem Thema Linked Open Data auseinanderzusetzen. Der Ausgangspunkt hierfür war zunächst die Frage nach den Nutzungsrechten an Katalogdaten. Auslöser der Diskussion war der Draft-Report der Working Group on Future of Bibliographic Control der Library of Congress, der am 13. November 2007 publiziert wurde.⁶⁶ Da dieser Bericht nicht auf die Thematik der Nutzungsrechte einging, formulierte die Open Knowledge Foundation mit Aron Schwartz einen entsprechenden Protest, der von 150 Gruppen und Einzelpersonen unterzeichnet wurde.⁶⁷ Die hervorgerufene Kritik führt dazu, dass Open Data ab 2008 ein Thema der bibliothekarischen Gemeinschaft wurde.⁶⁸

Einen Rückschritt stellte die im November 2008 durch den WordCat-Katalog veröffentlichte Policy for Use and Transfer of WorldCat Records dar.⁶⁹ Sie legte fest, dass an jeden Datensatz des weltweit größten Verbundkatalogs ein Copyright-Vermerk angebracht werden muss und die Nutzung der Daten generell Restriktionen unterliege. Eine längere Diskussion führte jedoch 2012 dazu, dass zumindest alle in worldcat.org als Linked Data eingebetteten Daten unter eine ODC-BY-Lizenz⁷⁰ gestellt wurden. In den folgenden Jahren begannen weltweit immer mehr Bibliotheken, ihre Metadaten als Open Data zu veröffentlichen, und schufen somit eine Grundvoraussetzung für Linked Open Data.

Der erste bibliografische Linked-Data-Dienst entstand im April 2008.⁷¹ Ed Summers, ein Programmierer der Library of Congress, veröffentlichte eine inoffizielle Version der Library of Congress Headings (LCSH) als Linked Data. Da er aber nicht die Rechte an den Daten besass, musste der Service bereits im Dezember desselben Jahres wieder eingestellt werden. Einige Monate später veröffentlichte jedoch die Library of Congress eine neue, verbesserte und rechtlich einwandfreie Version.⁷² Die erste Veröffentlichung von Titeldaten als Linked Data geschah durch die Schwedische Nationalbibliothek mit der Publikation des Verbundkatalogs LIBRIS als Linked Data.⁷³

Das erste Linked-Data-Projekt im deutschsprachigen Raum initiierte die Zentralbibliothek für Wirtschaftswissenschaften (ZWB) in Leipzig gegen 2009, indem sie ihren Standard Thesaurus Wirtschaft (STW) als Linked Data veröffentlichte.⁷⁴ Schnell folgten weitere wichtige Projekte. So veröffentlichte 2010 das Hochschulbibliothekszentrum des Bundeslandes Nordrhein-Westfalen (hbz) den Linked-Open-Data-Dienst lobid.org.⁷⁵ Ebenfalls in diesem Jahr startete die Deutsche Nationalbibliothek einen Linked-Data-Service für Normdaten.⁷⁶

In der Schweiz betreibt der Metakatalog Swissbib mit linked.swissbib.ch seit 2017 eine umfangreiche Linked-Open-Data-Plattform. Der Dienst basiert auf circa 21 Millionen MARC-XML-Dateien aller Schweizer Hochschulbibliotheken, der Schweizerischen Nationalbibliothek, zahlreicher Kantonsbibliotheken und weiterer Institutionen.⁷⁷ Wie in Abb. 8 zu sehen führen bzw. führten neben SwissBib auch noch weitere grosse Bibliotheken und Verbünde in der Schweiz Linked-Open-Data-Projekte.



Abb. 8

Linked-Open-Data-Projekte in der Schweizer Bibliothekslandschaft (Stand 2013)⁷⁸

Für die Realisierung wurden mehrere bibliotheksspezifische Vokabulare, Ontologien und Datenmodelle erstellt, zum Beispiel die «Bibliographic Ontology (bibo)», das «BIBFRAME vocabulary (bibframe)», die «RDA Registry (rdac, rdau, ...)», «Functional Requirements for Bibliographic Records (FRBR)» und die GND «Ontology (gnd)»⁷⁹. Viele Bibliothekssysteme wie Ex Libris Alma oder Koha beinhalten oder planen eine Linked-Data-Unterstützung.⁸⁰

Ebenfalls zeugen zahlreiche Konferenzen und Arbeitsgruppen von einem regen Austausch in der Bibliothekswelt über das Thema Linked Open Data. So beinhaltet die Open Knowledge Foundation (OKFN) eine Arbeitsgruppe zum Thema Open Bibliographic Data und die International Federation of Library Associations and Institutions (IFLA) eine Gruppe mit dem Namen Semantic Web Special Interest Group. Im deutschsprachigen Raum kann die Arbeitsgruppe Kompetenzzentrum Interoperable Metadaten (KIM) genannt werden.

Synergien und Abweichungen

Ausgehend von der Situation, dass die Archive mit ihren Linked-Open-Data-Bemühungen noch relativ am Anfang stehen und die Bibliotheken bereits über einen reichen Erfahrungsschatz verfügen, ist es sinnvoll, dass die Archive sich an den Erkenntnissen der Bibliotheken orientieren. Aus Sicht des Autors dieser Arbeit sollten dabei vor allem die Datenmodelle, die Migrationsworkflows und die Anwendungsfälle im Zentrum stehen. Ebenfalls sollte versucht werden, vorhandene Ressourcen wie die Normdatenbanken VIAF oder GND für die eigenen Zwecke zu verwenden. Gute Beispiele für einen solchen Erfahrungsaustausch stellen die Konferenzen des Netzwerks Linked Open Data in Libraries, Archives and Museums (LODLAM) dar. Dieser Austausch kann aber auch auf lokaler Ebene und zwischen einzelnen Institutionen oder in gemeinsamen kleineren Workshops realisiert werden.

Gleichzeitig darf nicht vergessen werden, dass es durchaus auch grössere Unterschiede gibt, die bei der Übernahme von Erfahrungen berücksichtigt werden müssen. So verfügen Archive oft über wesentlich mehr Inhalte mit Sperrfristen. Im Gegensatz zu vielen Bibliotheken führen die meisten Archive eigene Kataloge und sind nicht Teil eines gemeinsamen, institutsübergreifenden Verbundkatalogs. Dadurch sind die Datenstrukturen oft inhomogener und schwieriger zu verlinken. Dank gemeinsamer Portalen wie dem Archivportal Europa (APE), Standards wie ISAD(G) und dem Austauschformat EAD konnten hier bereits wesentliche Fortschritte erzielt werden.

Bei der Verwendung von normierten Begriffen für Personen, Geographika und Schlagworten verfügen die Bibliotheken über eine grosse Erfahrung. Über Jahrzehnte hinweg haben sie Institutionsübergreifende Normdaten wie die GND oder VIAF aufgebaut und mit ihren Medien verlinkt. Dank der fortschreitenden Öffnung⁸¹ dieser Normdaten steht den Archiven damit ein umfangreicher Datensatz zur Verlinkung zur Verfügung. Oft ist aber die Verwendung von Normdaten für Archivalien schwieriger als für Bibliotheken. Bei geographischen Normdaten spielt die Historisierung eine grössere Rolle und bei Personen stellen sich noch stärker die Fragen des Datenschutzes.

Auch konzeptionell stehen die Archive vor grösseren Herausforderungen als die Bibliotheken. Viele Archive haben ihre Bestände bisher als isolierte Einheiten verstanden und diese gemäss ISAD(G) als hierarchisch gegliederte Bäume modelliert. Querverlinke zu anderen Ressourcen (Bestände, Normdaten usw.) waren eher selten und wurden im besten Fall mit URLs verlinkt. Bibliotheken hingegen verwenden schon seit langem Graphen-basierte oder Graphen-ähnliche Modelle. Für sie ist das systematische Verlinken von Ressourcen wie Werk, geographischen Angaben, Schlagworten und Personennormdaten keine Neuheit.

Wichtige Datenmodelle im Kulturerbe-Bereich

Im Folgend sollen zwei weitere wichtige Datenmodelle aus dem Kulturerbe-Bereich vorgestellt werden.

Das CIDOC-CRM wurde durch das International Council of Museums (I- COM) entwickelt und als ISO Norm 1127:2014 zertifiziert. Es hat das Ziel, ein semantisches Framework zu schaffen, das von jeder Institution im Kulturerbe verwendet werden kann. Das Modell umfasst 26 Klassen, die mit zahlreichen Eigenschaften und Relationen versehen werden können. Daneben existieren eine Reihe von Erweiterungen, zum Beispiel CRMgeo für geographische Angaben.

Ein weiteres wichtiges Modell ist das Europeana Data Model (EDM).

⁸² Das Datenmodell besteht aus eigenen Ontologie mit Elementen aus bekannten Ontologien wie dc oder dcterms. Das Modell ist tendenziell einfacher gehalten und enthält nur acht Klassen. ⁸³

Von ISAD(G) und ISAAR(CPF) zu Open Data

ISAD(G), ISAAR(CPF), ISDF und ISDLAH

Will man die Metadaten von Archiven verlinken, so stellt sich die Frage nach deren Struktur. Bei vielen Archiven ist diese von den weitverbreiteten Archivstandards I- SAD(G) und ISAAR(CPF) beeinflusst. Diese sollen deswegen im Folgenden kurz vorgestellt werden.

Der International Standard Archival Description (General) oder kurz I- SAD(G) ist ein Standard zur Verzeichnung von Archivgut des International Council on Archives (ICA). Die Veröffentlichung erfolgte 1993/94. Die heute gültige zweite Fassung folgte im Jahr 2000. ISAD(G) sieht eine Verzeichnung in einem mehrstufigen hierarchischen Baum vor. Zu den wichtigsten Stufen gehören Bestand, Serie, Dossier und Dokument. Archivalien werden gemäss dem Provenienzprinzip nach ihrer Herkunft zusammengefasst.

ISAD(G) verfügt, wie in Tabelle 2 auf S. 345 zu sehen, über 26 Verzeich- nungselemente wie die Signatur oder den Entstehungszeitraum. ⁸⁴ Zur Vermeidung von Redundanz gilt die Regel, dass Informationen, die auf einer höheren Stufe angesiedelt sind, nicht auf einer unteren wiederholt werden dürfen. ⁸⁵

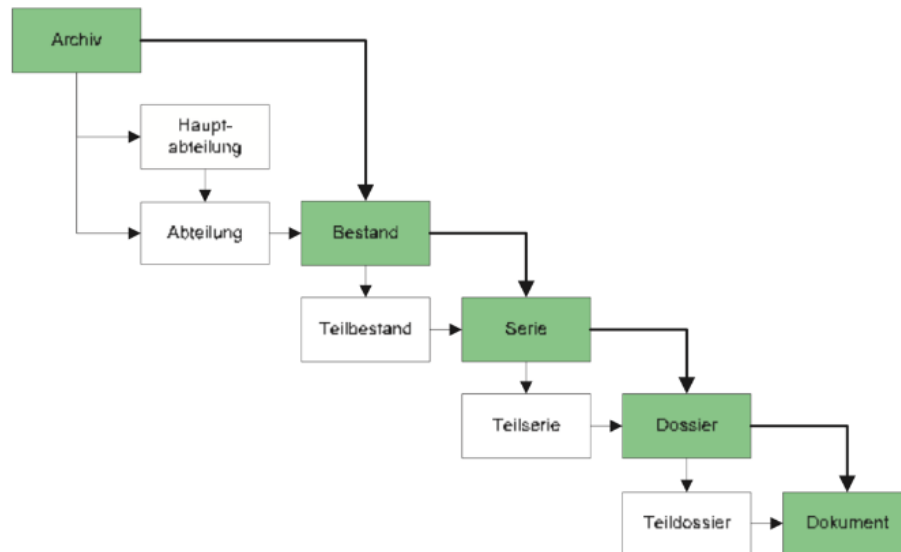


Abb. 9

Verzeichnisstufen nach den schweizerischen Richtlinien für die Umsetzung von ISAD(G).⁸⁶

Es gilt zu beachten, dass die meisten Archive ISAD(G) nicht eins zu eins umgesetzt haben. Viele benutzen weit mehr als die 26 Verzeichnungselemente. So hat beispielsweise das Staatsarchiv des Kantons Zürich das Verzeichnungselement Umfang in drei Felder aufgeteilt: Laufmeter, Gigabyte und Bestelleinheiten.⁸⁷

Der International Standard Archival Authority Record for Corporate Bodies, Persons, and Families oder kurz ISAAR(CPF) ist ein Standard zur Anlegung von Normdaten für Körperschaften, Personen und Familien. Die erste Veröffentlichung erfolgte im Jahr 1996 ebenfalls durch die ICA. Die heute gültige zweite Fassung folgte im Jahr 2004. Im Gegensatz zu ISAD(G) wird ISAAR(CPF) in den Schweizer Archiven kaum eingesetzt.⁸⁸ Anders als bei den Bibliotheken, existieren keine nationalen Gremien, die die Normdaten pflegen bzw. die Ansetzung von Körperschaftsnamen reglementieren.⁸⁹

Neben diesen beiden existieren seitens der ICA noch zwei weitere Normen: die International Standard for Describing Functions (ISDF) für die Beschreibung von Funktionen und die International Standard for Describing Institutions with Archival Holdings (ISDIAH) zur Wiedergabe von Archiv-Institutionen. Da diese aber ebenfalls weit weniger im Einsatz als ISAD(G), sollen sie im Folgenden nicht weiter diskutiert werden.

The LOCAH RDF Vocabulary & Ontologie L'ontologia della descrizione archivistica (OAD)

Ausgehend von ISAD(G) stellt sich die Frage, wie diese Archivbäume nach Linked Data übertragen werden sollen. Eine naheliegende Lösung besteht darin, die Struktur möglichst direkt in ein Linked-Data-Vokabular zu übertragen, was aufgrund der hohen Flexibilität

der Technologie problemlos möglich ist. Diesen Ansatz verfolgte die britische gemeinnützige Organisation Joint Information Systems Committee (JISC). Für das Projekt Linked Open Copac and Archives Hub (LOCAH) entwarf sie von 2010 bis 2011 das LOCAH RDF Vocabulary, das aus 12 Klassen und 61 Eigenschaften besteht. Auf der Website existiert sogar eine XSLT-Datei für die automatische Umwandlung von EAD zu RDF.⁹⁰ Das Vokabular kann somit als Versuch einer möglichst unveränderten Übertragung von ISAD(G) auf Linked Open Data gesehen werden.

Der gleichen Idee folgt auch die 2013⁹¹ entstandene Ontologie L'ontologia della descrizione archivistica (OAD)⁹². Die Ersteller sind das Archivio Centrale dello Stato und des Istituto per i beni artistici, culturali e naturali della Regione Emilia- Romagna (IBC) sowie die Firma regista.exe.⁹³ Die Beschreibungen der Attribute verweisen sogar auf die Kapitelnummern der ISAD(G)-Dokumentation.⁹⁴

Tabelle 2
Zuordnung von ISAD(G) zum LOCAH- und OAD-Vokabular.

Nr.	Verzeichnungselement gemäss ISAD(G)		LOCAH	OAD
	Deutsch95	English96		
1.1	Signatur(en)	Reference code(s)		referenceCode
1.2	Titel	Title	title	title
1.3	Entstehungszeit- raum/ Laufzeit	Dates of creation	dateCreatedAccumulatedString	date
1.4	Verzeichnungsstufe	Level of description	level	levelOfDescription
1.5	Umfang (Menge oder Abmessung)	Extent and medium of the unit	extend	extentAndMedium
2.1	Name der Provenienzstelle	Name of creator	origination	has_nameOfCreator
2.2	Verwaltungsgeschichte/ Biographische Angaben	Administrative/Biographical history	biographicalHistory	has_administrativeBiographicalHistory
2.3	Bestandsgeschichte	Archival history	custodialHistory	archivalHistory
2.4	Abgebende Stelle	Immediate source of acquisition	acquisitions	immediateSourceOfAcquisitionOrTransfer
3.1	Form und Inhalt	Scope and content	scopecontent	scopeAndContent
3.2	Bewertung und Kas- sation	Appraisal, destruction and scheduling	appraisal	appraisalDestructionAndSchedulingInformation
3.3	Neuzugänge	Accruals	accruals	accruals
3.4	Ordnung und Klassi- fikation	System of arrangement		systemOfArrangement
4.1	Zugangsbestimmungen	Conditions governing access	accessRestrictions	conditionsGoverningAccess
4.2	Reproduktionsbestimmungen	Conditions governing reproduction	useRestrictions	conditionsGoverningReproduction
4.3	Sprache/Schrift	Language/scripts of material		has_languageScriptsOfMaterial
4.4	Physische Beschaffenheit und technische Anforderungen	Physical characteristics	physicalTechnicalRequirements	physicalCharacteristicsAndTechnicalRequirements
4.5	Findhilfsmittel	Finding aids	findingAid	has_findingAid
5.1	Aufbewahrungsort der Originale	Existence and location of originals	locationOfOriginals	existenceAndLocationOfOriginals
5.2	Kopien bzw. Reproduktionen	Existence and location of copies	location	existenceAndLocationOfCopies
5.3	Verwandte Verzeichnungseinheiten	Related units of description	relatedMaterial	has_relatedUnitsOfDescription
5.4	Veröffentlichungen	Publication note		has_publicationNote
6.1	Anmerkungen	Note	note	note
7.1	Informationen des Bearbeiters	Archivist's note	isAdministeredBy	archivistsNote
7.2	Verzeichnisgrundsätze	rules or conventions		
7.3	Datum oder Zeitraum der Verzeichnung	date of descriptions		

Vorund Nachteile einer ISAD(G)/ISAAR(CPF)-nahen Modellierung

Der Vorteil dieser Eins-zu-eins-Modellierungen liegt vor allem in ihrer Einfachheit. Zudem weisen sie auch eine gewisse Universalität auf, da sie für alle ISAD(G)-basierten Archive einfach umzusetzen sind.

Daneben ergeben sich aus dem Vorgehen aber auch diverse Nachteile. Zunächst löst der Ansatz nicht das Problem, dass viele Archive weitaus mehr als die 26 Verzeichniselemente verwenden. Damit motiviert das Vorgehen dazu, die Lücke durch eigene Vokabulare zu ergänzen und somit zu einer Inhomogenität bei den Modellierungen beizutragen. Daneben widerspricht der Ansatz auch dem Linked-Open-Data-Grundsatz, bestehende Vokabulare und Ontologien zu verwenden. Zudem mindert ein archivbezogener Ansatz die Interoperabilität zu anderen Datenanbietern wie Bibliotheken oder Museen.

Weiterhin verleitet eine allzu nahe Datenübernahme dazu, die Tatsache auszublenden, dass ISAD(G) und ISAAR(CPF) nicht im Hinblick auf Maschinen-Lesbarkeit und die Verknüpfung von Ressourcen erstellt wurden. Ein Beleg hierfür sind die in den Standards aufgeführten Beispiele für die Verzeichniselemente. Für das Element Sprache/Schrift werden unter anderem «In Dakota, with partial English translation (File)», «English (File)» und «Latin. Ecriture insulaire (noter en particulier l'abréviation utilisée pour per) (Item)» aufgeführt. Für Maschinen ist diese Auswahl schwer zu verstehen und zu vergleichen. Aus Linked-Data-Sicht wäre zudem eine Verlinkung zu einer Ressource wie lexvo einem Literal vorzuziehen.

In Linked Data ist auch keine Vererbung von Inhalten an Kind-Elemente vorgesehen. Eine bessere Lösung wäre hier die mehrfache Verlinkung zu einer separaten Ressource. Konzeptionell unterscheidet sich ISAD(G) von Linked Data dadurch, dass es Bestände lediglich horizontal und isoliert als gewurzelte Bäume modelliert. Verlinkungen zu anderen Ästen des Baumes oder zu gemeinsamen Ressourcen sind grundsätzlich nicht vorgesehen.⁹⁷

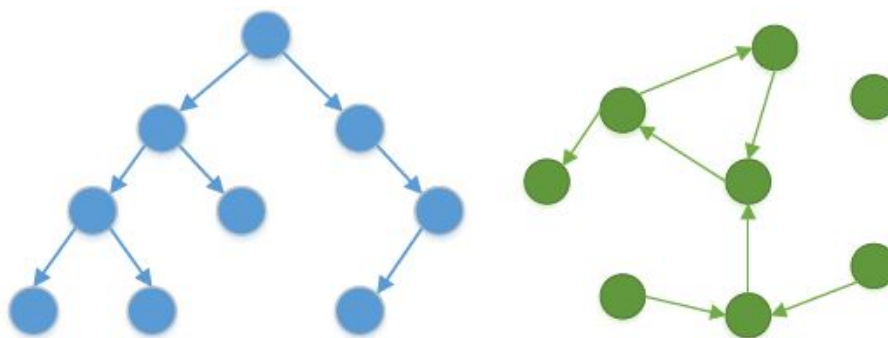


Abb. 10

Ein gewurzelter Baum (blau) und ein frei verknüpfbarer Graph (grün)

Ebenfalls besteht die Gefahr, bei einer direkten Datenmigration einen Linked-Data-Datensatz ohne externe Verlinkungen zu kreieren. Vor allem die Praxis, bei ISAAR(CPF) Personen und Körperschaften als

internen Normdateneinträge ohne Verlinkungen zu GND oder VIAF zu führen, ist aus Linked-Data-Sicht nicht optimal. Zusammenfassend kann festgehalten werden, dass Metadaten, die nach ISAD(G) und ISAAR(CPF) erfasst wurden, durchaus nach Linked Data transformiert werden können. Das Ausschöpfen der vollen Möglichkeiten von Linked Data ist jedoch nur mit einer angepassten Datenstruktur möglich.

Records in Context (RiC)

RiC-CM und RiC-Ontology

Seit 2012 entwickelt die Expert Group on Archival Description (EGAD) der ICA einen neuen Archivstandard. Im September 2016 veröffentlichte sie dazu einen konsultativen Entwurf.⁹⁸ Der Standard baut nach eigenen Angaben auf den vier bestehenden ICA-Beschreibungsstandards ISAD(G), ISAAR(CPF), ISDF und ISDIAH auf.⁹⁹ Anders als die bisherigen Standards ist das darin enthaltene Konzeptmodell RiC-CM graphenbasiert¹⁰⁰ und explizit für die Verwendung von Linked Open Data konzipiert¹⁰¹. Das Modell besteht primär aus 14 Entitäten¹⁰² (Entities):

- | | |
|--------------------|-----------------|
| - Record | - Function |
| - Record-Component | (Abstract) |
| - Record-Set | - Activity |
| - Agent | - Mandate |
| - Occupation | - Documentary- |
| - Position | Form |
| - Function | - Date |
| | - Place |
| | - Concept/Thing |

Diese Entitäten besitzen wiederum vorgegebene Eigenschaften (Properties) und können durch vorgegebene Relationen (Relations) miteinander verbunden werden.

