(Open) Linked Data in Bibliotheken [Patrick Danowski, Adrian Pohl]

Autor(en): **Holländer, Stephan**

Objekttyp: **BookReview**

Zeitschrift: **Arbido**

Band (Jahr): **- (2013)**

Heft 4: **Linked Open Data Big Data : alles vernetzt = tout en réseau = tutto in rete**

PDF erstellt am: **20.05.2024**

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.
Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.
Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

Ein Dienst der *ETH-Bibliothek*
ETH Zürich, Rämistrasse 101, 8092 Zürich, Schweiz, www.library.ethz.ch

http://www.e-periodica.ch

Open Linked Data in Bibliotheken – Herausforderungen beim Aufbau eines «Linked Open Data»-Systems

Stephan Holländer

Patrick Danowski und Adrian Pohl: (Open) Linked Data in Bibliotheken, De Gruyter Verlag, Berlin 2013

Rechtzeitig zur Buchmesse in Frankfurt erschien von den Herausgebern Patrick Danowski und Adrian Pohl eine Publikation zu (Open) Linked Data in Bibliotheken. Die Herausgeber warten gleich in einem ersten Kapitel mit einem eigenen Beitrag zu den Grundlagen und einem Überblick zu Linked Open Data (LOD) auf. Beide Verfasser stellen zu Recht fest, dass die Anzahl der Daten und Datenquellen, die als Linked Data zur Verfügung gestellt werden, seit 2007 stetig zugenommen hat. Grund genug, ein Buch zu diesem Thema zu veröffentlichen. Mit den beiden Herausgebern stehen zwei profilierte Experten zum Thema zur Verfügung. Patrick Danowski ist leitender Bibliothekar am Institute of Science and Technology Austria (IST Austria) in Klosterneuburg in Österreich. Adrian Pohl arbeitet seit 2008 im Hochschulbibliothekszentrum des Landes Nordrhein-Westfalen (hbz). Seit Juni 2010 ist er Koordinator der Arbeitsgruppe der Open Knowledge Foundation zu Open Bibliographic Data.

Gerade diesen beiden erfahrenen Autoren und Referenten geschieht im Eröffnungskapitel ein Fehler, der mit der Thematik sehr vertrauten Autoren öfters passiert. Sie fragen sich nicht, wo man die Einsteiger ins Thema abholen sollte und legen gleich voller Elan mit der Schilderung der Entwicklungsgeschichte von Linked Open Data (LOD) seit 2007 los. Und die gerät sehr ausführlich, ersetzt aber die Erklärung der Grundlagen nicht. Wo findet man eine Erklärung des LOD-Konzepts und wozu Graphen und Triples in diesem Zusammenhang dienen? Suchen im Abschnitt «Grundlagen»? Fehlanzeige! Die Autoren setzen das unbewusst als bekannt voraus. Hingegen wird die Entwicklungsgeschichte des LOD im angelsächsischen und deutschsprachigen Raum in allen seinen Verästelungen mit Schneid und Akribie aufgezeigt. Gerade für das Verständnis der Ausführungen im später folgenden Kapitel von Pascal Christoph sind diese einführenden Erklärungen unerlässlich. Irgendwann scheint den Autoren aber die Geduld ausgegangen zu sein, denn die Übersicht zu den Entwicklungen zwischen 2007 und 2012 kommen als schnöde Tabellenübersicht daher. So einfach kann man es sich machen. Hier hätte ein Fachlektorat eingreifen sollen, das für eine ausbalanciertere Darstellung der Gesamtthematik hätte besorgt sein müssen.

Carsten Klee, der Autor des nachfolgenden Kapitels zum Thema «Vokabulare für bibliografische Daten. Zwischen Dublin Core und bibliothekarischem Anspruch» hat sich seiner Aufgabe souveräner entledigt. Ein konzises und gut geschriebenes Kapitel, das den Weg von Bibliotheken mit ihren Vokabularen für bibliographische Daten auf dem Weg ins Netz aufzeigt. Es werden verschiedene Vokabulare aufgezeigt und erläutert. Gut zu gefallen weiss die Beschreibung bibliografischer Daten im Web. Von solchen Daten gibt es eine grosse Auswahl, die hier nur überblicksweise und in Ausschnitten beleuchtet werden konnte, wie der Autor entschuldigend meint. Dies ist aber für den geforderten Zweck völlig ausreichend. Carsten Klee ist zuzustimmen, wenn er schreibt, dass im Gegensatz zu den bibliothekarischen Austauschformaten MARC oder MAB die in Linked-Library-Data-Projekten eingesetzten Vokabulare einfach und gut überschaubar sind. Die Zusammenfassung am Schluss ist das Tüpfelchen auf dem i zu dem sehr gelungenen Beitrag.

Ein sehr solides Kapitel trägt Hans-Georg Becker zur Beschreibung von fortlaufenden Sammelwerken unter Verwendung von Functional Requirements for Bibliographic Records (FRBR) bei. Der Autor entwickelt die von ihm gewählte Thematik gut und führt übersichtlich durch die verschiedenen Aspekte des Themas. Über Diagramme werden den Lesern die Publikations- typen in FRBR-Darstellung nahegebracht. Der Autor fügt diesen Diagrammen noch die Darstellung zu den Publikationstypen in validem Turtle (Terse RDF Triple Language) hinzu, die der Serialisierung für Resource Description Framework (RDF-)Graphen dient. So stellt man sich das Rundumsorglos-Paket für den interessierten Praktiker vor. Der Verfasser weist in seinem Fazit darauf hin, dass seine Ausführungen zum einen auf dem Linked-Open-Library-Data-Portal der Universitätsbibliothek Dortmund und zum anderen auf das DFG-Projekts ArcheoInf basiert.

