

Zeitschrift: Arbido
Herausgeber: Verein Schweizerischer Archivarinnen und Archivare; Bibliothek Information Schweiz
Band: - (2013)
Heft: 4: Linked Open Data Big Data : alles vernetzt = tout en réseau = tutto in rete

Artikel: Linked Open Data in archivischen Fachinformationssystemen
Autor: Sandner, Peter
DOI: <https://doi.org/10.5169/seals-769741>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. [Mehr erfahren](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. [En savoir plus](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. [Find out more](#)

Download PDF: 14.01.2026

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

données d'Europeana permettra aux données du Service de la culture d'être présentées selon la structure du RDF (Ressource Description Framework).

Le référentiel des lieux, par exemple, sera aligné sur GeoNames, base de données géographique de plus de 6,5 millions de lieux existants¹⁰. Des discussions sont par ailleurs en cours à l'interne pour aligner le référentiel des personnes avec le fichier d'autorités VIAF¹¹. Par contre, l'abandon – temporaire toutefois – du référentiel des matières à ce stade du projet reporte l'alignement de ce dernier avec le thésaurus-matières pour l'indexation des archives locales proposé par le Service interministériel des Archives de France¹² ou la CDU abrégée¹³.

10 www.geonames.org/ [consulté le 15 octobre 2013].

11 <http://viaf.org/> [consulté le 15 octobre 2013].

12 <http://data.culture.fr/thesaurus/page/ark:/67717/Matiere> [consulté le 15 octobre 2013].

13 <http://udcdata.info/> [consulté le 15 octobre 2013].

Au final, les réflexions autour de la création d'un outil de recherche fédérée permettront non seulement d'offrir à toute personne intéressée un portail d'accès unique au patrimoine valaisan, qui pourrait du reste tout à fait être enrichi par des ressources en lien avec le Valais provenant d'autres institutions, mais surtout de l'inscrire dans le sillage des

projets d'open data et de linked open data actuellement en cours en matière de valorisation du patrimoine. En somme d'inscrire pleinement les données du Service de la culture dans le Web pour les valoriser au mieux.

Contact: alain.dubois@admin.vs.ch

ABSTRACT

Die Valorisierung des Kulturguts durch Onlinestellung von Daten der Dienststelle für Kultur des Kantons Wallis

Die Überlegungen zur Onlinestellung und Interoperabilität zwischen verschiedenen Informationssystemen, ebenso wie der Austausch und/oder die Wiederverwendung von Daten stehen im Zentrum der Überlegungen der Dienststelle für Kultur des Kantons Wallis, die im Gefolge der Inbetriebnahme der Speicher- und Archivierungsplattform für die digitalen und digitalisierten Bestände des Staatsarchivs des Kantons Wallis, der Mediathek und der kantonalen Museen angestellt wurden. Am Ende des ersten Semesters 2014 sollen diese Überlegungen ganz konkret zur Onlinestellung einer gemeinsamen Suche für das Walliser Kulturgut, zur Öffnung eines Datenlagers Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) und zum Abgleich der beschreibenden Daten nach den Grundsätzen des semantischen Webs führen.

Der vorliegende Artikel zeigt die wichtigsten Etappen des Projekts auf, das 2011 begonnen wurde und das zum Ziel hat, allen Interessierten den Zugang zu den kulturellen Ressourcen der Dienststelle für Kultur des Kantons Wallis über ein Guichet virtuel («One-Stop-Portal») zu ermöglichen.

(Übersetzung: as)

Linked Open Data in archivischen Fachinformationssystemen

Peter Sandner, Archivar,
Leiter der Abteilung Archivfachliche
IT/Digitales Archiv im Hessischen
Hauptstaatsarchiv in Wiesbaden

Die grosse Idee der vernetzten offenen Daten erreicht über den Umweg der übergreifenden Portalangebote auch die Archive und ihre Fachinformationssysteme. Bei der Erneuerung des hessischen Archivinformationssystems¹, das gemeinsam für das Hessische Landesarchiv und das Niedersächsische Landesarchiv entwickelt wird, ist von Anfang an erörtert worden, wie die Idee der Linked Open Data (LOD) in der Praxis realisiert werden kann.