Mit Records in Context Ontology (RiC-O) plant die ICA eine offizielle Ontologie des Konzeptmodells. Auch wenn sich die Ontologie noch in der Entwicklung befindet, wurden bereits zahlreiche Informationen publiziert: Ein Namensraum wurde in PURL eingefügt und lässt sich unter <http://purl.org/ica/ric> aufrufen. Eine Dokumentation ist unter <http://skos.um.es/TR/ric/einsehbar>. Als Präfix ist «ric» vorgesehen. Die Umsetzung wurde, wie bei Linked-Open-Data-Ontologien üblich, mit OWL realisiert. Die Ontologie bildet das gesamte Konzeptmodell ab und umfasst 14 Klassen, 792 Relationen und 166 Eigenschaften.¹⁰³

Wie nun der weitere Zeitplan der Entwicklung genau aussieht, geht aus den öffentlichen Informationen nicht hervor. Alain Dubois rechnet aber damit, dass eine Beta-Version im Laufe des Jahres 2018 veröffentlicht wird.¹⁰⁴

Matterhorn-RDF-Datenmodell

Neben der offiziellen Ontologie der ICA ist mit Matterhorn RDF noch eine alternative Implementierung von RiC im Entstehen begriffen, die durch eCH entwickelt wird. Der in der Schweiz beheimatete Verein hat bereits diverse andere Standards im Bereich E-Gouvernement veröffentlicht. Beteiligt an der Entwicklung des Datenmodells sind unter anderem die Firma Docuteam und die Staatsarchive Wallis, Basel-Stadt und St. Gallen.¹⁰⁵

Den Grund für Erstellung einer Implementierung sieht die Gruppe in dem Umstand, dass RiC-O lediglich 60 % der benötigten Felder für die OAIS-Implementierung der Docuteam Feeder abdeckt.¹⁰⁶ Ebenfalls führen sie den W3C-Grundsatz ins Feld: «It is best practice to use or extend an existing vocabulary before creating a new vocabulary.»

¹⁰⁷ Denn im Gegensatz zu RiC-O ist Matterhorn RDF keine wirkliche Ontologie; es ist vielmehr ein Konglomerat von 17 bereits bestehenden Vokabularen und Ontologien.¹⁰⁸ Dazu gehören unter anderem die bereits vorgestellten dc, dcterms und prov.

Für die geographische und temporale Verortung benutzt das Datenmodell das im Kultgüterbereich häufig verwendete CIDOC Conceptual Reference Model (CRM). Wie in Abb. 10 zu sehen ist, definiert das Datenmodell einen Rahmen aus Klassen. Ähnlich wie in PROV-O und teilweise darauf basierend, gibt es eine Entity, einen Agent und eine Activity. Für die Entity gibt es ein ISAD(G)-, für die Activity ein ISDF- und für die den Agent ein ISAAR-Mapping. Die Mappings geschehen mehrheitlich mit Attributen aus RDA-Vokabularen. Zusätzlich wird durch RDA eine Verbindung zur Welt der Bibliotheken geschaffen. Hinzu kommen zwei Klassen für die Rechteverwaltung. Digitale Dateien werden unter anderem mit Hilfe des Premis- und Pronom-Vokabulars modelliert.

Aufgrund der Mappings und der Breite der RDA-Vokabularen erreicht das Datenmodell für die meisten Archive eine sehr gute Abdeckung. Interessante Möglichkeiten über den Archivbereich hinaus bietet die Einbindung von RDA und CIDOC. Im Hinblick auf die aktive Pflege und Bewirtschaftung ist die Idee, sich aus einem Fundus von lang existierenden und intensiv genutzten Vokabularen und Ontologien zu bedienen, ebenfalls lobenswert. Durch die Beteiligung namhafter Staatsarchive und eines bekannten OAIS-Anbieters ist auch eine praxisnahe und baldige Einführung gegeben. Bemängelt werden kann hingegen die fehlende Standardisierung. Hier bleibt zu beobachten, wie sich die RiC-Ontology in Zukunft entwickelt.

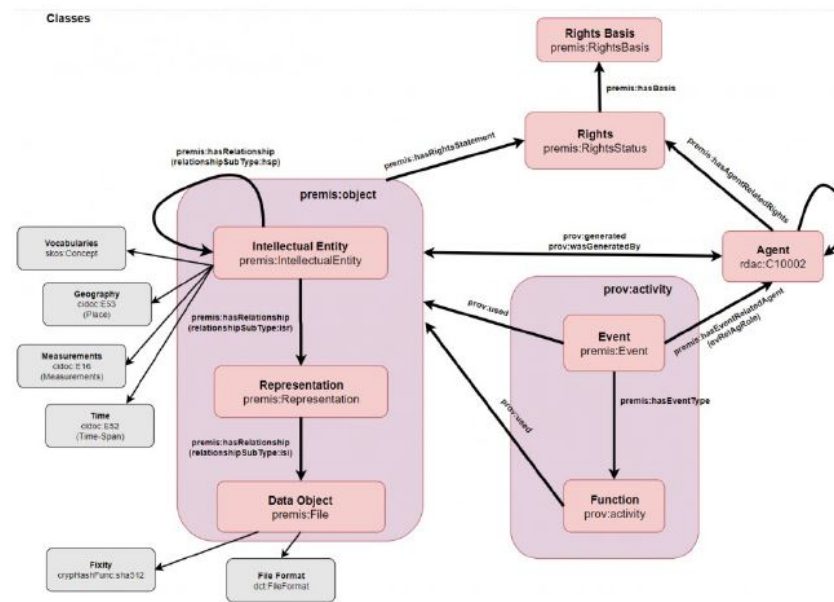


Abb. 11

Das Matterhorn-RDF-Datamodel¹⁰⁹

Existierende Anwendungsfälle

Verbreitung und Herstellerunterstützung

Momentan gibt es noch kaum Linked-Open-Data-Anwendung im Archivbereich. Auch grosse Portale wie Archives Portal Europe (APE) bieten keine Linked-Open-Data-Funktionen. Ein wesentlicher Grund hierfür mag darin liegen, dass die Hersteller der grossen Archivinformationssysteme (AIS) bislang kaum Lösungen für dieses Thema bieten. Dem Autor dieser Arbeit ist kein AIS bekannt, welches als Linked Open Data implementiert wurde. Führende Anbieter in der Schweiz wie Scope Solutions AG oder die CM Informatik AG haben auf ihren Websites noch nicht einmal Informationen zu Linked Open Data.

Bei den wenigen bisher existierenden Anwendungen handelt es sich zumeist um Daten-Exporte aus klassischen AIS-Lösungen, die in Linked-Open-Data-Form umgewandelt werden. Im Folgenden sollen vier solche Lösungen kurz vorgestellt werden.

UK Archives Hub

Der UK Archives Hub ist eine Plattform der britischen gemeinnützigen Organisation Joint Information Systems Committee (JISC). Er beinhaltet Linked Open Data aus 250 Institutionen.

Im Jahr 2010 wurde das einjährige Projekt Linked Open Copac and Archives Hub (LOCAH) gestartet, welches es sich zum Ziel gesetzt hatte, die Daten von Archives Hub und Copac als Linked Data zur Verfügung

zu stellen.¹¹⁰ Wie im folgenden Beispiel zu sehen ist, wurde hierfür das eigens entwickelte «The LOCAH RDF Vocabulary» verwendet.¹¹¹

RDF 14

Auszug des Eintrags der «British Australian New Zealand Antarctic Research Expedition» im UK Archives Hub

```
1 @prefix archiveshub: <http://data.archiveshub.ac.uk/def/> .
2 @prefix dcterms: <http://purl.org/dc/terms/> .
3 @prefix rdf: -schema#> .
4
5 - banzare>
6 a archiveshub:ArchivalResource ;
7 dcterms:identifier "GB 15 BANZARE" ;
8 dcterms:title "British Australian New Zealand Antarctic Research
9 Expedition"@EN ;
9 rdf:label "British Australian New Zealand Antarctic Research
10 Expedition"@EN ;
10 archiveshub:accessProvidedBy
11 <http://data.archiveshub.ac.uk/id/repository/gb15> ;
11 archiveshub:accessRestrictions "By appointment. Some materials
12 deposited at the Institute are NOT owned by the Institute. In such cases
13 the archivist will advise about any requirements imposed by the owner.
14 These may include seeking permission to read, extended closure, or
15 other specific conditions."@EN ;
12 archiveshub:accruals "Further accessions possible"@EN ;
13 dcterms:date "1929-1934";
14 archiveshub:dateCreatedAccumulatedString "1929-1934" ;
15 archiveshub:dateCreatedAccumulatedEnd
16 "1934"^^<http://www.w3.org/2001/XMLSchema#gYear> ;
16 archiveshub:dateCreatedAccumulatedStart
17 "1929"^^<http://www.w3.org/2001/XMLSchema#gYear> ;
18 archiveshub:extent "Expedition material (3 volumes)" ;
19 archiveshub:scopecontent "The collection comprises of press cuttings
20 relating to the expedition"@EN .
```

Etwas ungewöhnlich sind die identischen Doppel-Ausführungen, zum Beispiel durch dcterms:title und rdf:label. Die Modellierung beinhaltet mit PROV-O einen Nachweis der Provenienz. Insgesamt umfasst das Projekt 1'495'168 Tripel.¹¹²

Nach Projektende wurden die Daten nur noch bis 2013 aktualisiert. Inzwischen funktioniert auch der SPARQL-Endpoint nicht mehr vollständig. Die Anwendung erfuhr somit keine Übernahme in den Regelbetrieb.

LINDAS und Archival Linked Open Data (aLOD)

Die beiden folgenden Projekte LINDAS und alod.ch stellen die ersten beiden grossen Linked-Open-Data-Bemühungen in der schweizerischen Archivlandschaft dar. Aufgrund der E-Government-Strategie Schweiz 2012–2015 entwickelte das Staatssekretariat für Wirtschaft (SECO) den Linked Data Service LINDAS.¹¹³ Der seit 2017 durch das betriebene Dienst ermöglicht es dem Bund, den Kantonen und den Gemeinden, strukturierte Daten den Nutzern organisationsübergreifend

zur Verfügung zu stellen.¹¹⁴ Der Zugriff erfolgt über einen öffentlichen SPARQL-Endpoint.¹¹⁵

Seit der Inbetriebnahme wurden mehrere Anwendungsfälle erstellt. Dazu gehören unter anderem die historischen Bundesbudgets und das historisierte Gemeindeverzeichnis.¹¹⁶ Wie im folgenden Beispiel zu sehen ist, wird primär eine Ontologie mit dem Namen «Governmental Ontology Switzerland» verwendet. Ergänzt wird diese durch häufig verwendete Ontologien wie dcterms, xsd und rdf.

RDF 15

Eintrag des Kantons Aargau in LINDAS

```
1 @prefix gont: <https://gont.ch/> .
2 @prefix dcterms: <http://purl.org/dc/terms/> .
3 @prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
4 @prefix rdf: -rdf-syntax-ns#> .
5
6
7 rdf:type gont:Canton;
8 dcterms:identifier "19"^^xsd:integer ;
9 gont:longName "Aargau";
10 gont:id "19"^^xsd:integer ;
11 gont:date "1960-01-01"^^xsd:date ;
12 gont:cantonAbbreviation "AG" .
```

Eine allgemeine Kennzeichnung der Provenienz und eine Versionierung, wie sie in Kapitel 3 vorgestellt wird, sind nicht in Lindas modelliert. Lediglich einzelne Typen verfügen über individuelle Kennzeichnungen oder Versionierungen. So hat zum Beispiel die Gemeinde Marly aufgrund einer Gemeindefusion drei Versions-Einträge.

¹¹⁷ Mit Eigenschaften wie «gont:abolitionEvent» werden zudem Veränderungen an den Einträgen dokumentiert.

Im Jahr 2014 entstand aus einer informellen Zusammenarbeit des Schweizerischen Bundesarchives, der Kantonsarchive Neuenburg, Wallis, Genf, Basel-Stadt und dem Stadtarchiv Baden das Projekt aLOD. Das Ziel des Projekts war es, eine Pilotinfrastruktur mit einem visuellen Ergebnis zu erstellen. Die verwendeten Daten bestehen aus Teildatensätzen aus allen Archiven.¹¹⁸

Wie im folgenden Beispiel zu sehen ist, verwendet aLOD vor allem das LOCAH-Vokabular. Ergänzt wird dieses noch durch eine eigene Ontologie sowie bekannte Vokabulare wie dcterms.

RDF 16

Eintrag eines Archivals aus dem Staatsarchiv Basel-Stadt in aLOD

1	@prefix archiveshub: <http://data.archiveshub.ac.uk/def/> .
2	@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
3	@prefix time: <http://www.w3.org/2006/time#> .
4	@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
5	@prefix dcterms: <http://purl.org/dc/terms/> .
6	@prefix dc: <http://purl.org/dc/terms/> .
7	
8	-bs.ch/id/archivalresource/CH-000027-1/pa-633c-b-2-fasc-21>
9	a archiveshub:ArchivalResource ;
10	archiveshub:level ;
11	time:intervalStarts "1768-12-";
12	time:intervalEnds "1768-12-";
13	archiveshub:maintenanceAgency -000027-1> ;
14	archiveshub:maintenanceAgencyCode "CH-000027-1";
15	<http://data.alod.ch/alod/referenceCode> "PA 633c B 2, fasc. 21";
16	<http://data.alod.ch/alod/databaseID> 588691 ;
17	<http://data.alod.ch/alod/createdAt> "2007-09-19"^^xsd:date ;
18	<http://data.alod.ch/alod/changedAt> "2008-10-31"^^xsd:date ;
19	;
20	archiveshub:accessRestrictions .:
21	B76a0eac5246b404d729aaef2218a3a56 ;
22	<http://data.alod.ch/alod/physTech> "uneingeschränkt";
23	archiveshub:isRepresentedBy ;
24	tion/detailliert> ;
25	<http://data.alod.ch/alod/hasFindingAid> false ;
26	<http://data.alod.ch/alod/recordID> "pa-633c-b-2-fasc-21";
27	<http://data.alod.ch/alod/legacyTimeRange> "1768.12.6";
28	dc:title "Diverse Papiere von Albert Ochs, Mme. François His, Louise Ochs-
29	His" ;
30	dc:relation - bs.ch/id/archivalresource/CH-000027-1/pa-633c-b-2> ;
31	dcterms:isPartOf - bs.ch/id/archivalresource/CH-000027-1/pa-633c-b-
32	2> ;
33	skos:hiddenLabel "Diverse Papiere von Albert Ochs, Mme. François His,
34	Louise Ochs-His, Rechnungen, Vorfahren und Verwandte (Pierre Ochs,
35	Albrecht Ochs, Familie de Dietrich), Nachlass Peter Ochs-Vischer (1752-
36	1821), □" .

Europeana

Die Europeana ist ein Webportal, auf welchem Inhalte aus mehreren tausend europä- ischen Archiven, Bibliotheken und Museen gesammelt sind.¹¹⁹ Insgesamt umfasst sie mehr als 50 Millionen Objekte.¹²⁰ Als Datenmodell wird EDM verwendet.¹²¹ Auch einige Bestände aus dem Sozialarchiv sind über die Europeana abrufbar.¹²² Das Mapping geschieht hier eher rudimentär; so werden alle Felder als Zeichenkette übernommen. Eine Normierung oder das Anlegen von Ressourcen wird nicht vorgenommen. Auch ist es teilweise schwer, für alle Felder Äquivalente zu finden, was dazu führt, dass sich in der Europeana nicht alle Metadaten der importierten Objekte befinden.

International Institute of Social History (IISH)

Aufgrund seiner inhaltlichen Nähe soll im Folgenden noch das IISH vorgestellt werden. Wie das Schweizerische Sozialarchiv auch, sammelt das IISH im Bereich der Sozialgeschichte. Ebenfalls bestehen seine Bestände aus zahlreichen, vornehmlich privaten Organisationen wie Gewerkschaften und Vereinen. Teilweise befinden sich in den beiden Institutionen sogar Archive der gleichen Organisationen. So archiviert zum Beispiel das Schweizerische Sozialarchiv die Unterlagen von Greenpeace Schweiz und das IISH von Greenpeace International. Folglich wären Verlinkungen zwischen den Datenbeständen besonders interessant.

Das IISH stellt seine Metadaten in einem Repository online zur Verfügung.¹²³ Wie im unteren Beispiel zu sehen ist, benutzt das IISH die drei häufigen Web-Ontologien rdf, schema und foaf. Zusätzlich verwendet es eine ausführliche eigene Ontologie. Diese beinhaltet Elemente wie die Höhe in Zentimetern (<https://iisg.amster-dam/vocab/heightInCm>).

RDF 17

Eintrag zu einer IISH-Archivale auf datadruid.

1	@prefix rdf: -rdf-syntax-ns#> .
2	@prefix schema: <https://schema.org/> .
3	@prefix vocab: <https://iisg.amsterdam/vocab> .
4	@prefix foaf: <http://xmlns.com/foaf/0.1/> .
5	@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
6	
7	
8	rdf:type <http://purl.org/dc/dcmitype/StillImage> ;
9	rdf:type vocab:Poster ;
10	schema:inLanguage -3/> ;
11	schema:name "Poster Ernesto Che Guevara" ;
12	schema:about
13	<https://iisg.amsterdam/authority/person/96961> ;
14	vocab:dateOfManufacture "1998"^^xsd:gYear ;
15	vocab:dateOfPublication "1998"^^xsd:gYear;
16	vocab:form <https://iisg.amsterdam/authority/form/3248>
17	;
18	vocab:heightInCm "61"^^xsd:integer ;
19	vocab:manufacturer "Pyramid" ;
20	vocab:placeOfManufacture "Leicester" ;
21	vocab:size "86x61 cm" ;
22	vocab:topic
23	<https://iisg.amsterdam/authority/topic/306035> ;
24	vocab:widthInCm "86"^^xsd:integer;
25	foaf:depiction
26	
27	> .
28	
29	vocab:possibleViafMatch <http://viaf.org/viaf/31991306> ;
30	schema:name "Guevara, Ernesto Che" ;
31	rdf:type schema:Person .

Wie im Beispiel zu sehen ist, sind Sprachen mit der Datenquelle lexvo.org verlinkt. Personen und Organisationen sind mit der Eigenschaft possibleViafMatch mit der Normdatenbank VIAF verknüpft.

Weitere

In der Schweiz kann noch ein weiteres Archiv genannt werden, das gerade dabei ist, ein Linked-Open-Data Projekt zu realisieren. Es handelt sich um das aus der Fusion von Tanzarchiv und Theatersammlung hervorgegangene Schweizer Archiv der Darstellenden Künste (SAPA). Zusammen mit der Fachhochschule Bern ist es dabei, das «Data Model for the Swiss Performing Arts Platform» zu entwickeln. Ein Entwurf des Datenmodells steht bereits online, die Anwendung selbst jedoch noch nicht.¹²⁴ Das Datenmodell umfasst eine eigene umfangreiche Ontologie, deren Klassenmodell sich an Functional Requirements for Bibliographic Records (FRBR) und RiC orientiert.¹²⁵ Ebenfalls mit dem Thema beschäftigt sich Memoriav, der Verein zur Erhaltung des audiovisuellen Kulturgutes der Schweiz. Eine Studie über die zukünftige Ausrichtung des Vereins, die ebenfalls durch die Fachhochschule Bern erstellt wurde, spricht Open Data und Linked Data wichtige Funktionen für die Zukunft zu.¹²⁶

Tabelle 3
Vergleich der vorgestellten Linked-Open-Data-Anwendungen.

	UK Archives HUB	aLOD	Europeana	IISH	SEPA
Inhalt	Ca. 250 Archive aus dem UK Archives HUB	Ausgewählte Bestände aus dem Schweizerischen Bundesarchiv, den	Objekte aus diversen Kultur-erbinstitutionen aus ganz Europa	Archivbestand des IISH	Archivbestand des SEPA
		Kantonsarchiven Neuenburg, Wallis, Genf und Basel-Stadt sowie dem Stadtarchiv Baden			
Datenmodell	LOCAH in Kombination mit dcterms, skos, ore, lode, prov-o ...	LOCAH in Kombination mit dcterms, skos ...	EDM in Kombination mit dc, dcterms, foaf, ore, owl, skos ...	Eigenes Vokabular in Kombination mit schema, rdf, foaf ...	Eigenes Vokabular
Inbetriebnahme	2010	2014	2008	k.A.	Noch nicht in Betrieb
Archivstandard des Ursprungsmaterials	ISAD(G)	ISAD(G)	Diverse	EAD/MARC21	ISAD(G)
Enge Verbindung mit AIS	Nein	Nein	Nein	Nein	
Umfang	1'495'168 Tripel	7'706'028 Tripel	58'573'202 Objekte	20'837'889 Tripel	
Aktualisierungen	Keine	Keine	Fortlaufend	Fortlaufend	
Metadaten-Lizenz	CC0	Open Data Swiss: Freie Nutzung	CC0	CC0	
Externe Verlinkungen	dbpedia	-	-	VIAF, lexvo.org	
Metadaten-Provenienz	Ja	Nein	Nein	Nein	
Versionierung	Nein	Nein	Nein	Nein	
Dateiformate	turtle, RDF/XML	json-ld, turtle, n3	JSON	-	
SPARQL-Endpoint	Nicht mehr in Betrieb	Ja	Nein	Ja	
API	Nein	Nein	Ja	Ja	

Arbeitsgruppen und Weiterbildungen

International beschäftigt sich die bereits im Kontext von RiC genannte Expert Group on Archival Description (EGAD) der ICA mit dem Thema Linked Open Data. In der Schweiz ist hier vor allem die Arbeitsgruppe Normen und Standards des Vereins Schweizerischer Archivarinnen und Archivare (VSA) aktiv. Im Jahr 2018 veranstaltete sie einen gutbesuchten Workshop zum Thema «Linked Data zwischen Theorie und praktischer Anwendung». ¹²⁷ Ebenfalls einen wichtigen Beitrag liefern die universitären Weiterbildungskurse. So bot zum Beispiel der

Master of Advanced Studies in Archival, Library and Information Science der Universität Bern und Lausanne in der Periode 2017/18 drei Lektionen zu Linked Open Data an.

