

Informationen finden und Wissen verlinken - Der Weg der Metadatenstandards vom Archivregal zu den Linked Data

Die Informatisierung der Archivarbeit hat die Notwendigkeit von Metadatenstandards aufgezeigt, denn die Metadaten ermöglichen es, analoge und elektronische Unterlagen effizient zu verwalten und zu erschliessen. Die Aufbereitung der Metadaten als Linked Data wird den Archivnutzenden neue Zugangsmöglichkeiten eröffnen.

Wer schon vor einer Blechdose ohne Etikett stand, konnte sich selber überzeugen, wie wichtig Metadaten sind. Metadaten, die „Daten über Daten“, begleiten Archivare und Archivarinnen seit der Entstehung der ersten Archive. Die Metadaten und deren Standardisierung haben mit der Digitalisierung der Archivarbeit ständig an Bedeutung gewonnen. In diesem Artikel wird gezeigt, wie die Metadaten sich entwickelt haben, wie sie die Arbeit der Archivare verändert haben und wie sich mit den neuen Standards zu Linked Open Data neue Möglichkeiten für die Benutzer eröffnen.

Metadaten für Menschen. Die Erschliessungsstandards

Für Archivare und Archivarinnen ist es lebensnotwendig, dass sie ihre Unterlagen im eigenen Magazin finden. Dazu benutzten sie bereits im Mittelalter Archivinventare.

Normen für die Erschliessung waren aber in der Zeit der Archivinventare in Papierform, zumindest in der Schweiz, kaum vorhanden. Interne Richtlinien, die das Ziel hatten, die Arbeit zu rationalisieren und die Findmittel einheitlich zu gestalten, liessen den Archivmitarbeitenden immer noch sehr viele Freiheiten.

Erst mit der Implementierung der ersten digitalen Archivinformationssysteme wurde eine Normierung unumgänglich. Das International Council of Archives (ICA) reagierte auf dieses Bedürfnis und begann in den 1980er Jahren mit seinen Überlegungen zu Erschliessungsstandards. 1994 erschien die erste Version der Norm ISAD(G) [1](#). Diese wurde sechs Jahre später, im Jahr 2000, revidiert.

ISAD(G) legt archivische Prinzipien wie die des Provenienzprinzips und jenes der mehrstufigen Verzeichnung fest. Die Norm stellt auch sicher, dass die Informationen in ihrem Entstehungskontext beschrieben werden. Mittlerweile wenden fast sämtliche Archive der Schweiz ISAD(G) an. Nur dank dieser gemeinsamen Norm sind heute beispielsweise gemeinsame Archivportale möglich.

ISAD(G) ist eine Norm, die das Provenienzprinzip abbildet, sie ermöglicht aber nur eine Suche in der Archivstruktur. Suche nach Personen, Prozessen oder Funktionen sind kaum möglich. Die Standards ISAAR(CPF) und ISDIAH ermöglichen durch die Erstellung von Normdateien andere Arten des Zugangs. ISAAR(CPF) hat sich allerdings nur selten und ISDF kaum durchgesetzt.

Metadaten für Maschinen. Das digitale Magazin und seine Normen

Kurz nachdem die Metadaten elektronisch wurden, entstanden auch immer mehr Unterlagen digital. Ein digitales Magazin, in dem die Informationen nicht aus festem Material sondern aus Bits bestehen, braucht aber mehr Metadaten als ein analoges, damit das Schriftgut gemäss Norm ISO 15489 die Merkmale Authentizität, Zuverlässigkeit, Integrität und Benutzbarkeit aufweisen kann². So dokumentieren zum Beispiel spezifische Metadaten, wann ein Dokument in ein anderes Format migriert worden ist.

Das *Referenzmodell OAIS* (Norm ISO-14721) stellt das logische Grundgerüst für die digitale Archivierung dar. Normierte Metadaten sind vor allem im Bereich des Imports notwendig, damit die Schnittstellen zwischen Aktenproduzenten und Archiv korrekt funktionieren können. Wie es oft bei neuen Technologien der Fall ist, sind in einer ersten Phase zahlreiche unterschiedliche Spezifikationen für die verschiedenen digitalen Pakete (SIP/AIP/DIP³) entstanden. Die Erarbeitung von nationalen und internationalen Standards ist im Gange.