URI für archivische Erschliessungsobjekte

Um Daten miteinander verknüpfen zu können, ist ein eindeutiger Identifikator unverzichtbar. Seit Langem werden Archivalien und Archivbestände, wie es ISAD (G)² beschreibt, mit Signaturen versehen, um sie eindeutig identifizieren zu können. Auch die Archive selbst haben in der Regel eindeutige Kürzel. In den meisten Fällen genügt also die Archivsignatur den Ansprüchen an einen einheitlichen Ressourcenidentifikator (URI – Uniform Resource Identifier), aber eben doch nicht immer. Probleme können sich ergeben bei Korrekturen, Umsignierungen oder bei

Mehrfachverzeichnungen eines Objekts. Daher ist die Verwendung maschinell generierter Identifikatoren der sicherere Weg zur eindeutigen Bezeichnung archivischer Erschliessungsobjekte.

1 Peter Sandner: Von der Findmitteldatenbank zum virtuellen Lesesaal im Netz. Rechtliche Fragen bei der Umgestaltung des Archivinformationssystems HADIS, in: Alles was Recht ist. Archivische Fragen – juristische Antworten. 81. Deutscher Archivtag in Bremen, Fulda 2012, S. 135–145

2 www.ica.org/10207/standards/isadg-general-international-standard-archival-description-second-edition.html

jekte. Dabei können beispielsweise Datenbank-ID verwendet werden. Sicherlich ist die längerfristige Stabilität solcher ID wünschenswert, noch wichtiger als ewige Geltung aber ist die Eindeutigkeit.

Verwendung von HTTP-Links für archivische Erschliessungsobjekte

Eine Vorschrift zur Verwendung von HTTP-Links für archivische Erschliessungsobjekte findet sich in den Archivgesetzen zwar nicht. Eine solche Forderung steht aber im Einklang mit dem Auftrag des neu gefassten Hessischen Archivgesetzes, Archivgut «unter Anwendung moderner Technologien für die öffentliche Nutzung zugänglich zu machen»³. Schon vergleichsweise früh haben die hessischen Staatsarchive ihre Erschliessungsdaten im Archivinformationssystem HADIS unter www.hadis.hessen.de/researcherbar gemacht⁴. Allerdings waren dort die standardmäßig angebotenen HTTP-Adressen keine festen, sondern nur temporäre Weblinks, deren Gültigkeit mit Ablauf der Session verlief. Diese Einschränkung führte zu Unmut bei jenen Anwendern, die Links zu HADIS-Objekten setzen wollten. Seit 2006 werden von HADIS auch permanente Links für diverse Portale bereitgestellt. Infolgedessen sind hessische Erschliessungsdaten u.a. aus spartenübergreifenden Angeboten wie dem BAM-Portal oder der Deutschen Digitalen Bibliothek (DDB) heraus verlinkt worden⁵. Von der HADIS-Nachfolganwendung werden durchweg feste HTTP-Adressen so angeboten, dass eine Verlinkung zu den Erschliessungsobjekten für jeden leicht möglich ist. Dabei kommt das archivische Credo über die Bedeutung des Kontextes zur Geltung, indem jeder Erschliessungsdatensatz im Zusammenhang mit der Gliederung (Klassifikation, Tektonik) angezeigt wird; ein Navigationsbaum ermöglicht das Überwechseln zu benachbarten Objekten.