Zwischenfazit

Wie steht es nun um Linked Open Data im Archivbereich? Aus Sicht des Autors die- ser Arbeit ist das Bild ambivalent. Auf der einen Seite deutet sich Linked Open Data als klare Zukunftsperspektive in der Archivwelt an. Viele der grossen Archive be- schäftigen sich inzwischen mit dem Thema oder bieten sogar Pilotprojekte an. Ebenfalls kündigt sich mit RiC ein neuer Archivstandard an, der auf den Ideen von Linked Data basiert.

Andererseits muss aber festgestellt werden, dass die ganze Entwicklung eher schleppend voranschreitet. RiC hat nach sechs Jahren Entwicklung noch nicht die Entwurfsphase verlassen. Ebenfalls ist die Zahl an existierenden Linked-Open-Data- Anwendungen im Archivbereich noch sehr überschaubar. Zudem wirken die meistens dieser Projekte noch sehr am Anfang oder haben im Fall von LOCAH gar nie wirklich den Regelbetrieb erreicht. Viele der Projekte erfüllen auch nur einen kleinen Teil der in Kapitel 0 vorgestellten Qualitätsmerkmale. Vor allem bei den externen Verlinkungen und der Metadaten-Provenienz stehen die Anwendungen eher am Anfang. Auf- fällig ist die hohe Diversität bei den Datenmodellen:

- Der UK Archives Hub, aLOD und das Archivio Centrale dello Stato verwenden mit LOCAH und OAD Ontologien, die versuchen, ISAD(G) in Linked Data umzuwandeln.
- Die Europeana und das IISH setzen primär häufig verwendete, existierende Vokabulare und Ontologien wie Dublin Core ein. Fehlende Elemente werden dabei durch umfangreich eigene Vokabulare abgedeckt, die sich losgelöst von den archivalischen Standards bewegen.
- Die ICA versucht mit RiC-O, eine eigene, vollkommen neue und umfassende Ontologie zu etablieren.
- Mit Matterhorn RDF steht zusätzlich eine weitere RiC-Ontologie zur Verfügung. Im Gegensatz zur Ontologie der ICA stellt Matterhorn RDF jedoch eine reine Kombination von bereits existierenden Vokabularen und Ontologien dar.

Doch welcher Weg ist der Beste für die Archive?

Zunächst gilt festzuhalten, dass eine starke Fragmentierung nicht gut wäre. Diese würde übergreifende Datenzugriffe massiv erschweren. Aus Sicht des Autors dieser Arbeit ist das Konzeptmodell von RiC ein Schritt in die richtige Richtung. Es schafft die Grundlagen die archivalischen Metadaten als Linked Data zu konzeptionieren. Dass die ICA eine offizielle Ontologie zu RiC plant, ist ebenfalls zu begrüßen. Allerdings sind die Kritikpunkte, der fehlenden Tiefe der Verzeichniselemente und die Forderung nach der Benutzung bereits bestehenden Ontologien durchaus berechtigt. Der Ansatz von Matterhorn RDF, einen gut

ausgewählten Satz bestehender Ontologien zu verwenden, um die ICA-Standards abzubilden, ist daher sehr interessant. Vor allem der Einbezug von Ontologien aus dem Bibliotheksbereich scheint im Hinblick auf eine mögliche Zusammenarbeit und Verlinkung sinnvoll. Andererseits trägt eine zweite RiC-Implementierung wiederum zu einer Fragmentierung bei. Hier gilt es abzuwarten, wie die Ontologien Verbreitung finden werden. Vor allem auch ob Mat-terhorn RDF sich in einem internationalen Kontext Einsatz findet.

ISAD(G)-nahe Ontologien sind nach Meinung des Autors dieser Arbeit eine Sackgasse beziehungsweise eine Zwischenlösung, da sie ein Modell nach Linked Data abbilden, was nicht in dessen Sinne konzipiert wurde. Für Archive die aber mit wenig Ressourcenaufwand einen ISAD(G)-Datenbestand nach Linked Data transformieren wollen können sie aber durchaus ein pragmatischer Weg sein. Für Modelle mit nur sehr wenig Verzeichniselementen kann durchaus auch der Ansatz des IISH verwendet werden. Aus Sicht des Autors sollte aber auf jeden Fall auf die Schaffung eigener Ontologien verzichtet werden, um eine noch grösseren Fragmentierung entgegenzuwirken.

Zu guter Letzt hängt die Wahl des geeignetsten Datenmodells sicherlich auch von den eigenen Voraussetzungen ab. Dabei sollten folgende Fragen berücksichtigt werden: Welches Modell deckt die eigenen Verzeichniselemente und Workflows am besten ab. Welche Datenmodelle benutzen ähnliche Institutionen, mit denen die eigenen Daten womöglich verlinkt werden. Welche Datenmodelle kennen die Benutzenden und sind für sie einfach anzuwenden.

Fallbeispiel Schweizerisches Sozialarchiv

Aufbauend auf den bisher gemachten Erkenntnissen, soll in diesem Kapitel eine Linked-Open-Data-Modellierung für das Schweizerische Sozialarchiv erstellt werden.

Dabei soll untersucht werden, wie sich die gesammelten Qualitätsmerkmale auf die Praxis übertragen lassen.

Die Vorgehensweise orientiert sich an den in Kapitel 2 vorgestellten Qualitätsmerkmalen. Dafür werden zunächst eine Analyse und eine Abgrenzung der Datenstruktur vorgenommen. Danach werden mögliche Ziele für das Sozialarchiv gesammelt. Weiterhin wird die Lizenzierung der Metadaten betrachtet und es folgt eine Erstellung eines Datenmodells. Im letzten Schritt wird bei einigen ausgewählten Verzeichniselementen die Umwandlung genauer betrachtet, um auf entsprechende Probleme noch besser eingehen zu können.

Grundvoraussetzung

Das *Schweizerische Sozialarchiv* wurde 1906 gegründet.¹²⁸ Sein Sammlungsgebiet beinhaltet den gesellschaftlichen, politischen und kulturellen Wandel vom 19. Jahrhundert bis in die Gegenwart mit

Fokus auf der Schweiz.¹²⁹ Im Zentrum steht dabei die Dokumentation von sozialen Bewegungen. Anders als sein Name vermuten lässt, ist das Schweizerische Sozialarchiv nicht nur ein Archiv, sondern auch eine Bibliothek und eine Dokumentation. Alle drei Bereiche arbeiten eigenständig und führen ihre Metadaten getrennt.

Die Bibliothek ist Teil des schweizweiten Nebis-Verbundes. Die Dokumentation erfasst ihre Datenbestände in einer eigenen Datenbank und ist über <http://www.sachdokumentation.ch> aufrufbar.¹³⁰ Mit Beginn der elektronischen Verzeichnung 1992 des Papierarchivs wurden zunächst alle Bestände in einer Anwendung verzeichnet.¹³¹ Da diese Software für die Integration von Bild-, Ton- und Video-Dokumenten nicht gut geeignet war, entschied man sich 2003 mit Beginn der Sammlung audiovisueller Materialien dafür, diese durch eine getrennte Software zu erfassen. Aus diesem Grund erfolgt auch der Zugriff auf die Archivbestände über zwei verschiedene Websites.¹³² Auch wenn diese Querverweise beinhalten, ist diese Lösung weder aus archivalischer Sicht noch aus Sicht der Benutzer optimal. In Zukunft wird zusätzlich eine Lösung zur Archivierung von digital erstellten Dokumenten hinzukommen.

Das *Schweizerische Sozialarchiv* hat sich bisher noch nicht mit dem Thema Linked Open Data auseinandergesetzt und bietet demzufolge auch keine Dienste in dieser Richtung an.

Schritt I: Analyse und Abgrenzung der Datenstruktur

Im Fallbeispiel werden die Metadaten des Archivs verwendet. Die Metadaten der Bibliothek und der Sachdokumentation sollen lediglich als Verlinkungsquellen betrachtet werden. Der Grund für diese Einschränkung ergibt sich aus dem begrenzten Umfang der Arbeit.

Wie bereits erwähnt, sind die Metadaten des Archivs auf zwei Anwendungen verteilt. Das Papierarchiv ist sehr einfach gehalten: Die Bestände werden in XML-Dateien katalogisiert und anschliessend in EAD-konforme XML-Dateien umgewandelt. Schliesslich werden daraus HTML-Dateien generiert, die als Benutzer-Frontend dienen.



Abb. 12
Daten-Workflow des Papierarchivs

Die Modellierung basiert auf ISAD(G) und umfasst die Stufen Bestand, Serie, Teilsérie und Dossier. Insgesamt kommen 29 Verzeichniselemente zur Anwendung. Während die Stufen Bestand

und Dossier fast alle Elemente verwenden, umfassen die Stufen Serie und Teilserie lediglich die Elemente Titel und Verzeichnungsstufe.¹³³ Insgesamt beinhaltet das Papierarchiv 763 Bestände aus Körperschaftsarchiven und Personennachlässen.

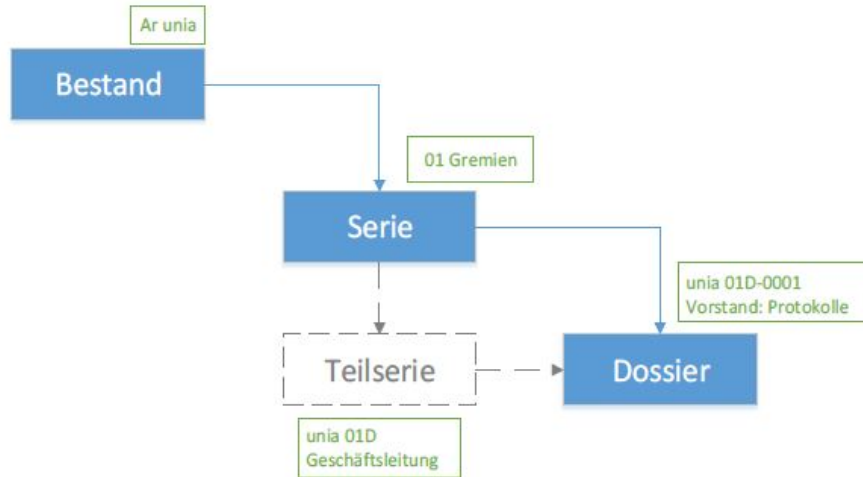


Abb. 13

Struktur des Papierarchivs nach ISAD(G) mit dem Unia-Archiv als Beispiel (grün)

Im Gegensatz zum Papierarchiv liegt dem audiovisuellen Archiv eine SQL-Daten-bank zu Grunde.

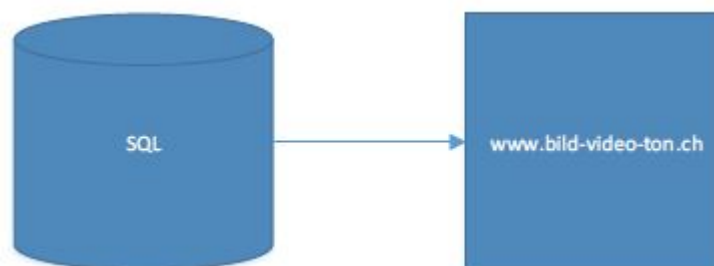


Abb. 14

Daten-Workflow des audio-visuellen Archivs

Die Modellierung orientiert sich ebenfalls an ISAD(G). Die Verzeichniselemente wurden hingegen davon losgelöst nach Bedarf angelegt. Insgesamt existieren 33 Elemente. Während die Stufen Bestand 15 und Dokumente 33 Elemente umfassen, beinhaltet die Stufe Serie nur die Verzeichniselemente Titel und Verzeichnungsstufe. Im Gegensatz zum Papierarchiv weisen Bestand und Dokument mehrheitlich verschiedene Verzeichniselemente auf. Insgesamt enthält das audiovisuelle Archiv 211 Bestände.

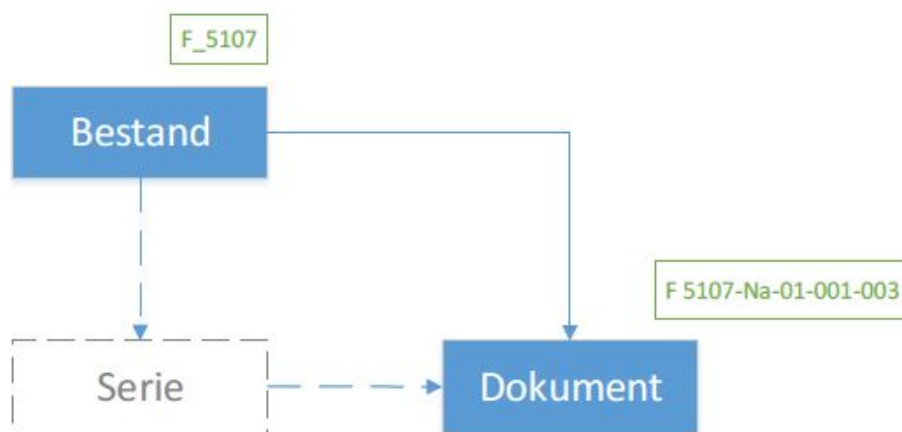


Abb. 15

Struktur des audio-visuellen Archivs nach ISAD(G) mit dem Vogler-Archiv als Beispiel (grün) Serie als optional

Ein Archiv für digital Ablieferungen ist momentan noch in der Entstehung begriffen. Das Schweizerische Sozialarchiv wird dafür die Softwareanwendung Feeder der Firma Docuteam verwenden. Der Feeder wird im Laufe des nächsten Jahres auf das Matterhorn-RDF-Datenmodell umgestellt und wird somit bereits über eine Linked-Data-Modellierung verfügen.

Schritt II: Zielsetzung

Linked Open Data bietet dem Sozialarchiv verschiedene Möglichkeiten:

1. Primär soll die Verlinkung zwischen dem Archiv, der Sachdokumentation und der Bibliothek verbessert werden. Vor allem im Bereich des Archivs soll es möglich sein, die Daten der drei Plattformen (Papierarchiv, Datenbank Bild + Ton und eArchiv) als Einheit zu verbinden.
2. Daneben soll auch die Verlinkung zu externen Datenquellen ermöglicht werden. Der Fokus liegt dabei auf Normdaten wie GND, VIAF und Metagrid. Mithilfe der Verlinkungen sollen Datenanreicherungen möglich werden. Bei Personen wäre beispielsweise die Anreicherung durch alternative Schreibweisen des Namens wünschenswert, um die Suchfunktion zu verbessern. Zusätzlich könnten dem Benutzer die verlinkten externen Ressourcen angezeigt werden, zum Beispiel in Form eines Links zur Eintragung in der GND-Datenbank. Weiterhin wäre auch eine Verlinkung zu verwandten Beständen in anderen Archiven interessant. Wünschenswert wäre auch eine bessere Verlinkung zwischen den Archivalien und den sie betreffenden Forschungsarbeiten.
3. Ein weiteres wichtiges Ziel besteht darin, maschinelle Anfragen zu ermöglichen. Dazu soll ein öffentlicher

SPARQL-Endpoint zur Verfügung stehen. Durch diesen könnten die Forschenden komplexere Suchanfragen stellen, zum Beispiel eine Suche nach Bildern, auf denen nur Frauen zu sehen sind. Auf Basis der maschinellen Anfragen könnten Funktionen geschaffen werden, die beispielsweise Netzwerke von Personen aufzeichnen, welche gemeinsam in den Archivalien vorkommen.

4. Durch eine Vereinheitlichung und eine bessere Dokumentation der Datenfelder soll zudem das Mapping in Portalen wie Europeana oder APE vereinfacht werden.
5. Zu guter Letzt soll mit der Transformation auch die Datenqualität erhöht werden. Inkonsistenzen und Fehler sollen erkannt und getilgt werden.

Schritt III: Lizenzierung der Metadaten

Wie bereits in Kapitel 0 ausgeführt, erfüllen die wenigsten Metadaten die im schweizerischen Recht formulierten Anforderungen an ein Werk und können somit ohne weiteres als Open Data verwendet werden. Um dennoch eine klare Situation zu schaffen, hat das Schweizerische Sozialarchiv alle archivalischen Metadaten als Creative Commons CC-0 freigeben.¹³⁴ Dies war problemlos möglich, da alle Metadaten selbst erstellt wurden. Bei Fremddatenanreicherungen wurde als Kriterium festgelegt nur Quellen zu verwenden, die Open-Data-konform sind.

Schritt IV: Modellbildung

Wie im dritten Kapitel gezeigt gibt es verschiedene Möglichkeiten für die Archive ihre Daten zu Modellieren. Für das Fallbeispiel soll das Datenmodell Matterhorn RDF verwendet werden. Fürs dieses sprechen vor allem drei Gründe:

- Das Archiv für digitale Ablieferungen (Feeder) wird auf diesem Modell basieren
- ISAD(G) lässt sich gut auf Matterhorn RDF abbilden
- Fast alle verwendeten Verzeichniselemente der drei Archivteile lassen sich mit Matterhorn RDF abbilden.

Die Metadaten-Provenienz und Referenzen werden wie in Kapitel 0 erarbeitet und in Abb. 16 ersichtlich mit Named Graphs moduliert. Eine ausführliche Auflistung aller Verzeichniselemente und ihrer Linked-Open-Data-Modellierungen ist im Anhang abS. 393 aufgeführt.

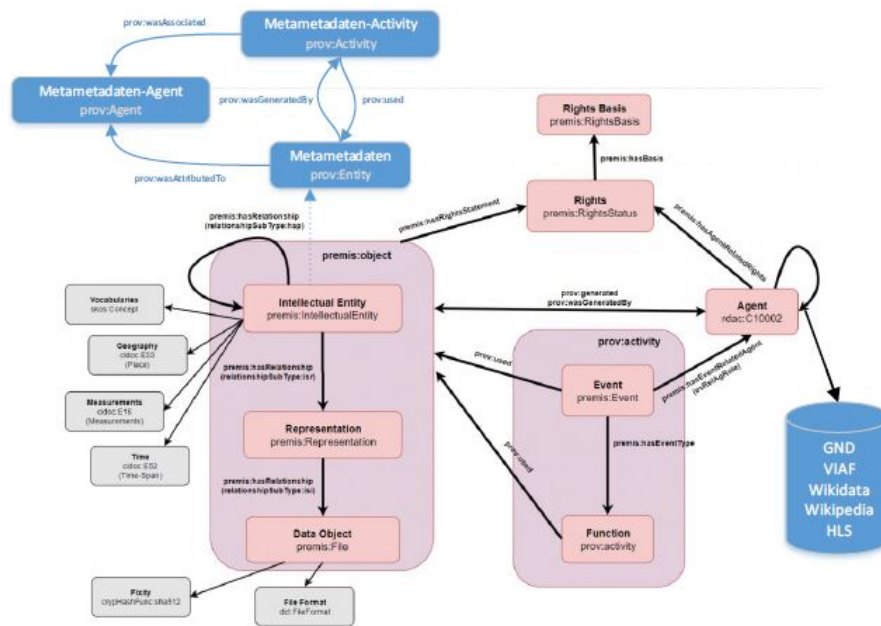


Abb. 16
Erweitertes Matterhorn-RDF-Datenmodell für die Modellierung des Fallbeispiels

Tabelle 4
Verwendete Ontologien und Vokabulare

Prefix	Beschreibung	Namespace URI
Matterhorn RDF		
bpmn	Business Process Model and Notation	http://dkm.fbk.eu/index.php/BPMN_Ontology#
cidoc	CIDOC Conceptual Reference Model	http://www.cidoc-crm.org/cidoc-crm/
dc	Dublin Core	http://purl.org/dc/elements/1.1/
dcterms	DC Terms	http://purl.org/dc/terms/
evReiAgRole	LoC Event Relations Agent Role	http://id.loc.gov/vocabulary/preservation/linkingAgentRoleEvent/
evType	LoC Event Type	http://id.loc.gov/vocabulary/preservation/eventType/
ebucore	Ebucore	http://www.ebu.ch/metadata/ontologies/ebucore/ebucore#
owl	OWL	http://www.w3.org/2002/07/owl
pronom	Pronom	http://reference.data.gov.uk/technical-registry/
prov	Provenance-Family PROV	http://www.w3.org/ns/prov#
premis	Premis	http://id.loc.gov/vocabulary/preservation
rdau	RDA Unconstrained proper-ties	http://rdaregistry.info/Elements/u/
rdac	RDA Class hierarchies	http://rdaregistry.info/Elements/c/
rdf	RDF	http://www.w3.org/1999/02/22-rdf-syntax-ns#
relSubType	LoC Relations Sub Type LOC	http://id.loc.gov/vocabulary/preservation/relationshipSubType/
skos	SKOS Simple Knowledge Organization System	http://www.w3.org/2004/02/skos/core#
time	Time Ontology in OWL	http://www.w3.org/2006/time#
Zusatz		
rdai	RDA Items	http://rdaregistry.info/Elements/i/
xsd	XML Schema	http://www.w3.org/2001/XMLSchema#
schema	Schema.org	https://schema.org/
helvetosaurus	Helvetosaurus	http://helvetosaurus.sozialarchiv.ch/
foaf	The Friend of a Friend RDF vocabulary	http://xmlns.com/foaf/0.1
geo	WGS84 Geo Positioning	http://www.w3.org/2003/01/geo/wgs84_pos#
gsp	GeoSPARQL	http://www.opengis.net/ont/geosparql
rdfs	The RDF Schema vocabulary	http://www.w3.org/2000/01/rdf-schema#

Schritt V: Festlegung von URIs

Die Festlegung der URIs soll dem Vorschlag von Berners-Lee folgen, diese losgelöst von Klassen und Themen zu bilden. Dadurch erübrigt sich

die Frage nach der Benennung der Klassen und die URLs sind kürzer. Ebenfalls sind die URIs damit vollkommen flexibel und nicht an ein Modell gebunden. Da die HTML-Seiten der Ressourcen ein integraler Teil der Website des Sozialarchivs werden sollen, erhalten die URIs deren Autorität (sozialarchiv.ch). Als Identifikator soll für alle Einheiten, die bisher keine Signatur hatten (Personen, geographische Angaben, ...) ein automatisch generierter numerischer Wert verwendet werden:

<http://www.sozialarchiv.ch/id/1>

Aus Gründen der besseren Lesbarkeit und Integration in die jetzige Website sollen Einheiten deren bisherige URL mit der Signatur gebildet wurden dies auch weiterhin so anmelden dürfen. Für den Bestand von Enrique Puelma (F 5015) würde das wie folgt aussehen:

http://www.sozialarchiv.ch/id/F_5015

Die RDF-Darstellungen können entweder über Content Negotiation oder ihre Endung abgerufen werden

http://www.sozialarchiv.ch/id/F_5015.rdf http://www.sozialarchiv.ch/id/F_5015.ttl http://www.sozialarchiv.ch/id/F_5015.trig

Um Metametadaten, deren Aktivitäten und deren Agenten besser von den Metadaten zu trennen sollen diese mit Klassenpfaden modelliert werden: <http://www.sozialarchiv.ch/metametadaten/1>

<http://www.sozialarchiv.ch/agent/1> <http://www.sozialarchiv.ch/activity/1>

Schritt VI: Umwandlung

Im Folgenden werden sieben Verzeichniselemente genauer betrachtet, anhand derer die Herausforderung der Datenumwandlung erläutert werden können. Dabei stehen die Verlinkung, die Fremddatenanreicherungen, die Darstellung von Provenienz sowie Datenbereinigungen im Fokus.