In der Schweiz erarbeitete das Bundesarchiv seine SIP-Spezifikationen basierend auf Dublin Core⁴ und ISAD(G). Diese Spezifikation diente als Grundlage für weitere Arbeiten des Vereins eCH, der im Jahr 2012 den Standard *eCH-0160: Archivische Ablieferungsschnittstelle* herausgab⁵.

Docuteam hat in Zusammenarbeit mit dem Staatsarchiv Wallis bei der Erarbeitung seiner Ablieferungswerkzeuge keine eigene Metadatenpezifikation entwickelt, sondern sich hauptsächlich auf eine Palette von bestehenden Normen gestützt. Das Objektprofil wurde 2013 bei der Library of Congress als Matterhorn METS Profile registriert [6](#).

Ein Beispiel für eine internationale Standardisierung ist [das Projekt E-ARK](#) (2014-2017), bei dem auch die Schweiz mitgewirkt hat. Das E-ARK Projekt verfolgte unter anderem das Ziel, ein gemeinsames Information Package (IP) als Common Specification zu definieren. Anschliessend wurden für SIP/AIP/DIP resp. für die verschiedenen Anwendungsfelder (Ablieferung, Archivierung, Bereitstellung) wo nötig differenzierte Anforderungen definiert. Die neue gemeinsame Definition eines IP sollte insbesondere den Austausch von Daten zwischen Institutionen erleichtern, was gerade unter europäischen Archiven zunehmend an Bedeutung gewinnt. Der E-ARK Standard für das Information Package wurde 2016 publiziert. Die Zeit wird zeigen, ob er sich durchsetzen wird.

Die normierten Metadaten dienen nicht nur der Verwaltung von Informationen im digitalen Magazin, sie können auch automatisiert in bestehende Archivinformationssysteme importiert werden [7](#).

Metadaten für alle. Linked Open Data

Von dieser technischen Entwicklung und von den neuen Standards, welche die Archivarbeit grundlegend verändert haben, konnten die Nutzer wenig profitieren.

Das Provenienzprinzip oder die hierarchische Erschliessung, welche eine Archivplansuche ermöglichen würden, sind für viele Archivbenutzer sehr schwierig zu verstehen. Für die meisten Benutzer ist die Suche nach einem bestimmten Wort der Königsweg für die Suche von Informationen in den Archiven. Die Metadaten sind heute aber häufig nicht auf diese Art der Suche ausgerichtet. Als Konsequenz können viele Archive den Nutzern nur mit aufwändiger Beratung helfen [8](#).

Open Data könnte hier eine Lösung bieten. Nach diesem Prinzip werden die Daten frei im Netz zur Verfügung gestellt. Die Nutzer können diese herunterladen und nach Ihren Bedürfnissen weiter bearbeiten. Open (Government) Data ist Teil der E-Government Strategie der Schweiz. Auf dem Portal opendata.swiss, das unter Federführung des Schweizerischen Bundesarchivs entstanden ist, können öffentliche Organisationen ihre Daten hochladen. Um die Daten im Portal leichter zu finden, wurde ebenfalls ein Metadatenstandard entwickelt. Es handelt sich um das DCAT-Anwendungsprofil für Datenportale in der Schweiz (DCAT-AP CH), es dient zur Beschreibung der in Schweizer Datenportalen verzeichneten Daten [9](#).

Open Data sind aber nicht immer lesbar für Maschinen und sind in unterschiedlichen Formaten vorhanden. Werden sie aber nach dem Standard RDF (Resource Description Framework) gespeichert, werden sie lesbar für Maschinen. So bearbeitete Daten können miteinander in Beziehung gebracht werden, daher der Name „Linked Data“. Die Abfragen sind nicht mehr an die Archivstruktur gebunden und Daten aus mehreren Beständen können verknüpft werden. So sind neue Zugangsmöglichkeiten auf die Bestände möglich (z.B über Personen, Orte usw.) [10](#).

Der Bund hat mit LINDAS einen Prototyp eines Linked Data Service entwickelt. Seit 2017 betreibt das Bundesarchiv den Prototypen. Entwicklerinnen und Entwickler erhalten über lindas-data.ch Zugang zu diesen Daten. Sie können sie mittels einer SPARQL-Suchanfrage abrufen.

Ein weiterer Prototyp im Bereich Linked Data ist [aLOD](#) (*Archival Linked Open Data*) (vgl. [den Artikel \(auf Französisch\)](#)). Es handelt sich um eine Suchmaschine, welche die Möglichkeit der Nutzung von Linked Data für die Archive prüfen will und dank der Zusammenarbeit von mehreren Schweizer Archiven entwickelt worden ist.