Eignung archivischer Erschliessungsdaten als Open Data

Manche hinterfragen, ob archivische Erschliessungsdaten überhaupt als Open Data geeignet seien. Diese Zweifel betreffen zum einen die – verglichen mit Bibliotheksdaten – heterogene Struktur der Archivdaten. Mehr noch

aber führen die hohen Anforderungen des Persönlichkeitsschutzes zu einer Zurückhaltung mancher Archive bei Open Data. Die öffentlichen Archive verstehen sich zu Recht auch als Datenschutzbehörden. Sie haben die gesetzliche Aufgabe, das Freiheitsrecht auf informationelle Selbstbestimmung⁶ auch nach Abgabe der Unterlagen ans Archiv zu gewährleisten. Dazu dienen vor allem die archivgesetzlichen Schutz- oder Sperrfristen. Aber deshalb

Die öffentlichen Archive verstehen sich zu Recht auch als Datenschutzbehörden. Sie haben die gesetzliche Aufgabe, das Freiheitsrecht auf informationelle Selbstbestimmung auch nach Abgabe der Unterlagen ans Archiv zu gewährleisten.

müssen die betreffenden archivischen Erschliessungsdaten keineswegs vollständig unter Verschluss gehalten werden. Vielmehr ist es Aufgabe der Archive, das Recht auf Persönlichkeitsschutz mit dem Grundrecht der Wissenschafts- und Forschungsfreiheit⁷ in Einklang zu bringen. Deshalb hat die bundesdeutsche Archivreferentenkonferenz 2007 eine Empfehlung zum Umgang mit sensiblen Erschliessungsdaten verabschiedet⁸. Dieses Papier arbeitet heraus, dass Erschliessungsdaten durchaus auch dann schon in öffentlich zugänglichen Netzen publiziert werden dürfen, wenn die Akteninhalte selbst noch Schutzfristen unterliegen. Entscheidend ist, dass der Erschließungstext die schutzwürdige Information (etwa einen Personennamen) nicht preisgibt. Archivinformationssysteme sollten diese Unterscheidung – so wie in Hessen seit vielen Jahren der Fall – abbilden und die beiden Merkmale «Zeitpunkt der Onlinepublikation der Erschließung» und «Ende der Schutzfrist» getrennt speichern. Dies sollte auf der Ebene der einzelnen Verzeichnungseinheit geschehen, und nicht pauschal für den ganzen Archivbestand, denn unter Onlinebedingungen entfällt die Notwendigkeit, die Fertigstellung eines kompletten Onlinefindbuchs abzuwarten. Dieses Verfahren erfordert unter den Vorzeichen von Linked Open Data eine erhöhte Wachsamkeit bei der Bearbeitung, denn ein versehentlich veröffentlichter Datensatz lässt sich kaum zurückholen.

Bereitstellung standardisierter Informationen zu den Objekten

Inzwischen hat sich EAD⁹ zum Standardaustauschformat für archivische Informationen entwickelt. In den letzten Jahren ist es im Förderprogramm der Deutschen Forschungsgemeinschaft (DFG) zur Retrokonversion von Findmitteln verwendet worden. Auch die Deutsche Digitale Bibliothek (DDB) nutzt EAD, um Daten von den Archiven in Empfang zu nehmen. Inzwischen haben die deutschen Archive sich über ein einheitliches EAD-Profil verständigt¹⁰. Die DDB reicht die Daten weiter an die Europeana, die diese in das klassische LOD-Format RDF verwandelt¹¹ und so öffentlich bereitstellt.

3 § 1 (1) Hessisches Archivgesetz (HArchivG). Vom 26. November 2012, Gesetz- und Verordnungsblatt für das Land Hessen Teil I, S. 458

4 Bernward Helfer: Das hessische Erschließungssystem HADIS 2000, in: Angelika Menne-Haritz (Hg.): Online-Findbücher, Suchmaschinen und Portale, Marburg 2002, S. 153–171, auch unter http://archivschule.de/uploads/Publikation/VOE35/Voe35_14HELPER.pdf; Peter Haberkorn: Das Fachinformationssystem HADIS der hessischen Staatsarchive, in: Gerald Maier, Thomas Fritz (Hg.): Archivische Informationssystem in der digitalen Welt. Aktuelle Perspektiven, Stuttgart 2010, S. 181–195

5 www.bam-portal.de; www.deutsche-digitale-bibliothek.de

6 In Deutschland abgeleitet aus der Menschenwürde und dem Recht auf freie Entfaltung der Persönlichkeit: Artikel 1 (1) bzw. Artikel 2 (1) Grundgesetz