Sprache

Das Verzeichniselement Sprache¹³⁵ wird sowohl in der Datenbank Bild + Ton als auch im Papierarchiv verwendet. Während Erstere die ISO-Norm 639-3:136 nutzt, wird im Papierarchiv das Verzeichniselement in natürlicher Sprache befüllt: «Unterlagen in deutscher Sprache.» Um Anfragen mit anderen Datenquellen zu erleichtern, ist die Verwendung der ISO-Norm zu bevorzugen. Eine Umformung der Werte im Papierarchiv wäre deswegen sinnvoll. Dabei muss jedoch darauf geachtet werden, ob die Angaben neben der Sprache noch weitere Informationen umfassen. Enthält ein Element zum Beispiel die Beschreibung «Enthält zur Hälfte handschriftliche französische Dokumente», muss dies bei der Modellierung berücksichtigt werden.

Da dies beim vorliegenden Fallbeispiel nicht zutrifft, würde sich eine Umwandlung leicht realisieren lassen.

RDF 18

Der Bestand Ar 137 mit Archivgut in deutscher Sprache

```
1 @prefix dbpedia-owl: <http://dbpedia.org/ontology/> .
2 @prefix ssa: <http://www.sozialarchiv.ch/id/> .
3 @prefix ssaMetaMetaData:
  <http://www.sozialarchiv.ch/metametadatas/> .
4
5 ssaMetaMetaData:731 {
6 ssa:Ar_137 dbpedia-owl:iso6393Code "ger" . }
```

Da es verschiedene ISO-Normen für Sprachen gibt, weist diese Modellierung auch gewisse Schwächen auf. Weiterhin bietet sie keine Sprachbezeichnungen an.

Aus diesem Grund verwenden viele Datenanbieter wie WorldCat, die British Library und DBpedia die Datenquelle lexvo.org.¹³⁷ Neben der Verbreitung spricht für diese Datenquelle ihr grosser Umfang von mehr als 7000 Sprach-Identifikatoren¹³⁸ und ihre freie Lizenz (CC BY-SA 3.0).¹³⁹ Eine Modellierung mit lexvo.org würde wie folgt aussehen:

RDF 19

Der Bestand Ar 137 mit Archivgut in deutscher Sprache mit lexvo.org

```
1 @prefix dcterms: <http://purl.org/dc/terms/> .
2 @prefix ssa: <http://www.sozialarchiv.ch/id/> .
3 @prefix ssaMetaMetaData:
  <http://www.sozialarchiv.ch/metametadatas/> .
4
5 ssaMetaMetaData:731 {
6 ssa:Ar_137 dcterms:language -3/ger> . }
```

Versionierung

Aufgrund ihrer Einfachheit und dem Umstand, dass sich die Metadaten im Archiv- Bereich meist selten ändern, soll eine Versionierung genutzt werden, die bei jeder Version alle Eigenschaften aufführt. Im Folgenden Beispiel wird die Sprache des Bestands Ar_137 von Englisch auf Deutsch gewechselt. Der Titel bleibt unverändert.

RDF 20

Die Archivale Ar 137 mit versionierten Metadaten.

```

1 @prefix dcterms: <http://purl.org/dc/terms/> .
2 @prefix prov: <http://www.w3.org/ns/prov#> .
3 @prefix rdau: <http://rdaregistry.info/Elements/u/> .
4 @prefix xsd:
5   <http://www.w3.org/2001/XMLSchema#> .
6 @prefix ssa: <http://www.sozialarchiv.ch/id/> .
7 @prefix ssaAgent: /> .
8 @prefix ssaActivity:
9   <http://www.sozialarchiv.ch/activity/> .
10 @prefix ssaMetaMetaData:
11   <http://www.sozialarchiv.ch/metametadatas/> .
12
13 ssaMetaMetaData:5 {
14   ssa:Ar_137 dcterms:language -3/ger> ; }
15 ssaMetaMetaData:4 {
16   ssa:Ar_137 dcterms:title "Robert Grimm als
17   Redner"^^xsd:string . }
18 ssaMetaMetaData:0 {
19   ssaMetaMetaData:5 prov:wasGeneratedBy
20   ssaActivity:38 .
21   ssa:Ar_137 dcterms:hasVersion ssa:Ar_137v1 . }
22 ssaMetaMetaData:4 {
23   ssa:Ar_137v1
24   dcterms:title "Robert Grimm als Redner"^^xsd:string ;
25   dcterms:language - 3/eng> . }
26 ssaMetaMetaData:0 {
27   ssa:Ar_137v1 dcterms:isVersionOf ssa:Ar_137 .
28   ssaMetaMetaData:4 prov:wasGeneratedBy
29   ssaActivity:37 . }
30 ssaMetaMetaData:0 {
31   ssaActivity:38
32   a prov:activity;
33   prov:startedAtTime "2018-04-
34   15T13:00:00+02:00"^^xsd:dateTime;
35   prov:endedAtTime "2018-04-
36   15T14:00:00+02:00"^^xsd:dateTime;
37   prov:wasAssociatedWith ssaAgent:1 .
38   ssaActivity:37
39   a prov:activity;
40   prov:startedAtTime "2017-04-
41   15T13:00:00+02:00"^^xsd:dateTime;
42   prov:endedAtTime "2017-04-
43   15T14:00:00+02:00"^^xsd:dateTime;
44   prov:wasAssociatedWith ssaAgent:1 . }
45 ssaMetaMetaData:400 {
46   ssaAgent:1
47   a prov:Agent, prov:Person ;
48   rdau:P60368 "Erika Mustermann"^^xsd:string . }

```

Die aufgezeichneten Versionen können dann direkt über ihre URI oder mit SPARQL- Anfragen aufgerufen werden. Mit der folgenden SPARQL-Abfrage könnte man sich zum Beispiel alle versionierten Einträge des Sprachfelds vor 2018 anzeigen lassen. Damit würde es vollkommen transparent, wer, wie und wann ein Metadaten-Eintrag verändert hat.

SPARQL 2

Anzeige aller Metadaten-Version der Archive Ar 137 vor dem 1. Januar 2018.

```

1 PREFIX dcterms:
2 PREFIX prov:
3 PREFIX xsd: xmlns:xsd="">
4 PREFIX ssa:
5
6 SELECT ?version ?language ?date WHERE {
7   ssa:Ar_137 dcterms:hasVersion ?version .
8   graph ?g {
9     ?version dcterms:language ?language
10  } .
11  ?activity prov:endedAtTime ?date .
12  ?g prov:wasGeneratedBy ?activity .
13 FILTER (?date < "2018-01-
14 01T00:00:00+02:00"^^xsd:dateTime)

```

Personen

Die Datenbank Bild+Ton besitzt ein Verzeichniselement, in welchem Personen angegeben werden können, die in Beziehung zu dem betreffenden Archive stehen.¹⁴⁰ Im Papierarchiv können zusätzlich Personen aufgrund von Named-Entity-Recognition aus diversen Verzeichniselementen extrahiert werden. Will man diese Angaben archivübergreifend vernetzen, ist eine Verlinkung zu Normdatenbanken wie der GND oder der VIAF sinnvoll. Daneben bietet sich auch eine Verlinkung zu Datenquellen wie zum Beispiel Wikidata oder dem Historischen Lexikon der Schweiz (HLS/DHS/DSS) an, um die eigenen Daten mit weiteren Informationen anzureichern. Im vorliegenden Fallbeispiel könnte diese wie folgt aussehen.

RDF 21

Verlinkung einer Persson (Robert Grimm) mit dem Bilddokument F 5069-Fa-031.

```

1 @prefix prov: <http://www.w3.org/ns/prov#> .
2 @prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
3 @prefix schema: <https://schema.org/> .
4 @prefix rdai: <http://rdaregistry.info/Elements/i/> .
5 @prefix rdau: <http://rdaregistry.info/Elements/u/> .
6 @prefix ssa: <http://www.sozialarchiv.ch/id/> .
7 @prefix ssaPerson: <http://www.sozialarchiv.ch/p/> .
8 @prefix ssaAgent: /> .
9 @prefix ssaActivity:
10 <http://www.sozialarchiv.ch/activity/> .
11 @prefix ssaReference:
12 <http://www.sozialarchiv.ch/activity/> .
13 @prefix ssaMetaMetaData:
14 <http://www.sozialarchiv.ch/metametadate/> .
15
16 ssaMetaMetaData:731 {
17 ssa:F_5069-Fa-031 rdai:P40073 ssaPerson:1 .
18
19 ssaPerson:1
20 schema:sameAs <http://d-nb.info/gnd/11869779X>;
21 schema:sameAs <http://www.hls-dhs-
22 dss.ch/textes/d/D4516.php>;
23
24 schema:sameAs
25 <https://www.wikidata.org/wiki/Q115645>;
26 rdau:P60368 "Robert Grimm"^^xsd:string . }
27
28 ssaMetaMetaData:732 {
29 ssaPerson:1
30 rdau:P60599 "1881-04-16"^^xsd:date . }
31
32 ssaMetaMetaData:0 {
33 ssaMetaMetaData:731
34 prov:wasGeneratedBy ssaActivity:1 . }
35
36 ssaMetaMetaData:732
37 prov:wasGeneratedBy ssaActivity:1;
38 prov:wasDerivedFrom: ssaReference:1;
39 prov:wasDerivedFrom: ssaReference:2 .
40
41 ssaActivity:1
42 a prov:activity;
43 prov:startedAtTime "2018-04-
44 15T13:00:00+02:00"^^xsd:dateTime;
45 prov:endedAtTime "2018-04-
46 15T14:00:00+02:00"^^xsd:dateTime;
47 prov:wasAssociatedWith agent:1 .
48
49 ssaReference:1
50 dcterms:references:
51 <http://d-nb.info/gnd/11869779X>;
52 prov:wasGeneratedBy: ssaActivity:1;
53 prov:wasAttributedTo ssaAgent:1 .
54
55 ssaReference:2
56 dcterms:references: <http://www.hls-dhs-
57 dss.ch/textes/d/D4516.php>;
58 prov:wasGeneratedBy: ssaActivity:1;
59 prov:wasAttributedTo ssaAgent:1 . }
60
61 ssaMetaMetaData:401 {
62 ssaAgent:1
63 a prov:Agent, prov:Person;
64 rdau:P60368 "Erika Mustermann"^^xsd:string . }

```

Der Hauptaufwand bei einer Umwandlung in eine solche Modellierung ist der Zustand der jetzigen Metadaten: So wurden Personen bisher lediglich mit Vor- und Nachname sowie manchmal noch mit Geburts- und Sterbedatum verzeichnet. Die sichere Verknüpfung zu Normdaten-Einträgen aus der GND, VIAF, Metagrid und Wikidata ist dementsprechend schwer. Für eine Lösung wurde testweise ein Skript erstellt, welches alle extrahierten Personeneinträge in klare Übereinstimmungen (gleicher Name und Geburts-, Sterbedaten), manuell zu prüfende Übereinstimmungen (z.B. ähnlicher Name oder fehlendes Todesdatum) sowie Einträge ohne Normdaten einteilte. Um später eine möglichst gute Erkennung zu gewährleisten, müssten die manuell zu prüfenden Einträge im AIS mit einem Identifikator einer Normdatenbank ergänzt werden.

Helvetosaurus

Die Datenbank Bild+Ton verfügt auf der Ebene Dokument über eine inhaltliche Erschließung durch einen Thesaurus.¹⁴¹ Das Schweizerische Sozialarchiv entschied sich dazu, eine bereits existierende Lösung zu verwenden, nämlich den Helvetosaurus der Schweizerischen Parlamentsdienste. Es handelt sich dabei um eine an Schweizer Bedürfnisse angepasste Version des EuroVoc, eines Thesaurus des Europäischen Parlaments und des Amtes für amtliche Veröffentlichungen der Europäischen Union.¹⁴² Nachdem die Parlamentsdienste allerdings 2003 überraschend beschlossen, den Helvetosaurus nicht mehr einzusetzen, verwendet diesen nun nur noch das Sozialarchiv. Da eine Migration zu einem anderen Thesaurus als zu aufwendig eingestuft wurde, ist er auch weiterhin im Sozialarchiv in Gebrauch.

Aus Sicht der Verlinkung schien das Verzeichniselement Helvetosaurus deswegen zunächst uninteressant. Im Zuge dieser Arbeit konnte aber folgende Idee realisiert werden: Zunächst wurden aus dem letzten Snapshot der Helvetosaurus-Website sämtliche Daten (ID, Bezeichnung, Synonyme, Verwandte Terme usw.) der einzelnen Helvetosaurus-Einträge extrahiert. Aus diesen wurde unter Verwendung von SKOS eine RDF-Version erstellt. Um den Helvetosaurus von seinem Inseldasein zu befreien, wurde durch ein Skript und manueller Nacharbeit ein Abgleich zum EuroVoc, dem STW und der GND vorgenommen und die Verweise der RDF-Version hinzugefügt. Das Ergebnis ist online einsehbar¹⁴³ und verfügt über CoolURIs mit Content-Negotiation.¹⁴⁴

Bei der Verlinkung der erstellten Linked-Open-Data-Version mit der Datenbank Bild+Ton konnten diverse kleinere Abweichungen aufgespürt und korrigiert werden. Dazu gehören vor allem Begriffe im Verzeichniselement, die nicht in der Version 9 des Helvetosaurus enthalten sind oder manuell bei der Katalogisierung hinzugefügt wurden.

Dank der Verlinkung sind nun auch komplexe Anfragen realisierbar. Beispielsweise wäre es denkbar, sich zu einem Bild im Archiv die Bücher aus der Bibliothek mit den gleichen GND-Schlagworten anzeigen zu lassen.

RDF 22

Das Bilddokument F 5030-Fa-0074 mit den Helvetosaurus-Schlagwörtern Auto, Polizei und Landverkehr

```
1 @prefix rdau: <http://rdaregistry.info/Elements/u/> .
2 @prefix helvetosaurus:
3   <http://helvetosaurus.sozialarchiv.ch/> .
4 @prefix ssa: <http://www.sozialarchiv.ch/id/> .
5 @prefix ssaMetaMetaData:
6   <http://www.sozialarchiv.ch/metametadate/> .
7 ssa:F_5030-Fa-0074 rdau:P60805
8 helvetosaurus:L05K1803010101;
9 helvetosaurus:L04K04030304;
10 helvetosaurus:L02K1803 .
```

Geographische Angaben

Die Datenbank Bild+Ton enthält zwei Felder für geographische Angaben: Geopolitik und Landschaft.¹⁴⁵ Geopolitik beinhaltet eine hierarchische Auswahl, die maximal bis zur politischen Gemeinde reicht: Europa # Schweiz # Zürich, Kanton Zürich, Stadt. Das Feld Landschaft umfasst natürliche geographische Angaben wie Berge, Seen oder Flüsse. Das Papierarchiv enthält ein Verzeichniselement zu den geographischen Angaben.

Die bisherige Verzeichnungspraxis weist gewisse Nachteile auf: (1) Da die Angaben nicht normiert sind, lassen sie sich schlecht mit anderen Archiven vergleichen. (2) Da die Daten keine geographischen Koordinaten beinhalten, ist keine «in der Nähe von»-Suche möglich. (3) Die Angaben beinhalten keine alternativen Schreibweisen und Übersetzungen. All diese Probleme lassen sich jedoch durch eine gute Modellierung in Linked-Data lösen. Primär müssen dafür Ressourcen statt Literalen verwendet werden. Diese können dann mit externen Fremddaten verknüpft und angereichert werden. Eine Auflistung möglicher externer Quellen findet sich in Tabelle 5 auf S. 377.

Da das Verzeichniselement der Datenbank Bild+Ton Orte auf der ganzen Welt umfasst, wäre eine Verlinkung zu GeoNames am besten geeignet. Mit zusätzlichen Verlinkungen zu ld.geo.admin.ch, Wikidata und ortsnamen.ch könnte eine zusätzliche Informationstiefe erreicht werden.

Im folgenden Beispiel ist eine solche Lösung modelliert:

RDF 23

Verknüpfung des Bildes F_5030-Fa-0074 mit der Stadt Zürich

```

1 @prefix dct: <http://purl.org/dc/terms/> .
2 @prefix schema: <https://schema.org/> .
3 @prefix geo: <http://www.opengis.net/ont/geosparql#> .
4 @prefix gsp: <http://www.opengis.net/ont/geosparql#> .
5 @prefix rdf: <http://www.w3.org/2001/XMLSchema#> .
6 @prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
7 @prefix prov: <http://www.w3.org/ns/prov#> .
8 @prefix foaf: <http://xmlns.com/foaf/0.1/> .
9 @prefix ssa: <http://www.sozialarchiv.ch/metametadata/> .
10 @prefix ssaGeo: <http://www.sozialarchiv.ch/geo/> .
11 @prefix ssaAgent: <http://www.sozialarchiv.ch/agent/> .
12 @prefix ssaActivity: <http://www.sozialarchiv.ch/activity/> .
13 @prefix ssaGeometry: <http://www.sozialarchiv.ch/geometry/> .
14 @prefix ssaReference: <http://www.sozialarchiv.ch/reference/> .
15
16 ssa:MetaMetadata:731 {
17   ssa:F_5030-Fa-0074 dct:location ssaGeo:1 .
18 }
19
20 ssaGeo:1
21   rdf:type <http://schema.org/AdministrativeArea> ;
22   rdf:type <http://www.opengis.net/ont/geosparql#Feature> ;
23   rdf:type <http://www.geonames.org/ontology#ADM3> ;
24   rdf:label "Zürich"@de .
25 }
26
27 ssa:MetaMetadata:731 {
28   prov:wasGeneratedBy ssaActivity:1 .
29 }
30
31 ssaActivity:1
32   a prov:Activity ;
33   prov:startedAtTime "2018-04-15T13:00:00+02:00"^^xsd:dateTime ;
34   prov:endedAtTime "2018-04-15T14:00:00+02:00"^^xsd:dateTime ;
35   prov:wasAssociatedWith ssaAgent:11 .
36 }
37
38 ssa:MetaMetadata:801 {
39   ssaAgent:11
40     a prov:Agent, prov:Person ;
41     foaf:name "Erika Mustermann"^^xsd:string ;
42 }
43
44 ssa:MetaMetadata:732 {
45   ssaGeo:1
46     schema:sameAs ;
47     schema:sameAs <http://www.hls-dhs-dss.ch/textes/d/D171.php> .
48 }
49
50 ssa:MetaMetadata:732 {
51   prov:wasGeneratedBy ssaActivity:2 .
52 }
53
54 ssaActivity:2
55   a prov:Activity ;
56   prov:startedAtTime "2009-06-30T18:30:00+02:00"^^xsd:dateTime ;
57   prov:endedAtTime "2009-06-30T18:30:00+02:00"^^xsd:dateTime ;
58   prov:wasAssociatedWith ssaAgent:15 .
59 }
60
61 ssa:MetaMetadata:400 {
62   ssaAgent:15
63     a prov:Agent, prov:SoftwareAgent ;
64     dct:title "geo-bot"^^xsd:string ;
65     schema:version "1.0.0"^^xsd:string ;
66     foaf:homepage "https://www.sozialarchiv.ch/bots/geo/"^^xsd:string ;
67 }
68
69 ssa:MetaMetadata:733 {
70   ssaGeo:1
71     schema:sameAs <http://www.wikidata.org/entity/Q72> ;
72     gsp:hasGeometry ssaGeometry:1 .
73 }
74
75 ssaGeometry:1
76   rdf:type <http://www.opengis.net/ont/geosparql#POLYGON> ;
77   gsp:asWKT "POLYGON((8 5736871209768 47 420034458028, ...))"^^xsd:string .
78 }
79
80 ssa:MetaMetadata:733 {
81   prov:wasGeneratedBy ssaActivity:2 ;
82   prov:wasDerivedFrom ssaReference:80 .
83 }
84
85 ssaReference:80
86   dct:references ;
87   prov:wasGeneratedBy ssaActivity:2 .
88 }
89
90 ssa:MetaMetadata:734 {
91   ssaGeo:1
92     geo:lat "47.36667"^^xsd:string ;
93     geo:long "8.55"^^xsd:string ;
94     schema:sameAs <http://sws.geonames.org/2657896> ;
95     schema:sameAs <https://de.wikipedia.org/wiki/Z%C3%BCrich> .
96 }
97
98 ssa:MetaMetadata:734 {
99   prov:wasGeneratedBy ssaActivity:2 ;
100   prov:wasDerivedFrom ssaReference:81 .
101 }
102
103 ssaReference:81
104   dct:references ;
105   prov:wasGeneratedBy ssaActivity:2 .
106 }

```


Tabelle 5
Quellen für geographische Angaben mit schweiz- oder weltweitem Fokus. ¹⁴⁶