Damit Archive Linked Data in Zukunft nutzen können, sind spezifische Standards nötig, welche Linked Data und die Prinzipien der Archivwissenschaft verknüpfen. Die ICA hat auf diese Herausforderung reagiert und erarbeitet gerade die Norm „[Records in Context](#)“ (vgl. [den Artikel](#)).

Ziel dieses Standards ist die Harmonisierung und Integration der vier bestehenden Metadatenstandards: ISAD(G), ISAAR(CPF), ISDF und ISDIAH. Es werden somit die bewährten Prinzipien beibehalten, sie können aber mit den oben erwähnten neuen Technologien verbunden werden.

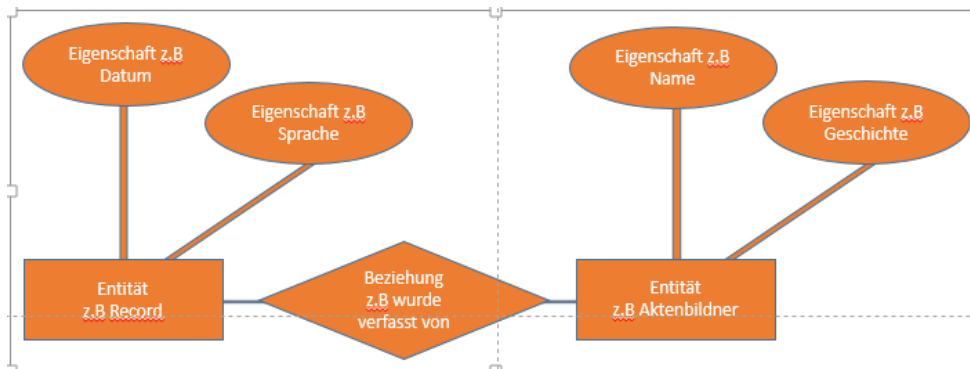


Abb. 1: Verknüpfungen zwischen Entitäten, Eigenschaften und Beziehungen

Diese neue Norm wird in zwei Teile gegliedert:

- *RiC-CM (Records in Contexts Konzeptmodell)*: Dieses Modell beschreibt die wichtigsten Entitäten (z.B. „Record“ oder „Aktenbildner“) des Modells, deren Eigenschaften und wie sie zueinander in Beziehung stehen. Für jede Entität werden zudem die Eigenschaften festgelegt (z.B. Datum, Sprache oder Material für einen Record). Bild 1: Verknüpfungen zwischen Entitäten, Eigenschaften und Beziehungen
- *RIC-O (Records in Contexts Ontologie)*: Ausgedrückt mittels Web Ontology Language (OWL) sollen hier die Entitäten und Eigenschaften genau normiert und maschinenlesbar im Sinne von Linked Data werden. Ein Entwurf sollte bis Ende 2017 zur Verfügung zu stehen.

Zusammen sollen es Modell und Ontologie ermöglichen, die Metadaten von Archiven zu bearbeiten und sie mit denen anderer Gedächtnisinstitutionen zu verknüpfen, ohne auf die Schlüsselkonzepte der archivischen Erschließung zu verzichten.

Fazit

Internationale Standards und die Informatik haben die Erschließung und Verwaltung von Archivunterlagen effizienter gemacht. Bei den Entwicklungen der Standards konnten die Archivare die Prinzipien Ihrer Arbeit konsolidieren und erhielten so wichtige Instrumente für ihre Arbeit. Mit Linked Data werden neue Verknüpfungen von Daten möglich. Dazu braucht es aber entsprechende Standards. Die Archivgemeinschaft hat bereits reagiert: Mit „Records in Contexts“ ist ein neuer Standard in Erarbeitung und hat viele Erwartungen geweckt.

¹ International Standard Archival Description (General).

² Siehe. Norm ISO 15489. Kapitel 7.2.

3 SIP - submission information package werden, AIP - archival information package, DIP - dissemination information package.

4 Dublin Core ist ein standardisiertes Set von Konventionen zur Beschreibung von Dokumenten und Objekten im Internet, um diese mit Hilfe von Metadaten in stichwortbasierten Suchmaschinen gezielt und schnell finden zu können. Siehe: URL: <http://www.vsa-aas.org/de/aktivitaet/ag-normen-und-standards/normenkatalog/> <http://www.vsa-aas.org/de/akti...> [Stand 07.07.2017].