7 Artikel 5 (3) Grundgesetz

8 www.bundesarchiv.de/imperia/md/content/bundesarchiv_de/fachinformation/ark/20070320_veroeffentlichungsgrundsaeetze_ark.pdf

9 EAD = Encoded Archival Description; www.loc.gov/ead

10 www.landesarchiv-bw.de/ead/; Ulrich Fischer, Sigrid Schieber, Wolfgang Krauth, Christina Wolf: Ein EAD-Profil für Deutschland. EAD(DDB) als Vorschlag für ein gemeinsames Austauschformat deutscher Archive, in: Archivar. Zeitschrift für Archivwesen 2/2012, S. 160–162, auch unter www.archive.nrw.de/archivar/hefte/2012/ausgabe2/ARCHIVAR_02-12_internet.pdf

11 <http://pro.europeana.eu/web/guest/in-brief>

Offenheit der Daten darf nicht auf Kosten der Datenqualität gehen. Daher widmete der LOD-Summit 2013 im kanadischen Montreal der Datenpflege («Curation») bei LOD eine eigene Sektion¹². Unter diesem Gesichtspunkt sind auch die Modalitäten der Datenaktualisierung zu klären. Die hessischen Staatsarchive hatten 2006 mit dem BAM-Portal eine tägliche Aktualisierung vereinbart, um mögliche Fehler – insbesondere solche, die den Persönlichkeitsschutz betrafen – ohne Zeitverzug bereinigen zu können. Derzeit werden die Daten bei der DDB noch nach Absprache im Einzelfall aktuali-

Offenheit der Daten darf nicht auf Kosten der Datenqualität gehen. Daher widmete der LOD-Summit 2013 im kanadischen Montreal der Datenpflege («Curation») bei LOD eine eigene Sektion.

siert, aber das ist nur eine Übergangslösung in der Startphase. Nach gegenwärtiger Planung werden die hessischen Archive ihre EAD-Dateien ohne Zugangsbeschränkung im Internet verfügbar machen. Über einen Zeitstempel lässt sich erkennen, wann eine Aktualisierung erfolgt ist. Ob darüber hinaus auch unmittelbar RDF-Daten publiziert werden, ist noch zu klären.

¹² <http://summit2013.lodlam.net/2013/08/18/curation-of-lod/>

¹³ <http://pro.europeana.eu/web/guest/linked-open-data>; <http://creativecommons.org/publicdomain/zero/1.0/deed.de>

¹⁴ <http://summit2013.lodlam.net/2013/08/18/curation-of-lod/>; <http://summit2013.lodlam.net/2013/07/08/notes-from-normalizing-licensing-and-data-models/>

¹⁵ <http://de.wikipedia.org/wiki/Wikipedia:Lizenzbestimmungen>; <http://creativecommons.org/licenses/by-sa/3.0/>

¹⁶ Richtlinie 2013/37/EU des Europäischen Parlaments und des Rates vom 26. Juni 2013, Amtsblatt Nr. L 175 vom 27/06/2013 S. 0001–0008, auch unter <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:175:0001:01:DE:HTML>