Anbieter/Dienst	Anzahl Einträge (Varianten)	Geographische Abdeckung	Historisierung	Alt. Namen / Varianten	Mundartname	Beschreibung	Namensdeutung	Quellen	Koordinaten	Administrative Unterteilung	M. u. M.	Nur Orte	Open Data	Linked Data
ortsnamen.ch	300'000 (500'000)	Teile der Deutschschweiz			x	x	x	x	x	x		x		
Rechtsquellenstiftung	3'790 (10'718)	Vorwiegend Schweiz	x	x						x			x	
Dodis	8636	Weltweit, mit Fokus auf der Schweiz		x										
Schweizerische Eidgenossenschaft (ld.geo.admin.ch)	300'000	Schweiz und Liechtenstein							x	x	x		x	x
GeoNames	9'000'000 (10'000'000)	Weltweit		x					x	x	x		x	x
Getty Thesaurus of Geographic Names	2,550.000 (4,100.000)	Weltweit		x					x	x	x		x	x
GND	300.138	Weltweit		x					x	x			x	x

Webarchiv

Das Sozialarchiv beteiligt sich am Webarchiv Schweiz, welches durch die Schweizer Nationalbibliothek betrieben wird. ¹⁴⁷ Viele der Websites, die das Sozialarchiv durch die Nationalbibliothek archivieren lässt, stehen in direktem Zusammenhang mit Beständen aus dem Papierarchiv, so zum Beispiel die Website des Schweizerischen Friedensrats, deren Papierarchiv sich im Sozialarchiv befindet. Die Daten der Harvests sind über OAI-PMH aufrufbar und als Open Data freigegeben. ¹⁴⁸

Bei der Provenienz geht es darum, die Datenebene von der Metadatenebene zu unterscheiden. Auf Datenebene kann der Herausgeber als Provenienz der Website genannt werden (RDF 24, Zeile 38). Auf der Metadatenebene ist hingegen die Nationalbibliothek zu nennen (RDF 24, Zeile 44ff.) – einseits aufgrund der Software, die sie zum Harvesten einsetzt, und anderseits durch die Katalogisierung der Einträge. Zusätzlich wäre es auch von Vorteil, den Importvorgang durch das Sozialarchiv zu dokumentieren. Ein Modellierung könnte wie folgt aussehen:

RDF 24

Beispiel der Harvests der Website <http://www.friedensrat.ch> durch das Webarchiv Schweiz

```

1 @prefix dcterms: <http://purl.org/dc/terms/> .
2 @prefix foaf: <http://xmlns.com/foaf/0.1/> .
3 @prefix rdfs: <schema#> .
4 @prefix schema: <https://schema.org/> .
5 @prefix prov: <http://www.w3.org/ns/prov#> .
6 @prefix rdac: <> .
7 @prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
8 @prefix ssa: <http://www.sozialarchiv.ch/id> .
9 @prefix ssaMetaMetaData:
10 <http://www.sozialarchiv.ch/metametadada> .
11 @prefix ssaAgent: <http://www.sozialarchiv.ch/agent/> .
12 @prefix ssaActivity: <http://www.sozialarchiv.ch/activity/> .
13 ssaMetaMetaData:821 {
14 ssa:Ar_76 dcterms:hasPart ssa:sz001617658 .
15 ssa:sz001617658 schema:sameAs
16 <http://permalink.snl.ch/bib/sz001617658> . }
17 ssaMetaMetaData:0 {
18 ssaMetaMetaData:821 prov:wasGeneratedBy ssaActivity:1 .
19 ssaActivity:1
20 a prov:Activity;
21 prov:startedAtTime "2018-04-
22 15T13:00:00+02:00"^^xsd:dateTime;
23 prov:endedAtTime
24 "2018-04-15T14:00:00+02:00"^^xsd:dateTime;
25 prov:wasAssociatedWith ssaAgent:1 . }
26 ssaMetaMetaData:400 {
27 ssaAgent:11
28 a prov:Agent, prov:Person;
29 foaf:name "Erika Mustermann"^^xsd:string; }
30
31
32 ssaMetaMetaData:822 {
33 ssa:sz001617658
34 dcterms:title "Schweizerischer Friedensrat : SFR = Swiss Peace
35 Council = Conseil suisse pour la paix = Consiglio svizzera per
36 la pace";
37 schema:url "http://www.friedensrat.ch";
38 dcterms:hasPart ssa:bel-174767;
39 rdac:C10002 "Zürich : Schweizerischer Friedensrat"@de .
40 ssa:bel-174767
41 dcterms:isPartOf ssa:sz001617658;
42 foaf:page <http://www.e-
43 helvetica.nb.admin.ch/directAccess?callnumber=bel-174767>
44 . }
45 ssaMetaMetaData:439 {
46 ssaAgent:10
47 a prov:Agent, prov:SoftwareAgent;
48 dcterms:title "webarchiv-import-bot";
49 rdfs:seeAlso - professionell/e-helvetica/infos-fuer-anbieter-
50 innen/websites-_webarchiv-schweiz.html>;
51 rdfs:seeAlso <https://www.sozialarchiv.ch/bot/webarchiv-
52 import> .
53 ssaAgent:11
54 a prov:Agent, prov:SoftwareAgent;
55 dcterms:title "webarchiv-crawler";
56 dcterms:description "Das Webarchiv Schweiz wird durch die
57 Schweizerische Nationalbibliothek betrieben. Es verwendet den
58 Crawler Heritrix und Open Wayback. Der Import in die
59 Linked-Open-Data-Umgebung des Schweizerischen
60 Sozialarchiv übernimmt das Skript
61 WebArchiveSwitzerlandImport"@de;
62 rdfs:seeAlso ;
63 rdfs:seeAlso - archiving/openwayback/>;
64 rdfs:seeAlso - professionell/e-helvetica/infos-fuer-anbieter-
65 innen/websites-_webarchiv-schweiz.html> . }
66 ssaMetaMetaData:0 {
67 ssaActivity:3
68 a prov:Activity;
69 prov:startedAtTime "2018-04-15T13:00:00"^^xsd:dateTime;
70 prov:endedAtTime "2018-04-15T13:00:00"^^xsd:dateTime;
71 prov:wasAssociatedWith ssaActivity:10;
72 prov:wasAssociatedWith ssaActivity:11 . }

```

Veröffentlichungen

Eine von den Benutzenden oft gewünschte Information zu den Archivalien ist, ob diese bereits in einer Publikation oder einer Arbeit verwendet wurden. Da die Benutzungsordnung die Abgabe eines Belegexemplars vorsieht, sind diese Informationen zumindest auf Titlebene vorhanden. So findet sich zum Beispiel auf der Website des Sozialarchivs eine Auflistung von Publikationen und Arbeiten mit den Quellenmaterialien des Archives.¹⁴⁹ Eine Zuordnung zu den Archivalien wurde jedoch nicht vorgenommen; lediglich auf Bestands-Ebene finden sich vereinzelt Nennungen.¹⁵⁰

Eine Modellierung könnte wie folgt aussehen:

RDF 25

Bestand Ar 459 mit der Veröffentlichung «Formierter Widerstand»

```

1 @prefix rdau: <http://rdaregistry.info/Elements/u/>.
2 @prefix dc: <http://purl.org/dc/terms/>.
3 @prefix dcterms: <http://purl.org/dc/terms/>.
4 @prefix schema: <https://schema.org/>.
5 @prefix xsd: <http://www.w3.org/2001/XMLSchema#>.
6 @prefix foaf: <http://xmlns.com/foaf/0.1/>.
7 @prefix prov: <http://www.w3.org/ns/prov#>.
8 @prefix ssaMetaMetaData:
9   <http://www.sozialarchiv.ch/metametadatas/>.
10 @prefix ssa: <http://www.sozialarchiv.ch/id/>.
11 @prefix ssaAgent: />.
12 @prefix ssaActivity: <http://www.sozialarchiv.ch/activity/>.
13 ssaMetaMetaData:821 {
14   ssa:Ar_459
15   rdau:P60313 ssa:111618;
16 }
17
18 ssaMetaMetaData:0 {
19   ssaMetaMetaData:821 prov:wasGeneratedBy ssaActivity:1 .
20 }
21 ssaActivity:1
22 a prov:Activity;
23 prov:startedAtTime "2018-04-
24   15T13:00:00+02:00"^^xsd:dateTime;
25 prov:endedAtTime
26   "2018-04-15T14:00:00+02:00"^^xsd:dateTime;
27 prov:wasAssociatedWith ssaAgent:14 . }
28 ssaMetaMetaData:400 {
29   ssaAgent:14
30   a prov:Agent, prov:Person;
31   foaf:name "Erika Mustermann"^^xsd:string . }
32 ssaMetaMetaData:822 {
33   ssa:111618
34   rdau:P60369 "Formierter Widerstand. Geschichte der
35     kommunistischen Bewegung in der Schweiz 1944-1991";
36   dc:identifier: "SozArch 111618". }
37 ssaMetaMetaData:0 {
38   ssaMetaMetaData:822 prov:wasGeneratedBy ssaActivity:1 . }
39
40 ssaMetaMetaData:401 {
41   ssaAgent:19
42   a prov:Agent, prov:SoftwareAgent;
43   dcterms:title "nebis-import-bot"^^xsd:string;
44   schema:version "1.0.0"^^xsd:string;
45   foaf:homepage "https://www.sozialarchiv.ch/bots/nebis-
46     import"^^xsd:string . }

```

Neben den selbstgesammelten Informationen könnten zudem Angaben aus Forschungs-Repositoryen wie Zora oder Boris der Universitäten Zürich und Bern extrahiert werden. Da Anreicherungen jedoch oft viel manuelle Arbeit erfordern, wäre für die Zukunft ein Crowdsourcing-Ansatz anstrengenswert. Die Benutzer sollten zudem die Möglichkeit haben, Verlinkungen von Publikationen oder Arbeiten direkt auf der Archiv-Website einzureichen. Die Herkunft dieser Daten müsste natürlich wiederum abgebildet werden.

Zwischenfazit

Das Kapitel hat gezeigt, dass eine Modellierung mit den Metadaten des Schweizerischen Sozialarchivs möglich ist. Die aufgestellten Qualitätsmerkmale können durchaus erreicht werden, auch wenn damit teilweise eine grössere Menge an Tripeln einhergeht. Den grössten Aufwand stellt die Umformung von Fliesstexten und unnormierten Daten dar. So zum Beispiel bei Personen. Ebenfalls eine Herausforderung ist die Wahl eines geeigneten Datenmodells.

Empfehlungen

Wie im vierten Kapitel gezeigt, befindet sich Linked Open Data noch am Anfang seines Weges im Archivwesen. Es fehlen noch entsprechende Standards und aus den wenigen vorhandenen Anwendungen kann sich keine Best-Practice ableiten lassen. Dementsprechend ist es schwer, zukunftsichere Empfehlungen aufzustellen. Den- noch lässt sich aus den Erkenntnissen dieser Arbeit durchaus eine Einschätzung ableiten, was Archive, die eine Linked-Open-Data-Anwendung erstellen möchten, tun können.

Ganz allgemein müssen die archivalischen Metadaten in eine Form umgewandelt werden, die für Maschinen zu verstehen ist. Um dies zu erreichen sollten die im dritten Kapitel aufgeführten Qualitätsmerkmale berücksichtigt werden.

Eine erste konkrete Massnahme sollte die Entwicklung eines guten Verständnisses der eigenen momentanen Datenstruktur sein. Dies kann zum Beispiel - falls noch nicht vorhanden - durch die Erstellung einer ausführlichen Dokumentation der genutzten Verzeichniselemente erreicht werden.¹⁵¹ Ebenfalls zu Beginn sollte auch analysiert werden, welche externen Datenquellen für den eigenen Datensatz interessant sein könnten. Für vermutlich viele Archive kommen zum Beispiel Normdaten für Personen und geographische Angaben in Frage.

Eine weitere wichtige Voraussetzung ist die Lizenzierung der Metadaten. Am besten ist es, möglichst schnell abzuklären, ob die Metadaten als Open Data veröffentlicht werden können. Eine gute Wahl für die Lizenzierung wäre die viel genutzte Lizenz CC0.

Danach soll ein Datenmodell gewählt werden. Hierzu gilt es die eigenen Voraussetzungen und Anforderungen zu evaluieren sowie zu analysieren, welche Datenmodelle fachverwandte Institute benutzen.¹⁵² Dabei sollte auch festgelegt werden, ob und wie ausführlich die Metadaten-Provenienz und eine Versionierung dargestellt werden sollen. Die Arbeit hat dafür eine Lösung vorgestellt, dank der ein hohes Mass an Transparenz und Authentizität im archivischen Sinn erreicht werden kann.¹⁵³

Ausgehend von diesen Schritten kann dann die eigentliche Umwandlung erfolgen. Da es noch kein Linked-Data fähiges AIS gibt, ist wohl im Moment der beste Weg eine separate Linked-

Data-Anwendung kontinuierlich mit den Daten aus dem AIS zu bespielen. Bei der Umwandlung gilt es möglichst, maschinenlesbare Ressourcen und Literale zu erzeugen. Dies kann zum Beispiel beim Verzeichniselement Sprache mit der Verknüpfung zum Linked-Open-Data-Sprachverzeichnis lexvo.org erreicht werden.¹⁵⁴ Umso mehr der eigene Datensatz mit allgemein gebräuchlichen Ressourcen verlinkt wird, desto höher wird dessen Nutzen sein. Denn diese Verknüpfungen vereinfachen den Abgleich mit anderen Datensätzen und machen so die Nutzung von Linked Open Data sinnvoll.

Da die Aktualisierung der Linked-Open-Data-Anwendung in regelmässigen und kurzen Abständen erfolgen sollte, muss die Umwandlung voll automatisch geschehen. In Fällen, bei denen bestehende Metadaten von Hand angepasst werden müssen¹⁵⁵, sollten diese Änderungen folglich im AIS geschehen. Zukünftige Metadaten sollten so eingegeben werden, dass sie problemlos umgewandelt werden können. Wenn nötig, sollten hierfür die Eingabemasken oder Verzeichnisregeln angepasst werden.

Zusammenfassend bleibt festzuhalten, dass momentan die Erstellung einer Linked-Data-Modellierung noch mit viel Arbeit verbunden ist. Dennoch ist die Erstellung von Prototypen sehr wichtig. Denn vor allem anhand der Praxis können die unterschiedlichen Datenmodelle, Provenienz-Modellierungen und andere Merkmale diskutiert werden.

Neben den konkreten Massnahmen sollten die Archive auch weiterhin ihr Wissen im Gebiet Linked Open Data ausbauen. Dafür sollten sie den Austausch innerhalb der Archivwelt und auch mit verwandten Fachgebieten wie den Bibliotheken fördern. Ebenfalls ist es wichtig, das Weiterbildungsangebot zu diesem Thema auszubauen. Des Weiteren sollten die Archive den Dialog mit den Herstellern intensivieren und auf eine möglichst hohe Kompatibilität der Lösungen hinwirken. Vor allem bei der Anschaffung von neuen OAIS-Systemen sollte auf eine Linked-Open-Data-Kompatibilität geachtet werden.

Schlusswort

Die Arbeit hat gezeigt, dass die Archivgemeinschaft einerseits begonnen hat, sich intensiv mit Linked Open Data auseinanderzusetzen, andererseits das Thema in diesem Bereich aber noch sehr am Anfang steht. So ist mit *Records in Context* ein neuer Verzeichnisstandard zwar in der Entwicklung, aber noch nicht fertiggestellt. Ebenfalls sind die wenigen bereits existierenden Linked-Open-Data-Anwendungen im Archivbereich eher noch in frühen Phasen und unterscheiden sich stark voneinander. Dementsprechend unmöglich ist es, eine Best-Practice aus ihnen abzuleiten und daraus zukunftsichere Empfehlungen auszusprechen.

Die Arbeit konnte aber auch aufzeigen, dass es durchaus schon jetzt möglich ist, zu erstrebende Qualitätsmerkmal für Metadaten im Bezug auf Linked Open Data zu formulieren. Anhand der Metadaten des Schweizerischen Sozialarchivs konnte auch nachgewiesen werden, dass

diese sich durchaus in der Praxis mit archivalischen Metadaten umsetzen lassen.

Doch welche Massnahmen sollen Archive bezüglich ihrer bestehenden Metadaten und welche bezüglich ihrer zukünftigen Bestände für Linked Open Data treffen? Ganz allgemein müssen die archivalischen Metadaten verständlicher für Maschinen werden. Konkret bedeutet das, dass Objekte nicht mehr vornehmlich durch Fliesstext, sondern möglichst als verlinkte Ressourcen verzeichnet werden sollen. Statt zum Beispiel «Der Bestand enthält Dokumente in deutscher Sprache» kann eine Verlinkung zur häufig verwendeten Linked-Open-Data-Sprachverzeichnis lexvo.org vorgenommen werden. Bei den Datenmodellen sollten die Archive versuchen eine möglichst einheitliche Basis zu nutzen, um so systemübergreifende Abfragen zu erleichtern. Daneben sollten die Archive anstreben, ihre Metadaten unter freien Lizenzen zu veröffentlichen. Weitere Aspekte wie zum Beispiel Empfehlungen zur Metadaten-Provenienz oder der Versionierung finden sich im fünften Kapitel dieser Arbeit. Um die Entwicklung von Linked Open Data voranzubringen sollten die Archive neben den konkreten Massnahmen auch den fachinternen und fachübergreifenden Austausch fördern. Ebenfalls sollten die Weiterbildungsangebote in diesem Bereich ausgebaut werden.

Für weiterführende Arbeiten bietet sich eine Fülle von Themen an. Sobald mehr Anwendungen im Regelbetrieb sind, könnten zum Beispiel deren Workflows untersucht und verglichen werden. Ebenfalls könnte dann deren effektive Nutzung und deren Mehrwert erforscht werden.

Bibliographie

- Arbeitsgruppe Normen und Standards (VSA), Tögel, Bettina. & Borrelli, Graziella: Schweizerische Richtlinie für die Umsetzung von ISAD(G) – International Standard Archival Description (General). Zürich 2009. Online: https://vsa-aas.ch/wpcontent/uploads/2015/06/Richtlinien_ISAD_G_VSA_d.pdf [Zugriff am 29. 6. 2018].
- Arbido: Linked open data, big data, alles vernetzt (2013/4). Aarau 2013. Online: https://arbido.ch/assets/files/arbido_4_2013_low.pdf [Zugriff am 5. 4. 2018]
- Archer, Phil: HTTPS and the Semantic Web/Linked Data. 2016. Online: <https://www.w3.org/blog/2016/05/https-and-the-semantic-web-linked-data> [Zugriff am 16. 6. 2018].
- Ayers, Danny. & Völkel, Max: Cool URIs for the Semantic Web. 2008. On-line: <https://www.w3.org/TR/cooluris/> [Zugriff am 16. 6. 2018].
- Berners-Lee, Tim: Universal Resource Identifiers in WWW. 1994. Online: <https://tools.ietf.org/html/rfc1630> [Zugriff am 7. 6. 2018].
- Berners-Lee, Tim: Cool URIs don't change. 1998. Online: <https://www.w3.org/Provider/Style/URI> [Zugriff am 16. 6. 2018].
- Berners-Lee, Tim: Linked Data - Design Issues. 2006. Online: <https://www.w3.org/DesignIssues/LinkedData.html> [Zugriff am 16. 6. 2018].

- Berners-Lee, Tim: Tim Bernes-Lee on the next Web. 2009. Online: https://www.ted.com/talks/tim_berners_lee_on_the_next_web [Zugriff am 16. 6. 2018].
- Berners-Lee, Tim, Fielding, Roy & Masinter, Larray: Uniform Resource Identifier (URI): Generic Syntax. 2005. Online: <https://tools.ietf.org/html/rfc3986>: [Zugriff am 15. 6. 2018].
- Berners-Lee, Tim, Hendler, James & Lassila, Ora: The Semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. 2001. *Scientific American*, 284((5)), S. 34–43.
- Borst, Tim & Neubert, Joachim: Case Study: Publishing STW Thesaurus for Economics as Linked Open Data. 2009. Online: <http://www.w3.org/2001/sw/sweo/public/UseCases/ZBW/> [Zugriff am 18. 6. 2018].
- Brüning, Rainer, Heegewaldt, Werner, Brübac, Nils & Archivschule Marburg (Hg.): ISAD (G): Internationale Grundsätze für die archivische Verzeichnung. Marburg 2002. Online: https://www.ica.org/sites/default/files/CBPS_2000_Guidelines_ISAD%28G%29_Second-edition_DE.pdf [Zugriff am 20. 6. 2018].
- data.archiveshub.ac.uk, kein Datum <http://data.archiveshub.ac.uk/>. Online: Archives Hub Linked Data [Zugriff am 12. 6. 2018].
- de Melo, Gerard: Lexvo.org: Language-Related Information for the Linguistic Linked Data Cloud. 2015. Online: <http://www.semantic-web-journal.net/system/files/swj420.pdf> [Zugriff am 24. 6. 2018].
- Deutsche Nationalbibliothek (Hg.): Jahresbericht 2017. Leipzig / Frankfurt am Main 2018. Online: <https://d-nb.info/1160486344/34> [Zugriff am 12. 7. 2018].
- Docuteam GmbH (Hg.): Matterhorn RDF Datamodel (aus dem Internen Wiki), Baden-Dättwil 2018.
- Dodis (Hg.): Diplomatische Dokumente der Schweiz 1848-1975. Kein Datum. Online: <https://dodis.ch/search?q=&c=Place&t=all&cb=doc> [Zugriff am 12. 7. 2018].
- Dubois, Alain, Kansy, Lambert, Lu#thi, Martin & Wildi Tobias: IP new generation: eCH-IP-Hackathon 16.5.2018 (Unveröffentlicht). 2018.
- Eckert, Kai: Die Provenienz von Linked Data. In: Pohl, Adrian (Hg.) & Danowski, Patrick (Hg.). (Open) Linked Data in Bibliotheken. Berlin/ Boston 2013, S. 97-120.
- Estermann, Beat.: Schweizer Gedächtnisinstitutionen im Internet-Zeitalter: Ergebnisse einer Pilotbefragung zu den Themenbereichen Open Data und Crowdsourcing. Bern 2013.
- Estermann, Beat & Schneeberger, Christian: Data Model for the Swiss Performing Arts Platform. Bern 2017. Online: https://www.wirtschaft.bfh.ch/uploads/tx_frppublikationen/SPA_Data_Model_v0-51_20170926.pdf [Zugriff am 15. 7. 2018].
- Europeana (Hg.): Europeana Data Model – Mapping Guidelines v2. 4. 2017. Online: https://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Mapping_Guidelines_v2.4_102017.pdf [Zugriff am 10. 7. 2018].
- Europeana (Hg.): About. Kein Datum. Online: <https://www.europeana.eu/portal/de/about.html> [Zugriff am 12. 7. 2018].