5 Hinweis zum Standard auf der Webseite des Vereins eCH. URL: <http://www.ech.ch/vechweb/page?p=dossier&documentNumber=eCH-0165&documentVersion=1.0> <http://www.ech.ch/vechweb/page...> [Stand 07.07.2017].

6 Siehe: URL: <http://www.loc.gov/standards/mets/news112912.html> <http://www.loc.gov/standards/m...> und https://wiki.docuteam.ch/lib/exe/fetch.php?media=ois:spezifikation_matterhorn-mets_20160803_wi.pdf <https://wiki.docuteam.ch/lib/e...> [Stand 07.07.2017].

7 Siehe hier zum Beispiel Extraktion deskriptiver Metadaten aus der archivischen Ablieferungsschnittstelle eCH-0160. <http://kost-ceco.ch/cms/download.php?78d4243b18aa483b219055afc1c39798> <http://kost-ceco.ch/cms/downlo...>

8 Siehe: Müller, Tamara: Die Schwierigkeiten bei der Recherche im Archiv(-katalog), Ursachenforschung und Vorschläge zur Problemstellung. In: Semar, Wolfgang, Brigitte Lutz: *Churer Schriften zur Informationswissenschaft* Nr 74. Chur 2015. URL: http://www.htwchur.ch/uploads/media/CSI_74_Mueller.pdf <http://www.htwchur.ch/uploads/...> [Stand 21.07.2017].

9 Die Standardisierung auf nationaler Ebene ist beim Verein eCH in Gange.

10 Die Semantic University bietet Tutorials an, um die Grundprinzipien von Linked Data, RDF und SPARQL zu verstehen. URL: <http://www.cambridgesemantics.com/semantic-university/getting-started-semantic> <http://www.cambridgesemantics....> [Stand 21.07.2017].

?Michele Merzaghi



Michele Merzaghi ist seit 2012 wissenschaftlicher Mitarbeiter und Berater im Bereich elektronische Geschäftsverwaltung (GEVER) beim Schweizerischen Bundesarchiv. Seit 2010 ist er der Präsident der „Arbeitsgruppe Normen und Standards“ des Vereins Schweizerische Archivare und Archivarinnen (VSA). Der gebürtige Tessiner hat Geschichte, Kulturwissenschaft und Kunstgeschichte in Basel studiert und 2007 abgeschlossen. 2008-2010 absolvierte er das Weiterbildungsprogramm in Archiv-, Bibliotheks- und Informationswissenschaften (MAS ALIS) in Bern und Lausanne und 2015 das CAS in Fundamental of Informatics an der Berner Fachhochschule.

Résumé

Deutsch

Metadaten beschäftigen Archivare und Archivarinnen seit langer Zeit, durch die Informatisierung der Archive wurden immer mehr Metadatenstandards festgelegt, diese ermöglichen analoge und elektronische Unterlagen effizient zu verwalten und zu erschliessen. Von den vielen Metadatenstandards konnten bis jetzt nur die ArchivarInnen profitieren, denn Archivnutzer hatten nur wenige Suchmöglichkeiten in den Metadaten. Die Aufbereitung der Metadaten als Linked Data aber wird den Archivbenutzenden nun neue Zugangsperspektiven ermöglichen.

Français

Les métadonnées sont depuis longtemps une préoccupation pour les archivistes. Depuis l'informatisation des archives, de plus en plus de normes de métadonnées ont été établies afin de décrire et gérer de manière efficace des documents analogiques et électroniques. Toutefois, ces normes restent actuellement profitables surtout aux archivistes; les utilisateurs d'archives n'ont encore que très peu de possibilités de faire des recherches avec les métadonnées. Le traitement des métadonnées dans le cadre des Linked Data devrait cependant offrir aux chercheurs de nouvelles perspectives pour la réutilisation des données d'archives.

Italiano

Da generazioni gli archivisti si occupano di metadati. Con l'informatizzazione sono stati definiti sempre più standard per i metadati che hanno permesso di descrivere e gestire i documenti in modo sempre più efficiente.

Degli standard definiti finora hanno maggiormente approfittato gli archivisti, mentre agli utenti restavano poche possibilità di ricerca. Con l'elaborazione dei dati nell'ambito dei Linked Data si aprono nuove possibilità di ricerca per gli utenti.