¹⁷ www.lagis-hessen.de/de/subjects/index/sn/hadis

¹⁸ www.regesta-imperii.de; <http://monasterium.net>

Offenheit

Es wird diskutiert, wann Daten wirklich als «offen» zu bezeichnen sind. Die Europeana publiziert Erschließungsdaten nur, wenn sie unter der (Nicht-) Lizenz CCo verwendet werden können¹³. Dagegen wurden beim LOD Summit sachliche und rechtliche Zweifel an dieser umfassenden Lizenzierung artikuliert¹⁴. Ein Verzicht auf die Nennung des Datenurhebers widerspricht dem Ziel der Archive, ihre Archivalien besser auffindbar zu machen. Und bei Weiterverwendung der Daten zu rechtswidrigen Zwecken könnte die Reputation des Archivs leiden. Die deutschsprachige Wikipedia erlaubt eine Nachnutzung ihrer Artikel unter der Creative-Commons-Lizenz BY-SA, die die Nennung des Urhebers fordert und die Weitergabe von Bearbeitungen unter den gleichen Bedingungen zulässt¹⁵. Die möglichen Nachteile einer CCo-Lizenzierung sind abzuwägen gegenüber den Vorteilen, die im positiven Fall einen gesellschaftlichen Mehrwert für alle Beteiligten erzeugen. Sicherlich wird die 2013 neu gefasste EU-Richtlinie über die Weiterverwendung von Informationen des öffentlichen Sektors, die innerhalb von zwei Jahren in nationales Recht umzusetzen ist, die Entwicklung beflügeln: Erstmals werden nun nämlich auch Bibliotheken, Museen und Archive verpflichtet, die freie (wenn auch nicht kostenfreie) Weiterverwendung ihrer bereitgestellten öffentlichen Daten zu ermöglichen¹⁶.

Adressaten der verlinkten Daten

Wer mit einem gewissen Aufwand Links und weitergehende Informationen zu seinen Erschließungsobjekten im Internet bereitstellt, möchte auch, dass die Angebote genutzt werden. Bisher zeigen sich – neben den nationalen und europäischen Archivportalen – vor allem andere Archiv- oder Kulturinstitutionen daran interessiert, Links zu Verzeichnungsobjekten der regionalen Archivportale wie dem in Hessen zu setzen. Dadurch werden Quellen im Zusammenhang präsentiert oder Recherchemöglichkeiten durch Zusatzinformationen optimiert. Die landesgeschichtliche Anwendung LAGIS beispielsweise setzt Links, um die von den hessischen Staatsarchiven online gestellten digitalisierten Perso-

nenstandsregister gezielter recherchierbar zu machen¹⁷. Mehrfach haben Projektförderungen dazu geführt, dass mittelalterliche Urkunden der hessischen Staatsarchive im Verbund online präsentiert worden sind¹⁸.

In der nächsten Ausbaustufe der Datenverknüpfung könnten Archivdaten selbst als historische Forschungsdaten genutzt werden – für die Geschichtsforschung und für andere Disziplinen. Dazu dürften Erschließungsdaten nicht mehr nur als Verweise auf einzelne Akten und Urkunden verstanden werden, sondern als ein eigener Datenkorpus, der weiter gehende Auswertungen zulässt. Beispielsweise könnte eine systematische Auswertung eines grösseren Korpus von Erschließungsdaten Erkenntnisse darüber erbringen, welche Krankheiten im Zeitverlauf diagnostiziert wurden, welche Bodenschätze wann an welchen Orten gewonnen wurden; welche Personennamen über die Jahrhunderte hinweg verwendet wurden usw. Im Moment ist bei derartigen Forschungsthemen ein direkter Kontakt zum Archiv unumgänglich, und er wird auch in Zukunft sicherlich weiterhin angeraten sein, denn die archivischen Erschließungsdaten sind teilweise erklärungsbedürftig, und aufgrund archivischer Bewertung und Kassation haben statistische Auswertungen ihre Grenzen. Dennoch werden die Möglichkeiten der Forschung deutlich erweitert, wenn Daten in einer strukturierten Form frei zugänglich sind.

Open Government

Über die Onlinebereitstellung von Erschließungsdaten hinaus können staatliche und kommunale Archive – wie alle anderen öffentlichen Verwaltungen auch – ihren Beitrag zu einer offenen und transparenten Verwaltungskultur leisten. Im Sinne des Open Government lassen sich statistische Daten zur archivischen Tätigkeit in einer weiterverarbeitbaren Form publizieren. Bei diesem Aspekt der Datenbereitstellung stehen die meisten deutschen Archive noch am Beginn der Überlegungen. Zu denken wäre etwa an die jährliche Anzahl der Zugänge, der neu verzeichneten Archivalieneinheiten oder der Nutzerinnen und Nutzer im Archiv usw. – kurzum: an alle

Daten, die auch heute schon in Jahres- und Tätigkeitsberichten publiziert werden¹⁹, aber bislang, da sie nur in Papier- oder PDF-Dokumenten enthalten sind, nur schwer automatisiert ausgewertet werden können.