- GeoNames.org (Hg.): About GeoNames. Kein Datum. Online: <http://www.geonames.org/about.html> [Zugriff am 12. 7. 2018].
- Geschäftsstelle E-Government (Hg.): Linked Data Service – LINDAS. Bern kein Datum. Online: <https://www.egovernment.ch/de/umsetzung/e-government-schweiz-2008-2015/lindas/> [Zugriff am 31. 5. 2018].
- Gonzenbach, Anouk Dunant: Archival linked open data – le projet suisse aLOD. Aarau 2017. Online: <https://arbido.ch/de/ausgaben-artikel/2017/metadaten-datenqualit%C3%A4t/archival-linked-open-data-le-projet-suisse-alod> [Zugriff am 21. 6. 2018].
- Gray, Jonathan: Response to ‘The Future of Bibliographic Control’ draft from the Library of Congress. 2007. Online: <http://blog.okfn.org/2007/12/19/response-to-the-future-of-bibliographic-control-draft-from-the-library-of-congress/> [Zugriff am 31. 5. 2018].
- Häusler, Jaqueline: 100 Jahre soziales Wissen: Schweizerisches Sozialarchiv 1906-2006. Zürich 2006. Online: https://e-monos.sozialarchiv.ch/HaeuslerJacqueline_100JahreSozialesWissen.pdf [Zugriff am 11. 07. 2018].
- Heath, Tom & Bizer, Chritian: Linked Data – Evolving the Web into a Global Data Space. San Rafael 2011.
- Hyönen, Eero: Publishing and Using Cultural Heritage Linked Data on the Semantic Web. San Rafael 2012.
- Hyland, Bernadette, Atemezing, G. & Villazón-Terrazas, B.: Best Practices for Publishing Linked Data. 2014, Online: <https://www.w3.org/TR/ld-bp/> [Zugriff am 16. 6. 2018].
- International Council on Archives (ICA) (Hg.): ISAD(G): General International Standard Archival Description, Second Edition. 2000. Online: https://www.ica.org/sites/default/files/CBPS_2000_Guidelines_ISAD%28G%29_Second-edition_EN.pdf [Zugriff am 31 05 2018].
- International Council on Archives (ICA) (Hg.): Records in Context: A conceptual model for archival description. 2016. Online: <https://www.ica.org/sites/default/files/RiC-CM-0.1.pdf> [Zugriff am 28. 5. 2018].
- J. Paul Getty Trust (Hg.): TGN: Frequently Asked Questions. Los Angeles 2017. Online: <http://www.getty.edu/research/tools/vocabularies/tgn/faq.html> [Zugriff am 12. 7. 2018].
- Jones, Ed & Seikel, Michele: Linked Data for cultural heritage. Chicago 2016.
- Klee, Carsten: Vokabulare für bibliographische Daten: Zwischen Dublin Core und bibliothekarischem Anspruch. In: (Open) Linked Data in Bibliotheken. Berlin/Boston 2013, S. 45-63.
- Koha Community (Hg.): Linked Data RFC. 2017. Online: https://wiki.koha-community.org/wiki/Linked_Data_RFC#Goals [Zugriff am 31. 6. 2018].
- Kompetenzzentrum in Digitalem Recht: Urheberschutz von Metadaten. Kein Datum. Online: <https://ccdigitallaw.ch/index.php?cID=977#metadaten> [Zugriff am 31 05 2018].
- labs.regesta.com: OAD Ontology. Kein Datum. Online: <https://labs.regesta.com/progettoReload/en/oad-ontology/> [Zugriff am 14. 7. 2018].

- Lexvo.org: About. Kein Datum. Online: <http://www.lexvo.org/> [Zugriff am 28. 6. 2018].
- Llanes-Padro#n, Dunia & Pastor-Sa#nchez, Juan-Antonio: Records in Contexts: the road of archives to semantic interoperability. 2017. Online: <http://eprints.rclis.org/31993/1/record-in-context-ontology-paper-revised.pdf> [Zugriff am 7. 20. 2018].
- lod-cloud.net: The linked Open Data Cloud. 2018. Online: <https://lodcloud.net> [Zugriff am 27. 6. 2018].
- Malmsten, Martin: LIBRIS available as Linked Data. 2008. Online: <http://librisbloggen.kb.se/2008/12/03/libris-available-as-linked-data/> [Zugriff am 31. 05.2018].
- Marden, Julia, Li-Madeo, Carolyn, Edelstein, Jegg & Whysel, Noreen: Linked Open Data for Cultural Heritage: Evolution of an Information Technology. 2013. On-line: <http://www.whysel.com/papers/LIS670-Linked-Open-Data-for-Cultural-Heritage.pdf> [Zugriff am 14. 7. 2018].
- Merzaghi, Michelle: Informationen finden und Wissen verlinken - Der Weg der Metadatenstandards vom Archivregal zu den Linked Data. Aarau 2017. Online: <https://arbido.ch/de/ausgaben-artikel/2017/metadaten-datenqualit%C3%A4t/vom-regalzum-word-wide-web-die-entwicklung-von-normen-und-standards> [Zugriff am 10. 6. 2018].
- Montiel-Ponsoda, Elena, Vila-Suero, Daniel & Villazo#n-Terrazas, Boris: Style Guidelines for Naming and Labeling Ontologies in the Multilingual Web. 2011. On-line: <http://dcevents.dublincore.org/index.php/IntConf/dc-2011/paper/download/47/15> [Zugriff am 17. 6. 2018].
- Open Knowledge Foundation Wiki: Response to Working Group on the Future of Bibliographic Control (Library of Congress). Kein Datum. Online: <http://wiki.okfn.org/FutureOfBibliographicControl> [Zugriff am 17. 7. 2018 (über Wayback Machine von archive.org)].
- Open Knowledge International: The Open Definition. Kein Datum. Online: <https://opendefinition.org/> [Zugriff am 17. 6. 2018-06].
- opendata.swiss: swissNAMES3D Geografische Namen der Landesvermessung. 2018. Online: <https://opendata.swiss/de/dataset/swissnames3d-geografische-namen-der-landesvermessung> [Zugriff am 12. 7. 2018].
- ortsnamen.ch: Datenbank der Schweizer Namenbücher. Kein Datum. Online: <https://www.ortsnamen.ch/index.php/datenbank-info.html> [Zugriff am 12. 7. 2018].
- Papakonstantinou, Vassilis et al.: Versioning for Linked Data: Archiving Systems and Benchmarks. 2016. Online: <http://ceur-ws.org/Vol-1700/paper-05.pdf> [Zugriff am 10. 7. 2018].
- Pohl, Adrian & Danowski, Patrick: (Open) Linked Data in Bibliotheken. Berlin/Boston 2013.
- Pohl, Adrian & Danowski, Patrick: Linked Open Data in der Bibliothekswelt: Grundlagen und Überblick. In: Pohl, Adrian & Danowski, Patrick (Hg.): (Open) Linked Data in Bibliotheken. Berlin/Boston 2013, S. 1-44.
- Prongu, Nicolas & Hu#gi, Jasmin, 2013. Les applications base#es sur les LOD en bibliothe#que: un tour d'horizon. arbido, 2013(4), Aarau 2013, S. 25-27.
- Prud'hommeaux, Eric: SPARQL 1.1 Query Language. 2013. Online: <https://www.w3.org/TR/sparql11-query/> [Zugriff am 16. 6. 2018].

- Prud'hommeaux, Eric & Seaborne, Andy: SPARQL Query Language for RDF. 2008. Online: <https://www.w3.org/TR/rdf-sparql-query/> [Zugriff am 16. 6. 2018].
- Rechtsquellenstiftung (Hg.): SSRQ Datenbank historischer Ortsnamen. 2018. Online: <https://www.ssrq-sds-fds.ch/places-db/search/search-form.xq> [Zugriff am 12. 7. 2018].
- Sakr, Sherik et al.: Linked Data: Storing Querying and Reasoning. Cham 2018.
- Sanders, Shlomo: Linked Library Data: It's Happening. 2017. Online: <http://www.exlibrisgroup.com/linked-library-data-its-happening/> [Zugriff am 17. 6. 2018].
- Schüpbach, Sebastian: Swissbib data goes linked 1: Transformation des métadonnées, modélisation, indexation. 2016. Online: <http://swissbib.blogspot.com/2016/04/swissbib-data-goes-linked-teil-1.html> [Zugriff am 17. 6. 2018].
- Schweizerische Nationalbibliothek. Open Data Strategie: Freigabe der Metadaten von „Helveticat“. Bern 2016. Online: <https://www.nb.admin.ch/snl/de/home/dienstleistungen/open-data-strategie--freigabe-der-metadaten-von-helveticat.html> [Zugriff am 10. 6. 2018].
- Schweizerisches Bundesarchiv: Kurzbeschrieb Linked Data Service – LIN- DAS (Unveröffentlicht). Bern kein Datum.
- Schweizerisches Sozialarchiv: Leitbild. Zürich 2018. Online: https://www.sozialarchiv.ch/wpcontent/uploads/fileadmin/user_upload/Sozialarchiv/Dokumente/PDFs/Sozialarchiv/leitbild.pdf [Zugriff am 11. 7. 2018].
- Schweizerisches Sozialarchiv: Nutzungsbestimmungen. Zürich 2018. Online: <https://www.sozialarchiv.ch/archiv/benutzung/nutzungsbestimmungen/> [Zugriff am 25. 6. 2018].
- Staatsarchiv des Kantons Zürich (Hg.): Erschliessungshandbuch (v2.3). Zürich 2018. Online: https://staatsarchiv.zh.ch/internet/justiz_innere/sta/de/ueber_uns/veroeffentlichungen/_jcr_content/contentPar/downloadlist_4/downloaditems/252_1491571647812.spooler.download.1524148408406.pdf/Erschliessungshand-buch_Version_2.3_extern_2018_03_27_bt.pdf [Zugriff am 20. 7. 2018].
- Stevenson, Adrian: LOCAH Project and Considerations of Linked Data Approaches. 2011 Online: <https://www.slideshare.net/adrianstevenson/locah-project-and-considerations-of-linked-data-approaches> [Zugriff am 28. 5. 2018].
- Swartz, Aaron: Aaron Swartz's A programmable Web: An unfinished work. San Rafael 2013.
- Tögel, Bettina: Archivische Normen und deren Umsetzung im Staatsarchiv Zu#rich (Referat im MAS ALIS 2017/18). Zürich 2018.
- The Basel Register of Thesauri, Ontologies & Classifications (BARTOC): Helvetosaurus. Basel kein Datum. Online: <https://bartoc.org/en/node/675> [Zugriff am 12. 7. 2018].
- Théreaux, Oliver: Common http implementation problems. 2003. Online: <http://www.w3.org/TR/chips/> [Zugriff am 28. 6. 2018].
- UK Cabinet Office: Designing URI Sets for the UK Public Sector. London 2009. Online: <https://assets.publishing.service.gov.uk/government/>

uploads/system/uploads/attachment_data/file/60975/designing-URI-sets-uk-public-sector.pdf [Zugriff am 3. 6. 2018].

Van Hooland, Seth & Verborgh, Ruben: Linked data for libraries, archives and museums: how to clean, link and publish your metadata. London 2014.

VSA: Protokoll der Sitzung der Arbeitsgruppe Normen und Standards vom 29. November 2017. 2017. Online: http://vsa-aas.ch/wp-content/uploads/2018/05/2017-11-29_Protokoll-AGNuS.pdf [Zugriff am 20. 7. 2018].

VSA: Linked Data zwischen Theorie und praktischer Anwendung. kein Datum. Online: <http://vsa-aas.ch/news/workshop-vom-21-februar-2018-linked-data-zwischen-theorie-und-praktischer-anwendung/> [Zugriff am 20. 7. 2018].

W3C, Lebo, T., Sahoo, S. & McGuinness, D.: PROV-O: The PROV Ontology. 2013. Online: <https://www.w3.org/TR/prov-o/> [Zugriff am 22. 7. 2018].

W3C: Ontologies. kein Datum. Online: <https://www.w3.org/standards/semanticweb/ontology> [Zugriff am 3. 6. 2018].

Wang, Richard Y. & Strong, Diana M.: Beyond Accuracy: What Data Quality Means to Data Consumers. In: *Journal of Management Information Systems*, Vol. 12 (No. 4). 1996. S. 5-33.

wiki.dnb.de: GNDCon 2018. 2018. Online: <https://wiki.dnb.de/display/GNDCON2018/GNDCon+2018> [Zugriff am 20. 7. 2018].

Wood, David, Zaidmann, Marsha & Luke, Ruth: Linked Data – Structured data on web. Shelter Island 2014.

Anhang Abkürzungsverzeichnis

AIS: Archivinformationssystem

CC: Creative Commons

ebd: ebenda

GND: Gemeinsame Normdatei

HTTP: Hypertext Transfer Protocol

HTTPS: HTTP Secure

ICA: International Council on Archives

ISAAR (CPF): International Standard Archival Authority Record for Corporate Bodies, Persons, and Families

ISAD(G): General International Standard Archival Description

ISDF: International Standard for Describing Functions

ISDIAH: International Standard for Describing Institutions with Archival Holdings

JSON: JavaScript Object Notation

JSON-LD: JSON for Linked Data

k. A.: Keine Angabe

LOCAH: Linked Open Copac and Archives Hub

LOD: Linked Open Data

OAIS: Open Archival Information System

RDF: Resource Description Framework

RiC: Records in Context

SPARQL: Protocol and RDF Query Language

Turtle: Terse RDF Triple Language

URI: Uniform Resource Identifiers

vgl.: vergleiche

VSA: Verein Schweizerischer Archivarinnen und Archivare

W3C: World Wide Web Consortium

XML: Extensible Markup Language

Anhang

Vokabular- und Ontologie-Prefix-Verzeichnis

Prefix	Beschreibung	Namespace URI
bibframe	The BIBFRAME vocabulary	http://bibframe.org/vocab/
bibo	The Bibliographic Ontology	http://purl.org/ontology/bibo/
bpmn	Business Process Model and Notation	http://dkm.fbk.eu/index.php/BPMN_Ontology#
cc	Creative Commons	https://creativecommons.org/ns
cidoc	CIDOC Conceptual Reference Model	http://www.cidoc-crm.org/cidoc-crm/
dc	Dublin Core	http://purl.org/dc/elements/1.1/
dcterms	Dublin Core Metadata Initiative (DCMI) Metadata Terms	http://purl.org/dc/terms/
ebucore	Ebucore	http://www.ebu.ch/metadata/ontologies/ebucore/ebucore#
evReiAgRole	LoC Event Relations Agent Role	http://id.loc.gov/vocabulary/preservation/linkingAgentRoleEvent/
evType	LoC Event Type	http://id.loc.gov/vocabulary/preservation/eventType/
foaf	The Friend of a Friend RDF vocabulary	http://xmlns.com/foaf/0.1
geo	WGS84 Geo Positioning	http://www.w3.org/2003/01/geo/wgs84_pos#
gnd	GND Ontology	http://d-nb.info/standards/elementset/gnd#
gsp	GeoSPARQL	http://www.opengis.net/ont/geosparql
lode	An ontology for Linking Open Descriptions of Events	http://linkedevents.org/ontology/
ore	The OAI ORE terms vocabulary	http://www.openarchives.org/ore/terms/

owl	The OWL (Web Ontology Language) 2 Schema vocabulary	http://www.w3.org/2002/07/owl#
premis	Premis	http://id.loc.gov/vocabulary/preservation
prov	The PROV Ontology	http://www.w3.org/TR/prov-o/
rdac	RDA Class hierarchies	http://rdaregistry.info/Elements/c/
rdaGr2	RDA Group 2 elements	http://rdvocab.info/ElementsGr2/
rdai	RDA Items	http://rdaregistry.info/Elements/i/
rdau	RDA Unconstrained properties	http://rdaregistry.info/Elements/u
rdf	The RDF Concepts Vocabulary	http://www.w3.org/1999/02/22-rdf-syntax-ns
rdfs	The RDF Schema vocabulary	http://www.w3.org/2000/01/rdf-schema#
relSubType	LoC Relations Sub Type LOC	http://id.loc.gov/vocabulary/preservation/relationships
schema	Schema.org	https://schema.org/
skos	Simple Knowledge Organization System	http://www.w3.org/2004/02/skos/core#
time	Time Ontology in OWL	http://www.w3.org/2006/time#
xsd	XML Schema	http://www.w3.org/2001/XMLSchema#

Verzeichniselemente und Linked-Data-Äquivalente

Die folgende Dokumentation der Verzeichniselemente des Papierarchivs und der Datenbank Bild+Ton wurde im Zuge dieser Arbeit erstellt.

Bezeichnung	ISAD-G	Linked Open Data	Bild + Ton Bestand	Bild + Ton Dokument	Papierarchiv Bestand	Papierarchiv Dossier
Signatur	1.1 Signatur(en)	dc:identifier		Importar 	Importar 	Importar 
Ursprüngliche Signatur	1.1 Signatur(en)					
Titel	1.2 Titel	dc:title	Importar 	Importar 	Importar 	Importar 
Entstehungszeitraum/ Laufzeit	1.3 Entstehungszeitraum/ Laufzeit	dc:terms.date, dc:doc.time hasBeginning, time hasEnd (=Periode)	Importar 	Importar 	Importar 	~
Aufnahmedatei - rum	1.3 Entstehungszeitraum/ Laufzeit	rdau:P60074 (has date of capture)		Importar 		
Sendungsda- tum	1.3 Entstehungszeitraum/ Laufzeit	rdau:P60073 (has date of publication)		Importar 		
Verzeichnungsstufe	1.4 Identifizierung der Verzeichnungsstufe	dc:type	~	~	Importar 	~
Bestand	1.4 Identifizierung der Verzeichnungsstufe	dc:terms.isPartOf		Importar 		~
Umfang	1.5 Umfang (Menge oder Ab- messung)	rdau:P60550 (has extent), rdau:P60134 (has note on extent of resource)			Importar 	Importar 
Spieldauer	1.5 Umfang (Menge oder Ab- messung)	rdau:P60557 (has duration)		Importar 		
Filesize	1.5 Umfang (Menge oder Ab- messung)	rdau:P60551 (has file size)		Importar 		
Objekträger	1.5 Umfang (Menge oder Ab- messung)			Importar 		
Name der Provenienzstelle	2.1 Name der Provenienzstelle	PROV-Agent	~	~	~	~
Verweilungsge- schichte/Bio- graphische Angaben	2.2 Verweilungs- geschichte/Bio- graphische Angaben	rdau:P60484 (has agent history)	Importar 		Importar 	
Bestandsge- schichte	2.3 Bestandsge- schichte	rdau:P60176 (has cus- tomal history of re- source)	Importar 		Importar 	
Übernahmemo- dalkisten	2.3 Bestandsge- schichte	PROV-Aktivität			Importar 	
Abgebende Stelle	2.4 Abgegebende Stelle	rdau:P60583 (has im- mediate source of ac- quisition of resource)	~		Importar 	
Form und In- halt	3.1 Form und In- halt	dc:description	Importar 	Importar 	Importar 	~
Thema, Helve- tosurus	3.1 Form und In- halt	rdau:P60805 (has sub- ject)	Importar 	Importar 	Importar 	
Abstract	3.1 Form und In- halt	rdau:P60375 (has summarization of con- tent)		Importar 		
GeoPolitik	3.1 Form und In- halt	schema:Place		Importar 		
GeoNatur	3.1 Form und In- halt	schema:Place		Importar 		
Bewertung und Kassation	3.2 Bewertung und Kassation	dc:terms accrualPolicy	Importar 		Importar 	
Neuzugänge	3.3 Neuzugänge	dc:terms accrual- Method	Importar 		Importar 	
Ordnung und Klassifikation	3.4 Ordnung und Klassifikation	rdau:P60348 (has system of organization)	~	Importar 	~	Importar 
Zugangs- bestimmungen	4.1 Zugangs-bestimmungen	premis hasRightsOran- tel (Access Rights)	Importar 	Importar 	Importar 	
Urheberrechte	4.1 Zugangs-bestimmungen	dc:terms license	Importar 	Importar 		
Urheber	4.1 Zugangs-bestimmungen	rdau:P60447 (has cre- ator)		Importar 		
Reproduktions- bestimmungen	4.2 Reprodukti- ons-bestimmun- gen	premis hasRights- Granted	~	~	~	~
Sprache/Schrift	4.3 Spra- che/Schrift	dc:terms language	Importar 	Importar 	Importar 	
Physische Beschaffenheit und technische Anforderungen	4.4 Physische Beschaffenheit und technische Anforderungen	rdau:P60528 (has equipment or system requirement)		Importar		
Zustand	4.4 Physische Beschaffenheit und technische Anforderungen			Importar		
Zustand Details	4.4 Physische Beschaffenheit und technische Anforderungen			Importar		
Findhilfsmittel	4.5 Findhilfsmittel	rdau:P60262 (is find- ing aid)			Importar	
Aufbewehrungsort der Originale	5.1 Aufbewehrungsort der Originale	dc:term isVersionOf		Importar	Importar	
Anzahl Original	5.1 Aufbewehrungsort der Originale			Importar		
Kopien bzw Reproduktionen	5.2 Kopien bzw Reproduktionen	rdau:P60272 (is repro- duced as)		Importar	Importar	
Verwandte Verzeichnungs- einheiten	5.3 Verwandte Verzeichnungs- einheiten	pcdm:hasRelatedOb- ject	Importar	Importar	Importar	
Veröffentli- chungen	5.4 Veröffentli- chungen	rdau:P60333 (has publication statement)			Importar	
Allgemeine An- merkungen	5.1 Allgemeine Anmerkungen	rdau:P60470 (has note on resource)		~	Importar	
Informationen des Bearbeiters	7.1 Informationen des Bearbeiters	PROV-O-Agent	Importar	Importar	Importar	
Datum oder Zeitraum der Verzeichnung	7.3 Datum oder Zeitraum der Verzeichnung	PROV-O-Aktivität	~	Importar	~	
Person		rdau:P40073 (relates an item to a person associated with an item being described)		Importar	~	~
Farbe	4.4 Physische Beschaffenheit und technische Anforderungen	rdau:P60761 (has details of colour content)		Importar		
Entity Status	2.2 Verweilungs- geschichte/Bio- graphische Angaben	Durch Körperschaften- Ressource			Importar	

Signatur

Zweck Eindeutige Identifikation der Verzeichnungseinheiten LOD
dc:identifier ISAD-G 1.1 Signatur(en), Reference code(s)

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung	Importar imagen	Importar imagen	Importar imagen	Importar imagen
Beispiel	F_5070	F_5070-Oa-054	Ar 201.138	Ar 201.138.2

Ursprüngliche Signatur

Zweck Bezeichnet allfällige Signierungen durch die abliefernde Körperschaft; wichtig, um dem ursprünglichen Ordnungszustand rekonstruieren zu können.

LOD

ISAD-G 1.1 Signatur(en), Reference code(s)

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung		Importar imagen		
Beispiel		F 106-26		

Titel

Zweck Benennung der Verzeichnungseinheit LOD dc:title
ISAD-G 1.2 Titel, Title

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung	Importar imagen	*	Importar imagen	Importar imagen
Beispiel	Aargauischer Gewerkschaftsbund (AGB)	Plakette "VHTL FCTA, die Gewerkschaft der Berufs- chauffeurs", um 1990	Arbeiterpartei Zü- rich III	Diverses

* Verwendung unter der Bezeichnung Haupttitel

Entstehungszeitraum/Laufzeit

Zweck Identifizierung des Entstehungsdatums bzw. -zeitraums der in der Verzeich- nungseinheit enthaltenen Unterlagen

Importar imagen Importar imagen LOD dcterms:date,
cidoc, time:hasBeginning, time:hasEnd (=Periode) ISAD-G 1.3
Entstehungszeitraum/Laufzeit, Dates of creation

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung	*	**	Importar imagen	***
Beispiel(e)	1971 1971-2017	Neuzeit; 20. Jh.; 1951-2000; 1981-1990; 1991 / Neuzeit; 20. Jh.; 1951-2000; 1981 -1990; 1993	1892- 1913	Abrechnungen und Rechnungsbelege 1897-1902

* Automatische Errechnung aus den Dokumenten. Verwendung unter der Bezeichnung Time Period.

** Verwendung unter der Bezeichnung Periode. Angaben von mehreren Daten möglich.

Aufnahmedatum

Zweck Bei Tonaufnahmen kann das Datum der Aufnahme angegeben werden. LOD rdau:P60074 (has date of capture) (=Aufnahme) ISAD-G 1.3 Entstehungszeitraum/Laufzeit, Dates of creation

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung		*		
Beispiel		10.12.2004 **		

* Nur bei Tonaufnahmen

** in Datenbank als Datumsfeld

Sendungsdatum

Zweck Bei Tonaufnahmen kann das Datum der Sendung angegeben werden. LOD rdau:P60073 (has date of publication) (=Sendung) ISAD-G 1.3 Entstehungszeitraum/Laufzeit, Dates of creation

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung		*		
Beispiel		10.12.2004		

* Nur bei Tonaufnahmen

Verzeichnungsstufe

Zweck Identifizierung der Verzeichnungsstufe. LOD dc:type ISAD-G 1.4 Identifizierung der Verzeichnungsstufe, Level of description

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung	~ *	~ *	Importar imagen	~ *
Beispiel	Bestand	Dokument	Bestand	Dossier

* Ergibt sich aus dem Kontext

Bestand

Zweck Gibt den zu einem Dokument zugehörigen Bestand wieder LOD dc:isPartOf ISAD-G 1.4 Identifizierung der Verzeichnungsstufe, Level of description

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung	-	~ *	-	~ *
Beispiel	-	F_5070		

* Ergibt sich aus dem Kontext

Umfang

Zweck Identifizierung des physischen Umfangs und der Art des Materials der Verzeichnungseinheit.

LOD rdau:P60550 (has extent), rdau:P60134 (has note on extent of resource) ISAD-G 1.5 Umfang (Menge oder Abmessung), Extent and medium of the unit

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung			Importar imagen	*
Beispiel			0.20 m	3 Mappen

* in Beschreibung von Dossier

Spieldauer

Zweck Identifizierung der Spieldauer eines Ton- und audiovisuellen Objekts. LOD rdau:P60557 (has duration) (=Spieldauer)

ISAD-G 1.5 Umfang (Menge oder Abmessung), Extent and medium of the unit

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung		Importar imagen		
Beispiel		00:32:28		

Filesize

Zweck Identifizierung der Dateigrösse eines Ton- und audiovisuellen Objekts. LOD rdau:P60551 (has file size) (=Filesize)

ISAD-G 1.5 Umfang (Menge oder Abmessung), Extent and medium of the unit

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung		Importar imagen		
Beispiel		1.404.898		

Objektträger

Zweck Auswahl des Trägers der Verzeichnungseinheit. LOD

ISAD-G 1.5 Umfang (Menge oder Abmessung), Extent and medium of the unit

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung		*		
Beispiel		Tonaufnahme; Magnetband; Kompaktkassette		

* Vorgegebene Auswahl

Name der Provenienzstelle

Zweck Identifizierung der Provenienzstelle(n), bei der (denen) die Verzeichnungs- einheit entstanden ist.

LOD prov:wasGeneratedBy ISAD-G 2.1 Name der Provenienzstelle, Name of creator

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung	~*	~*	~**	~
Beispiel	Radio Riesbach	F_5130 Frauen sehen Frauen	POCH Zürich	Einladungen, Pendenzenlisten, Projektbeschriebe Provenienz: Michael von Felten

* Die Provenienz ist im Bild+Ton-Bestand ebenfalls im Titel des Bestandes ausge- drückt

** Die Provenienz ist im Titel des Bestands ausgedrückt (UK)

Verwaltungsgeschichte/Biographische Angaben

Zweck Information über Verwaltungsgeschichte der Provenienzstelle bzw. die Bio- graphie, wenn es sich um natürliche Personen handelt, zum besseren Verständnis des zur Verzeichnungseinheit gehörenden Kontextes.

LOD rdau:P60484 (has agent history) ISAD-G 2.2
Verwaltungsgeschichte/Biographische Angaben, Administrative/
Biographical history

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung	Importar imagen		Importar imagen	
Beispiel	Im Vorfeld der Nationalratswah- len 1971 konnten sich alle gros- sen Parteien mit mehrminütigen Spots oder Kurzfilmen zur bes- ten Sendezeit am Schweizer Fernsehen präsentieren...		Das Frauen/Les- ben-Archiv (bis 1990 Frauenar- chiv) wurde 1985 gegründet...	

Bestandsgeschichte

Zweck Information über den Wechsel der Eigentums- und Besitzverhältnisse der Verzeichnungseinheit, die für deren Authentizität, Integrität, Vollständigkeit und Interpretation von wesentlicher Bedeutung sind.

LOD rdau:P60176 (has custodial history of resource) ISAD-G 2.3
Bestandsgeschichte, Archival history

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung	*		Importar imagen	
Beispiel	Tel. mit Mischa Brutschin, 3.3.2014: Er hat den Bestand von einem Radiopiraten übernommen und durfte ihn dem Sozialarchiv geben. Er fragt noch, ob die Person zu einem Gespräch bereit ist. 20160519: Adrian Scherrer hat via Rudolf Müller, Memo-riav, die offenbar allererste Kassette der Wellenhexen (damals noch unter dem Na- men "101" erhalten.		Die Akten der PdAZ wurden vor der Über- nahme durch das Schweizerische Sozial- archiv im Sekretariat der PdAZ an der Rot- wandstrasse 65 in Zü- rich aufbewahrt. Die Dossiers SOZARCH Ar 458.80.1-Ar 458.80.29 bildeten einen Teil des sogenannten KPS- /PdA-Archivs, das von der ehemaligen Histori- schen ...	

* Verwendung unter der Bezeichnung Custodial History

Übernahmемodalitäten

Zweck Information über die Übernahmемodalitäten. LOD PROV- Aktivität ISAD-G 2.3 Bestandsgeschichte, Archival history

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung			Importar imagen	
Beispiel			Das Archiv der Partei der Arbeit des Kan- tons Zürich gelangte am 22.05.2007 und am 21.06.2007 ins Schweizerische Sozi- alarchiv. Die Übergabe wurde von Manf- red Vischer und Daniel Brunner betreut.	

* Verwendung unter der Bezeichnung Custodial History

Abgebende Stelle

Zweck Darstellung der Umstände, die mit der direkten Übernahme der Verzeich- nungseinheit von der abgebenden Stelle verbunden sind.

LOD rdau:P60583 (has immediate source of acquisition of resource)
ISAD-G 2.4 Abgebende Stelle, Immediate source of acquisition

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung	~*		Importar imagen	
Beispiel	2017: Nachlieferung erhalten von Rita Lanz (erschlossen von Leonie Schmid)		Zentralsekretariat unia, Monbi- joustrasse 61, 3000 Bern	

* teilweise ausgedrückt im Feld Custodial History

Form und Inhalt

Zweck Feststellung von Hauptgegenstand und Form der Verzeichnungseinheit, um Benutzern eine Beurteilung ihrer Relevanz zu ermöglichen.

LOD dc:description (=Beschreibung) ISAD-G 3.1 Form und Inhalt, Scope and Content

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung	*	*, **	Importar imagen	~
Beispiel	Umfangreichster Bildbestand der drei Unia-Vorgängerge- werkschaften: über 3'400 Fotos, Objekte und Drucke. Heraus- ragend sind sicher....	[Track 01] 00:00: 00 Wei- tere Diskussion zur Forderung der Nichtverwen- dung von Gum- migeschossen...	Der Bestand enthält: Biographisches (Kindheit, Jugend, Blaues Kreuz, Fami- lie, Schule, Studium, Beruf); Korrespon- denz, ca. 1945-1999; Unterlagen zu...	Unterlagen betr. Jüng- lingsbund St. Johann vom Blauen Kreuz: BG-Stunden- Vorbereitun- gen 1946- 1955;...

* Verwendung unter der Bezeichnung Beschreibung

** nur bei Ton-, Video- und Filmdokumenten

Thema, Helvetosaurus

Zweck Auswahl von übergeordneten Themen, welche Bestand zugeteilt werden können. Der Helvetosaurus dient der thematischen Erschliessung des Dokumentin- halts.

LOD rdau:P60805 (has subject) (=Thema, Thema Hope, Helvetosaurus) ISAD-G 3.1 Form und Inhalt, Scope and Content

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung	*	*	*	
Beispiel	Gewerkschaften (B&T Datenbank) Anarchist (HOPE)	Leben in der Gesell- schaft (allgemein); Freizeit	Gewerkschaften	

* Vorgegebene Auswahl

Abstract

Zweck Kurze Zusammenfassung des Inhalts, die auch Angaben zum Genre des Films/Videos/der Tonaufnahme und Wertungen enthalten kann.

LOD rdau:P60375 (has summarization of content) (=Abstract)
ISAD-G 3.1 Form und Inhalt, Scope and Content

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung		*		
Beispiel		Ein Einwohner eines Dorfes erzählt wie der Wassermangel ...		

* nur bei Ton-, Video- und Filmdokumenten

GeoPolitik

Zweck Ein geografisches Schlagwort wird durch GeoPolitik ausgedrückt.
LOD schema:Place oder cidoc

ISAD-G 3.1 Form und Inhalt, Scope and Content

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung		*		
Beispiel		Europa; Schweiz		

* Vorgegebene Auswahl

GeoNatur

Zweck Bei stark landschaftlich ausgerichteten Beständen, wird GeoNatur verwendet.

LOD schema:Place oder cidoc ISAD-G 3.1 Form und Inhalt, Scope and Content

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung		*		
Beispiel		Gletscher; Aletschgletscher, Seen; Genfersee		

* Vorgegebene Auswahl

Bewertung und Kassation

Zweck Bereitstellung von Informationen über jede vorgenommene Bewertung und Kassation.

LOD ISAD-G 3.2 Bewertung und Kassation, Appraisal, destruction and scheduling

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung	*		Importar imagen	
Beispiel	20151104: Von 55 ursprünglich abgelieferten Einheiten wurden 44 kassiert (hauptsächlich Aufnahmen öffentlich. Radio - Abklärung mit Archiv SRF erfolgt). Siehe Kassationsliste (word).		Kassiert wurden Mehrfach-exemplare und Buchhaltungsbefuge.	

* Verwendung unter der Bezeichnung Kassationen

Neuzugänge

Zweck Angaben für den Benutzer über mögliche Veränderungen im Umfang der Verzeichnungseinheit.

LOD dterms:accrualMethod ISAD-G 3.3 Neuzugänge, Accruals

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung			Importar imagen	
Beispiel			Neuzugänge werden nicht erwartet.	

Ordnung und Klassifikation

Zweck Bereitstellung von Informationen über die Ordnung und Klassifikation der Verzeichnungseinheit.

LOD rdau:P60348 (has system of organization) ISAD-G 3.4 Ordnung und Klassifikation, System of arrangement

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung	~*	Importar imagen	~**	Importar imagen
Beispiel		F_1000 Vollver- sammlungen Ju- gendbewegung Zü- rich [TON]	Das Russlandschweizer- Archiv wurde an der Osteuropa-Abtei- lung des Historischen Seminars der Universität Zürich nach Do- kumentart geordnet	Kindheit und Ju- gend

* vereinzelt ausgedrückt im Feld Custodial History

** vereinzelt ausgedrückt im Feld Form und Inhalt

Zugangsbestimmungen

Zweck Angabe derjenigen Bestimmungen, die den Zugang zur Verzeichnungseinheit einschränken oder beeinflussen.

LOD premis:hasRightsGranted (=Bestand, Access Rights) ISAD-G 4.1 Zugangsbestimmungen, Conditions governing access

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung	Importar imagen	Importar imagen *	Importar imagen	
Beispiel	keine Benut- zungsbe- schränkungen	abspielbar	Für den Zugang muss die Bewilli- gung bei der Familie Braunschweig eingeholt werden. Die Aktenserien Private Korrespondenz und Amtsvor- mundschaft bleiben bis auf weiteres gesperrt.	

* Vorgegebene Auswahl

Urheberrechte

Zweck Angabe der Urheberrechte LOD dcterms:license

ISAD-G 4.1 Zugangsbestimmungen, Conditions governing access

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung		Importar imagen		
Beispiel		Andres, Peter		

Urheber

Zweck Angabe des Urhebers LOD rdau:P60447 (has creator)
ISAD-G 4.1 Zugangsbestimmungen, Conditions governing access

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung		Importar imagen		
Beispiel		Unbekannt		

Reproduktionsbestimmungen

Zweck Information über Beschränkungen bei der Reproduktion der Verzeichnungs- einheit.

LOD premis:hasRightsGranted ISAD-G 4.2
Reproduktionsbestimmungen, Conditions governing reproduction

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung	~*	~*	~*	~*
Beispiel				

* Festgelegt in den Nutzungsbestimmungen des Schweizerischen Sozialarchivs

Sprache/Schrift

Zweck Identifizierung der in der Verzeichnungseinheit enthaltenen Sprache(n), Schriftarten und Zeichensysteme.

LOD dcterms:language ISAD-G 4.3 Sprache/Schrift, Language/ scripts of material

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung	*	Importar imagen	Importar imagen	
Beispiel	ita,gsw,ger,fre	gsw	Unterlagen grösstenteils in deutscher Sprache.	

* Verwendung unter der Bezeichnung Languages of Items

Physische Beschaffenheit und technische Anforderungen

Zweck Bereitstellung von Informationen über wichtige physische Besonderheiten oder der technische Anforderungen, die die Benutzung der Verzeichnungseinheit beeinflussen.

LOD rdau:P60528 (has equipment or system requirement) ISAD-G 4.4 Physische Beschaffenheit und technische Anforderungen, Physical characteristics

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung			Importar imagen	
Beispiel				

Zustand

Zweck Beschreibung des materiellen Zustands der Verzeichnungseinheit.
LOD ISAD-G 4.4 Physische Beschaffenheit und technische Anforderungen, Physical characteristics

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung		*		
Beispiel		beschädigt		

* Vorgegebene Auswahl

Zustand Details

Zweck Beschreibung des materiellen Zustands der Verzeichnungseinheit.
LOD ISAD-G 4.4 Physische Beschaffenheit und technische Anforderungen, Physical characteristics

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung		Importar imagen		
Beispiel		Riss in Glasscheibe		

Findhilfsmittel

Zweck Identifizierung aller für die Verzeichnungseinheit vorhandenen Findhilfsmittel.

LOD rdau:P60262 (is finding aid) ISAD-G 4.5 Findhilfsmittel, Finding aids

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung			Importar imagen	
Beispiel			Im Schweizerischen Sozialarchiv ist ein detailliertes Findmittel zum Bestand SO- ZARCH Ar 201.240 (ELPOS Zürich, Elternverein für Kinder mit leichten psychischen Funktionsstörungen) vorhanden. Dieses kann nach Rücksprache mit der Leitung des Schweizerischen Sozialarchivs eingesehen werden.	

Aufbewahrungsort der Originale

Zweck Nachweise über die aufbewahrende Institution, die Zugänglichkeit oder die Vernichtung der Originale, falls es sich bei der Verzeichnungseinheit um eine Reproduktion handelt.

LOD dterm:isVersionOf ISAD-G 5.1 Aufbewahrungsort der Originale, Existence and location of originals

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung		*	Importar imagen	
Beispiel		Schweizer Radio und Fernsehen (SRF)	Die Originale befinden sich in der Regel im Besitz der Autorinnen und Autoren.	

* Verwendung unter der Bezeichnung Original Standort, nur bei Ton-, Video- und Filmdokumenten

Anzahl Original

Zweck Anzahl der vorhandenen Verzeichnungseinheiten LOD

ISAD-G 5.1 Aufbewahrungsort der Originale, Existence and location of originals

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung		Importar imagen		
Beispiel		1		

Kopien bzw. Reproduktionen

Zweck Verweis auf Kopien bzw. Reproduktionen der Verzeichnungseinheit und ihre Verfügbarkeit.

LOD rdau:P60272 (is reproduced as) ISAD-G 5.2 Kopien bzw. Reproduktionen, Existence and location of copies

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung		*	Importar imagen	
Beispiel		Lichtspiel Bern	Die Masterkopien im Format Beta Digital werden im Schweizerischen Bundesarchiv in Bern aufbewahrt; das Bundesarchiv verfügt wie das Schweizerische Sozialarchiv ebenfalls über Visionierungskopien	

* Verwendung unter der Bezeichnung Masterkopie Standort, nur bei Ton-, Video- und Filmdokumenten

Verwandte Verzeichnungseinheiten

Zweck Ermittlung von verwandten Verzeichnungseinheiten im selben Archiv oder in anderen Archiven.

LOD pcdm:hasRelatedObject ISAD-G 5.3 Verwandte Verzeichnungseinheiten, Related units of description

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung	*	(*)	Importar imagen	
Beispiel	ArSMUV http://...	http://findmittel.ch/archive/arch-Neu/ArVid.html	SOZARCH Ar 477: Demokratische Juristinnen und Juristen der Schweiz DJS, Juristes Démocrates de Suisse JDS, Giuristi e Giuriste Democratici Svizzeri . Jahresberichte Amnesty International, Schweizer Sektion: - SozArch K 747: Tätigkeitsbericht / Amnesty International, Schweizer Sektion. - Bern : Amnesty International [1996-]	

* Verwendung unter der Bezeichnung (weiterer) Archivbezug, siehe auch, Link Europeana, Link zu Memobase+

Veröffentlichungen

Zweck Ermittlung von Veröffentlichungen, die unter Benutzung oder Auswertung der Verzeichnungseinheit entstanden sind.

LOD rdau:P60333 (has publication statement) ISAD-G 5.4
Veröffentlichungen, Publication note

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung			Importar imagen	
Beispiel			50 Jahre Amnesty International - 1961-2011 [Hrsg.: Amnesty International - Sek- tion Schweiz; Red.: Jessica Cuerq ... et al.] Spezialausgabe von Amnesty...	

Allgemeine Anmerkungen

Zweck Bereitstellung von Spezialinformationen und Angaben, die in keinem der anderen Bereiche angebracht werden können.

LOD rdau:P60470 (has note on resource) ISAD-G 6.1 Allgemeine Anmerkungen, Note

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung	~*?	~*	Importar imagen	
Beispiel		[Freitextfeld für ergän- zende Infor- mationen.]	Der Bestand enthält diverse Überfor- mate. Aus diesem Grund werden die Unterlagen in einem Plakatschrank ("Sicherheitsschrank") aufbewahrt.	

* Verwendung unter der Bezeichnung Bemerkungen, Beschreibungen

Informationen des Bearbeiters

Zweck Erläuterungen zur Verzeichnung und über den oder die Bearbeiter. LOD PROV-O-Agent

ISAD-G 7.1 Informationen des Bearbeiters, Archivist's note

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung	*	**	Importar imagen	
Beispiel		[Freitextfeld für ergänzende Informationen.]	Die Erschliessung und Verzeichnung der ersten Aktenlieferung von 1972 erfolgte durch Karl Lang im Jahr 1978.	

* Verwendung unter der Bezeichnung VerzeichnerIn

** Verwendung unter der Bezeichnung IAModificationUser

Datum oder Zeitraum der Verzeichnung

Zweck Angabe von Datum oder Zeitraum der Verzeichnung und ggf. einer späteren Überarbeitung.

LOD PROV-O-Aktivität ISAD-G 7.3 Datum oder Zeitraum der Verzeichnung, date of descriptions

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung	~**	**	~*	
Beispiel	Erika Mustermann (2004, 2010)	15.01.2018 16:03:28	Bearbeitet im Winter 2002/2003 von Erika Mustermann.	

* Wird bereits im Feld VerzeichnerIn bzw. Informationen des Bearbeiters ausgedrückt.