Links zu Ressourcen anderer Anbieter

Ein Kernpostulat der Linked-Open-Data-Idee ist die Verlinkung zu Ressourcen anderer Informationsanbieter, um die Daten zu Elementen des Semantic Web zu machen. Hierbei erweist es sich als Problem, dass die archivischen Erschliessungsdaten bei den meisten Archiven bislang nur bedingt strukturiert oder standardisiert sind. Zwar bieten EAD-Elemente eine Zuordnung, jedoch findet dabei oft kei-

ne Entitätenprüfung statt: Wenn im Titelfeld das Wort «Müller» vorkommt, geht häufig nur aus dem Kontext hervor, ob es sich um den Beruf oder um einen

Eine Anreicherung der Erschliessungsdaten um eindeutige Informationen ist dazu notwendig und auch erstrebenswert, jedoch mit den personellen Kapazitäten der Archive allein kaum zu bewältigen.

Personennamen handelt – und im zweiten Fall weiss man nicht, ob es sich um dieselbe Person namens Müller handelt wie in einer anderen Verzeichnung. Daher sind solche Datei-

en nur eingeschränkt für das Semantic Web geeignet. Eine Anreicherung der Erschliessungsdaten um eindeutige Informationen ist dazu notwendig und auch erstrebenswert, jedoch mit den personellen Kapazitäten der Archive allein kaum zu bewältigen. Ziel muss daher eine automatisierte Anreicherung um Normdaten sein, sodass der Aufwand der Archive sich auf eine Restmenge und auf die Qualitätskontrolle beschränkt. Hier sind Verbünde und Kooperationen unerlässlich. Vordringlich erscheint vor allem die Erstellung oder Optimierung von Personen- und Ortsregistern. Erschliessungsdatensätze sollten dazu mit Datensätzen grösserer Anbieter verlinkt werden. Dabei kommen etablierte Dienste wie GeoNames für ein Ortsregister und DBpedia für ein Personenregister infrage, die auch bei der Europeana hierfür verwendet werden²⁰.

Letztlich gilt es gerade für mittlere und kleinere Archivverwaltungen, sich verbreiteten Lösungen anzuschliessen, die im Verbund genutzt und weiter entwickelt werden. Nur dann kann die grosse Vision der Linked Open Data in der Praxis auch mit Leben gefüllt werden.

Kontakt: Poststelle@hhstaw.hessen.de

ABSTRACT

Linked Open Data dans les systèmes d'information professionnels d'archives

La grande idée des données ouvertes et reliées entre elles (Linked Open Data) intéresse également les archives et leurs systèmes d'information professionnels par le biais d'offres de portails transversaux. Dans le cadre de la rénovation du système d'information archivistique du Land de Hesse, développé conjointement pour les archives de Hesse et de Basse-Saxe, la question de savoir comment concrétiser l'idée des Linked Open Data (LOD) s'est posée dès le début. Les Archives du Land de Hesse permettent depuis des années de relier leurs données via des liens HTTP. L'EAD (encoded archival description; description archivistique encodée) a fait ses preuves comme standard d'échange pour les archives. La bibliothèque numérique européenne Europeana publie également des données archivistiques au format RDF. L'entretien («Curation») des open data représente un véritable défi pour le futur. La question porte également sur le degré d'ouverture des données. Jusqu'ici, ce sont surtout des institutions patrimoniales qui utilisent les liens des systèmes d'information archivistiques. Les données archivistiques devraient également être utilisées à l'avenir pour la recherche, et les archives devraient elles-mêmes se considérer comme des administrations transparentes au sens de l'open government.

(traduction: sg)

¹⁹ www.hauptstaatsarchiv.hessen.de/irj/HHStAW_Internet?cid=b95c44612cd6101b5bb41fbf2f40116b

²⁰ <http://pro.europeana.eu/web/guest/in-brief>