** Verwendung unter der Bezeichnung IAModificationDate

Personen

Zweck Einstufige alphabetisierte Liste, die ständig ergänzt und mutiert wird. Einen Eintrag erhalten: abgebildete und identifizierbare Personen, Personen auf Grabsteinen, sinnbildlich gemeinte Personen, in Schriftzeichen erwähnte Personen

LOD rdai:P40073 (Relates an item to a person associated with an item being described.) ISAD-G -

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung		Importar imagen		
Beispiel		Hüsch, Oskar (1892-1979)		

Farbe

Zweck Unterscheidung von s-w und farbigen Verzeichnungseinheiten.
 LOD rdau:P60761 (has details of colour content) ISAD-G 4.4 Physische Beschaffenheit und technische Anforderungen, Physical characteristics

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung		Importar imagen		
Beispiele		Farbe Schwarzweiss s-w		

Entity Status

Zweck Gibt an ob die Abliefernde Organisation noch besteht LOD
 Durch Körperschaften-Ressource
 ISAD-G 2.2 Verwaltungsgeschichte/Biographische Angaben,
 Administrative/Biographical history

	Bild+Ton		Papierarchiv	
	Bestand	Dokument	Bestand	Dossier
Verwendung			Importar imagen	
Beispiele			Ja Nein ??	

Eintrag des Sozialarchiv in der Europeana

```
<?xml version="1.0" encoding="UTF-8"
standalone="yes"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-
syntax-ns#" xmlns:edm="http://www.europeana.eu/sche-
mas/edm/">
  <edm:ProvidedCHO rdf:about="http://data.euro-
peana.eu/item/2022081/10848_FFEB2153_DCBF_4C8E_9DCD_8C19
777EAE47"/>
    <edm:WebResource rdf:about="http://hdl.han-
dle.net/10848/3F57CF59-431A-4D05-B103-
6E308C3C3BD7?locatt=view:level2"/>
      <edm:rights rdf:resource="http://rightsstate-
ments.org/vocab/Inc/1.0/" />
    </edm:WebResource>
    <edm:WebResource rdf:about="http://hdl.han-
dle.net/10848/1F4FEC36-4A01-44F8-BB46-631E33438E2A#1"/>
      <ore:Aggregation xmlns:ore="http://www.openar-
chives.org/ore/terms/"
        rdf:about="http://data.euro-
peana.eu/aggregation/pro-
vider/2022081/10848_FFEB2153_DCBF_4C8E_9DCD_8C19777EAE47
">
        <edm:aggregatedCHO
          rdf:resource="http://data.euro-
peana.eu/item/2022081/10848_FFEB2153_DCBF_4C8E_9DCD_8C19
777EAE47"/>
        <edm:dataProvider>Schweizerisches Sozialar-
chiv</edm:dataProvider>
        <edm:isShownAt rdf:resource="http://hdl.han-
dle.net/10848/1F4FEC36-4A01-44F8-BB46-631E33438E2A#1"/>
        <edm:isShownBy
          rdf:resource="http://hdl.han-
dle.net/10848/3F57CF59-431A-4D05-B103-
6E308C3C3BD7?locatt=view:level2"/>
        <edm:object rdf:resource="http://hdl.han-
dle.net/10848/3F57CF59-431A-4D05-B103-
6E308C3C3BD7?locatt=view:level2"/>
        <edm:provider>HOPE - Heritage of the People's
Europe</edm:provider>
        <edm:rights rdf:resource="http://rightsstate-
ments.org/vocab/Inc/1.0/" />
        </ore:Aggregation>
      <ore:Proxy xmlns:ore="http://www.openar-
chives.org/ore/terms/"
        rdf:about="http://data.euro-
peana.eu/proxy/pro-
vider/2022081/10848_FFEB2153_DCBF_4C8E_9DCD_8C19777EAE47
">
      </ore:Proxy>
    </edm:WebResource>
  </edm:ProvidedCHO>
  <dc:creator xmlns:dc="http://purl.org/dc/ele-
ments/1.1/">Atelier Eidenbenz: Basel</dc:creator>
  <dc:description xmlns:dc="http://purl.org/dc/ele-
ments/1.1/" xml:lang="deu">Wirtschaft; Industrie; Maschi-
nen- und
  Metallindustrie; Maschine; K hlanlage
  </dc:description>
  <dc:description xmlns:dc="http://purl.org/dc/ele-
ments/1.1/" xml:lang="deu">Landwirtschaft
  (allgemein); landwirtschaftliche
  Produkte; Nahrungsmittel; Lebensmittelindustrie; Fleis-
  chindustrie
  </dc:description>
  <dc:description xmlns:dc="http://purl.org/dc/ele-
ments/1.1/" xml:lang="deu">Landwirtschaft
  (allgemein); landwirtschaftliche
  Produkte; Nahrungsmittel; zubereitetes Lebensmit-
  tel; Fleischerzeugnis
  </dc:description>
  <dc:format xmlns:dc="http://purl.org/dc/ele-
ments/1.1/">intakt</dc:format>
  <dc:format xmlns:dc="http://purl.org/dc/ele-
ments/1.1/">intakt</dc:format>
  <dc:identifier xmlns:dc="http://purl.org/dc/ele-
ments/1.1/">
    http://hdl.handle.net/10848/FFEB2153-DCBF-
    4C8E-9DCD-8C19777EAE47
  </dc:identifier>
  <dc:identifier>
  </dc:identifier>
</rdf:RDF>
```

```

xmlns:dc="http://purl.org/dc/ele-
ments/1.1/">Socarch_F_5030-Pb-0062</dc:identifier>
<dc:relation xmlns:dc="http://purl.org/dc/ele-
ments/1.1/"
    rdf:resource="http://www.peo-
plesheritage.eu"></dc:relation>
<dc:subject xmlns:dc="http://purl.org/dc/ele-
ments/1.1/">Kuhlanlage</dc:subject>
<dc:subject xmlns:dc="http://purl.org/dc/ele-
ments/1.1/">Fleischindustrie</dc:subject>
<dc:subject xmlns:dc="http://purl.org/dc/ele-
ments/1.1/">Fleischerzeugnis</dc:subject>
<dc:subject xmlns:dc="http://purl.org/dc/ele-
ments/1.1/">Schweiz</dc:subject>
<dc:title xmlns:dc="http://purl.org/dc/ele-
ments/1.1/" xml:lang="deu">Metzger mit Wursthagen im
Kuhraum
</dc:title>
<dc:type xmlns:dc="http://purl.org/dc/ele-
ments/1.1/">item</dc:type>
<dc:type xmlns:dc="http://purl.org/dc/ele-
ments/1.1/">Fotografie</dc:type>
<dc:terms:extent
xmlns:dc:terms="http://purl.org/dc/terms/">&lt;=130x185</
dc:terms:extent>
<dc:terms:isPartOf
xmlns:dc:terms="http://purl.org/dc/terms/"
    rdf:resource="http://hdl.han-
dle.net/10848/59C825DC-2983-4919-9E28-
B8510CF050F2"></dc:terms:isPartOf>
<dc:terms:medium
xmlns:dc:terms="http://purl.org/dc/terms/">Foto-
grafie;Positiv;Papierabzug</dc:terms:medium>
<dc:terms:provenance
xmlns:dc:terms="http://purl.org/dc/terms/">Schweizer-
isches Sozialarchiv</dc:terms:provenance>
<dc:terms:spatial
xmlns:dc:terms="http://purl.org/dc/terms/">Eu-
ropa;Schweiz</dc:terms:spatial>
<dc:terms:spatial
xmlns:dc:terms="http://purl.org/dc/terms/">CH</dc:terms:sp
atial>
<dc:terms:temporal
xmlns:dc:terms="http://purl.org/dc/terms/">1951-
1960</dc:terms:temporal>
<dc:terms:temporal
xmlns:dc:terms="http://purl.org/dc/terms/">1961-
1970</dc:terms:temporal>
<dc:terms:temporal
xmlns:dc:terms="http://purl.org/dc/terms/">1971-
1980</dc:terms:temporal>
<edm:europaeaProxy>false</edm:europaeaProxy>
<core:proxyFor rdf:resource="http://data.euro-
peana.eu/item/2022081/10848_FFB2153_DCBF_4C8E_9DCD_8C19
777EAE47"/>
<core:proxyIn
    rdf:resource="http://data.euro-
peana.eu/aggregation/pro-
vider/2022081/10848_FFB2153_DCBF_4C8E_9DCD_8C19777EAE47
"/>
<edm:type>IMAGE</edm:type>
</core:Proxy>
<core:Proxy xmlns:core="http://www.openar-
rdf:about="http://data.euro-
peana.eu/proxy/euro-
peana/2022081/10848_FFB2153_DCBF_4C8E_9DCD_8C19777EAE47
">
<edm:europaeaProxy>true</edm:europaeaProxy>
<core:proxyFor rdf:resource="http://data.euro-
peana.eu/item/2022081/10848_FFB2153_DCBF_4C8E_9DCD_8C19
777EAE47"/>
<core:proxyIn
    rdf:resource="http://data.euro-
peana.eu/aggregation/euro-
peana/2022081/10848_FFB2153_DCBF_4C8E_9DCD_8C19777EAE47
"/>
<edm:type>IMAGE</edm:type>

```

```

</ore:Proxy>
<ore:Proxy xmlns:ore="http://www.openar-
chives.org/ore/terms/"
  rdf:about="http://data.euro-
peana.eu/proxy/euro-
peana/2022081/10848 FFEB2153 DCBF 4C8E 9DCD 8C19777EAE47
">
  <edm:europenaProxy>true</edm:europenaProxy>
  <ore:proxyFor rdf:resource="http://data.euro-
peana.eu/item/2022081/10848 FFEB2153 DCBF 4C8E 9DCD 8C19
777EAE47"/>
  <ore:proxyIn
    rdf:resource="http://data.euro-
peana.eu/aggregation/euro-
peana/2022081/10848 FFEB2153 DCBF 4C8E 9DCD 8C19777EAE47
"/>
  <edm:type>IMAGE</edm:type>
</ore:Proxy>
<edm:EuropeanaAggregation
  rdf:about="http://data.europeana.eu/aggrega-
tion/euro-
peana/2022081/10848 FFEB2153 DCBF 4C8E 9DCD 8C19777EAE47
">
  <dc:creator xmlns:dc="http://purl.org/dc/ele-
ments/1.1/">Europeana</dc:creator>
  <edm:aggregatedCHO
    rdf:resource="http://data.euro-
peana.eu/item/2022081/10848 FFEB2153 DCBF 4C8E 9DCD 8C19
777EAE47"/>
  <edm:dataset-
Name>2022081 Ag EU HOPE SSASlot06</edm:datasetName>
  <edm:country>Switzerland</edm:country>
  <edm:preview
    rdf:resource="http://euro-
peanastatic.eu/api/image?uri=http%3A%2F%2Fhdl.han-
dle.net%2F10848%2F3F57CF59-431A-4D05-B103-
6E308C3C3BD7%3Flocatt%3Dview%3Alevel2&amp;size=LARGE&amp
;type=TEXT"/>
  <edm:landingPage
    rdf:resource="http://www.euro-
peana.eu/portal/rec-
ord/2022081/10848 FFEB2153 DCBF 4C8E 9DCD 8C19777EAE47.h
tml"/>
  <edm:language>mul</edm:language>
  <edm:rights rdf:resource="http://rightsstate-
ments.org/vocab/InC/1.0/" />
</edm:EuropeanaAggregation>
</rdf:RDF>

```

Fußnote

1. Gute Übersichtswerke sind zum Beispiel: Sakr, et al., 2018, Wood, et al., 2014, Swartz, 2013 und Heath & Bizer, 2011.
2. Zum Beispiel: Jones & Seikel, 2016, Van Hooland & Verborgh, 2014 und Hyönen, 2012
3. Pohl & Danowski, 2013.
4. Arbido, 2013.
5. Van Hooland & Verborgh, 2014, S. 44.
6. Das W3C ist ein Gremium zur Standardisierung von Techniken im World Wide Web.
7. Berners-Lee, et al., 2001.
8. Ebd., S. 41.
9. Pohl & Danowski, 2013, S. 5.

10. Berners-Lee, 2006. Zitiert aus Pohl & Danowski, 2013, S. 5.
- 11Ebd.
12. Pohl & Danowski, 2013, S. 5f.
- 13 lod-cloud.net, 2018.
- 14 lod-cloud.net, 2018.
- 15.Open Knowledge International, kein Datum.
- 16.Eine Auflistung von Lizenzen, die der Open-Definition genügen, finden sich unter: <https://opendefinition.org/licenses/>.
17. Pohl & Danowski, 2013, S. 10 und Berners-Lee, 2009.
18. Berners-Lee, 2006.
19. Pohl & Danowski, 2013, S. 22.
- 20.Dort wird das Akronym URI noch mit „Universal Resource Identifier“ aufgelöst. Vgl. Berners-Lee, 1994.
- 21.Berners-Lee, et al., 2005.
- 23.Siehe http://dbpedia.org/resource/Albert_Einstein.
- 24.Online verfügbar unter <http://purl.org/dc/terms/>.
- 25.W3C, kein Datum.
26. Heath & Bizer, 2011, S. 61f. und Wood, et al., 2014, S. 39-41.
27. Prud'hommeaux & Seaborne, 2008.
- 28Ebd.
- 29.Ebd., Prud'hommeaux, 2013.
- 30.Die Abfrage wurden den Beispielen auf <https://query.wikidata.org/> entnommen.
- 31.Hyland, et al., 2014.
32. Wang & Strong, 1996, S. 9.
- 33.Viele Archivalien weisen Sperrfristen auf. Gründe dafür sind unter anderem der Datenschutz und die Geheimhaltungspflicht. Manchmal betrifft diese Sperrung auch Teile der Metadaten. So kann es bereits problematisch sein, wenn der Name einer Person im Titel eines Archivals auftaucht.
- 34.Kompetenzzentrum in Digitalem Recht, kein Datum.
- 35.Hyland, et al., 2014.
36. Heath & Bizer, 2011, S. 62f.
- 37.Hyland, et al., 2014.
- 38.Hyland, et al., 2014.
39. Archer, 2016.
40. Ayers & Völkel, 2008.
- 41.Siehe Ayers & Völkel, 2008 für alternative Implementierung.

42. Ayers & Völkel, 2008.
44. Berners-Lee, 1998, zitiert aus Montiel-Ponsoda, et al., 2011, S. 108.
45. Théreaux, 2003.
46. UK Cabinet Office, 2009, zitiert aus Montiel-Ponsoda, et al., 2011, S. 108.
47. Ebd. S. 6, zitiert aus Montiel-Ponsoda, et al., 2011, S. 108.
48. Berners-Lee, 1998, zitiert aus Montiel-Ponsoda, et al., 2011, S. 108.
49. Ayers & Völkel, 2008.
50. Eckert, 2013, S. 97.
51. W3C, et al., 2013.
52. Eckert, 2013, S. 107.
53. Vgl. Papakonstantinou, et al., 2016.
54. data.archiveshub.ac.uk, kein Datum.
55. Ayers & Völkel, 2008.
56. Ebd.
57. Online abrufbar: <http://www.gahetna.nl/en/about-us/open-data>.
58. Siehe Kapitel 0.
59. Online abrufbar: <http://dati.acs.beniculturali.it/>.
60. Online abrufbar: <https://data.gov.uk/publisher/the-national-archives>.
61. Siehe Kapitel 0.
62. Online abrufbar: <https://archiveshub.jisc.ac.uk>.
63. Siehe Kapitel 0.
64. Marden, et al., 2013, S. 3.
65. Estermann, 2013, S. 14f.
66. Pohl & Danowski, 2013, S. 12.
67. Open Knowledge Foundation Wiki, kein Datum & Gray, 2007. Zitiert aus: Pohl & Danowski, 2013, S. 12.
68. Pohl & Danowski, 2013, S. 12.
69. Pohl & Danowski, 2013, S. 13f.
70. Open Data Commons Attribution License (ODC-By), vgl. <https://opendatacommons.org/licenses/by/1-0/>.
71. Ebd., S. 13.
72. Einsehbar unter <http://id.loc.gov/authorities/subjects>.
73. Malmsten, 2008. Zitiert aus: Pohl & Danowski, 2013, S. 13.
74. Pohl & Danowski, 2013, S. 14.
75. Ebd.

76Ebd.

77. Schüpbach, 2016.

78. Prongu & Hu#gi, 2013, S. 18.

79. Eine Übersicht der in dieser Arbeit erwähnten Vokabulare und Ontologien mit Präfixen und Name- spaces findet sich auf S. 48. Eine umfangreiche Auflistung von Vokabularen und Ontologien aus dem Bibliotheksbereich findet sich in: Klee, 2013.

80. Sanders, 2017 & Koha Community, 2017.

81. Inzwischen stehen sowohl die GND wie auch VIAF als Linked Open zur Verfügung. Die Deutsche Nationalbibliothek fördert sogar aktiv die Öffnung der GND gegenüber Archiven und anderen Fachbe- reichen. So lautet das Motto der GNDCon 2018 «Öffnung der GND». Vgl. wiki.dnb.de, 2018.

82. Europeana, 2017.

83. Ebd., S. 4.

84. International Council on Archives (ICA), 2000, S. 7.

85. Ebd. S. 12.

86. Arbeitsgruppe Normen und Standards (VSA), et al., 2009, S. 8.

87. Staatsarchiv des Kantons Zürich (Hrsg.), 2018, S. 50.

88. Tögel, 2016, S. 16.

89 Ebd.

90. Online verfügbar unter <http://data.archiveshub.ac.uk/ead2rdf/>.

91. Online verfügbar unter <https://labs.regesta.com/progettoReload/wp-content/uploads/2013/04/oad.rdf>. 92. labs.regesta.com, kein Datum.

92. labs.regesta.com, kein Datum.

93. Online verfügbar unter <https://labs.regesta.com/progettoReload/wp-content/uploads/2013/04/oad.rdf>.

94 Ebd.

95. Brüning, et al., 2002, S. 28-53.

96. International Council on Archives (ICA), 2000, S. 13-32.

97. Lediglich die normierten Personen, Körperschaften und Familien können als gemeinsame Ressour- cen verstanden werden.

98. International Council on Archives (ICA), 2016.

99. Ebd., S. 1.

100. Ebd., S. 10.

101. Ebd., S. 2.

102. Ebd., S. 9f

103. Llanes-Padro#n & Pastor-Sa#nchez, 2017, S. 16/18.

104. VSA, 2017, S. 4.

105. Dubois, et al., 2018, S. 1.
106. Dubois, et al., 2018, S. 14.
107. Docuteam GmbH, 2018.
108. Ebd.
109. Docuteam GmbH, 2018.
110. Stevenson, 2011, S. 29.
111. Siehe Kapitel 0.
112. Die Zahl wurde aus der zum Download angebotenen Datei berechnet: <http://data.archiveshub.ac.uk/dump/>.
113. Schweizerisches Bundesarchiv, kein Datum, S. 1.
114. Merzaghi, 2017.
115. Zugänglich unter <https://lindas-data.ch/sparql-ui/>.
116. Geschäftsstelle E-Government Schweiz, kein Datum.
117. Vergleiche <http://classifications.data.admin.ch/municipalityversion/13249>.
118. Gonzenbach, 2017
119. Europeana, kein Datum.
120. Ebd.
121. Siehe Kapitel 0.
122. Die Einträge sind über die Europeana API abrufbar. Ein Beispiel-Output befindet sich auf S. 140.
123. Online verfügbar unter <https://druid.datalegend.net/Hack-a-LOD/iisg-kg/>.
124. Estermann & Schneeberger, 2017.
125. Ebd., S. 16.
126. Estermann, 2013.
127. VSA, kein Datum.
128. Häusler, 2006, S. 5.
129. Schweizerisches Sozialarchiv, 2018, S. 2.
130. <http://www.sachdokumentation.ch>.
131. Schweizerisches Sozialarchiv, 2018, S. 32.
132. <http://www.findmittel.ch> (Papierarchiv) und <http://www.bild-video-ton.ch> (audiovisuelles Archiv).
133. Dies steht in Einklang mit der durch den VSA festgelegten Richtlinie, vgl. Arbeitsgruppe Normen und Standards (VSA), et al., 2009, S. 12.
134. Schweizerisches Sozialarchiv, 2018.
135. Eine Beschreibung des Verzeichniselements findet sich auf S. 133.

136. Die ISO-Norm 639-3 ordnet Sprachen und Dialekten eindeutige Kürzel mit drei Buchstaben zu. Zum Beispiel «gsw» für Schweizerdeutsch oder «fra» für Französisch.
137. Lexvo.org, kein Datum. 138. de Melo, 2015, S. 2.
- 138 de Melo, 2015, S. 2.
139. Lexvo.org, kein Datum.
140. Eine Beschreibung des Verzeichniselements findet sich auf S. 154.
141. Eine Beschreibung des Verzeichniselements findet sich auf S. 115.
142. The Basel Register of Thesauri, Ontologies & Classifications (BARTOC), kein Datum.
143. <https://helvetosaurus.sozialarchiv.ch/> und <https://github.com/helvetosaurus/helvetosaurus>.
144. Vgl. Kapitel 0 Festlegung des Schemas der URIs, S. 14.
145. Eine Beschreibung der Verzeichniselemente findet sich auf S. 119 und S. 121.
146. ortsnamen.ch, kein Datum; Rechtsquellenstiftung, 2018; Dodis, kein Datum; opendata.swiss, 2018; GeoNames.org, kein Datum; J. Paul Getty Trust, 2017; Deutsche Nationalbibliothek, 2018, S. 51.
147. Online abrufbar unter: https://www.nb.admin.ch/snl/fr/home/bn-professionnel/e-helvetica/infos-pour-les-fournisseurs/sites-web/_archives-web-suisse.html.
148. Schweizerische Nationalbibliothek, 2016.
149. Siehe https://www.sozialarchiv.ch/wp-content/uploads/fileadmin/user_upload/Sozialarchiv/Dokumente/PDFs/Publikationen/Publikationen_mit_Material_des_Schweizerischen_Sozialarchivs_seit_2008.pdf.
150. Vgl. http://www.findmittel.ch/archive/archNeu/Ar198_33.html.
151. Im Rahmen dieser Arbeit wurde eine solche für das Schweizerische Sozialarchiv erstellt (siehe im Anhang «Verzeichniselemente und Linked-Data-Äquivalente»).
152. Eine detaillierte Diskussion der verschiedenen Datenmodelle befindet sich im Zwischenfazit des vierten Kapitels.
153. Vgl. S. 23 und S. 28.
154. Vgl. S. 22 und S. 60.
155. Vgl. die Umwandlung von Personendaten auf S. 63.