Dominique Ritze, Kai Eckert und Magnus Pfeffer widmen sich dem wichtigen Thema Forschungsdaten, denn jede Forschung beginnt mit einer Recherche vorhandener Publikationen, aber es können auch die Forschungsdaten für den Wissenschaftler zu Beginn seiner Forschung von Nutzen sein. Deshalb mag es hilfreich sein, die Forschungsdaten direkt in ein Recherchesystem einer Bibliothek zu integrieren, in dem der Forscher nach existierenden Publikationen zu seinem Forschungsgegenstand recherchiert. Hier erkennen Hochschul- und Forschungsbibliotheken einen weiteren Bereich für ihre Informationsvermittlung. Die Autoren beschreiben in ihrem sehr interessanten Beitrag die Möglichkeiten für die Integration von Forschungs- daten anhand des Projektes «Integration von Forschungsdaten und Literatur in den

Sozialwissenschaften (InFoLiS)». InFoLiS ist ein DFG-gefördertes Projekt des Leibniz-Instituts für Sozialwissenschaften, der Universitätbibliothek Mannheim und des Lehrstuhls für Künstliche Intelligenz an der Universität Mannheim.

Pascal Christoph bringt uns die Datenanreicherung auf LOD-Basis näher. Die Herstellung von Kontext kann als Kataloganreicherung einen Mehrwert bei der Recherche für den Nutzer mit sich bringen. Durch Linked Open Data gehen wir einen Schritt weiter. Es ist nicht nur möglich, Daten zu integrieren, um damit den eigenen Katalog anzureichern. Vielmehr ist die Basiszielsetzung von LOD auch das Publizieren der Daten, wie der Autor zutreffend ausführt. Sein Beitrag beleuchtet einführend die theoretischen Aspekte, die im Zusammenhang mit der Katalog-anreicherung auf Basis von Linked (Open) Data wichtig sind. Anhand des vom Hochschulbibliothekszentrum des Landes Nordrhein-Westfalen (hbz) betriebenen LOD-Services lobid.org wird aufgezeigt, wie bibliografische Daten auf Basis anderer LOD-Quellen angereichert werden können. Abschliessend betrachtet der Verfasser künftige Chancen und Möglichkeiten der Kataloganreicherung. Er erläutert die drei möglichen Ansätze «blosse Verlinkung», «dynamische Kataloganreicherung» und «Datenübernahme mit ihren Vor- und Nachteilen». Schnell zeigt sich, dass die Lizenzierungsfrage bei der Entscheidung, welchen Ansatz die Bibliotheken wählen können, eine zentrale Rolle spielt, da sie bei der Nutzung von Fremddaten von den vertraglichen Nutzungsbedingungen abhängig sind. Ein guter Beitrag, der aber Vorkenntnisse in der Computerlinguistik voraussetzt.

Mit zwei weiteren Beiträgen gehen Markus M. Geipel, Christoph Böhme, Julia Hauser und Alexander Haffner («Herausforderung Wissensvernetzung: Impulsgebende Projekte für ein zukünftiges LOD-Konzept der Deutschen Digitalen Bibliothek») sowie Klaus Ceynowa, Matthias Gross, Andreas Kahl und Gabriele Messmer («Linked Open Data geht in die Fläche: Der B3Kat stellt seine Daten frei») Fragen der Umsetzung an Projekten wie der Deutschen Digitalen Bibliothek und in Länderverbundprojekten nach. Den gängigen bibliografischen Metadatenformaten liegt ein anderes Datenkonzept zugrunde als beim Resource Description Framework (RDF). Einen MARC-, Pica- oder MAB-Datensatz kann man sich als eine Art digitale Karteikarte vorstellen, deren Felder ausgefüllt sind. RDF hingegen betrachtet Daten von bibliografischen Angaben als Netzwerk. Es existieren keine einzelnen Datensätze in herkömmlichem Sinne mehr, sondern nur noch Knoten, die durch Aussagen verknüpft sind. Dieser Paradigmenwechsel macht die Umwandlung von bibliografischen Daten ins RDF schwierig, da noch keine standardisierten Konkordanztabellen zwischen den Bibliotheksformaten und dem RDF geschaffen wurden.

Diese Neuerscheinung sticht in zweierlei Hinsicht aus den bisherigen Publikationen der Reihe «Bibliotheks- und Informationspraxis» des Verlages De Gruyter hervor. Einerseits, da die Autoren von der bisherigen Konzeption des «Ready for use» der Reihe abrücken, was teilweise der Komplexität des Publikationsgegenstandes geschuldet ist, und andererseits, weil es das erste Werk in dieser Reihe des Verlages ist, welches von den Verlegern Open Access unter Creative Common Licence gestellt wird (http://www.degruyter.com/viewbook-toc/product/181080). Das verdient Anerkennung. Ein Gewinn der Reihe «Bibliotheks- und Informationspraxis» ist es bisher gewesen, dass man die Bücher auch als Berufspraktiker ohne allzu spezielle Vorkenntnisse lesen konnte. Daran sollte man seitens des Verlags weiterhin festhalten, denn das ist die eigentliche Stärke der Reihe. Wer sich andererseits tapfer auch ohne linguistische Vorkenntnisse durch die mehr theoretischen Beiträge des Buches durchgebissen hat, legt das Buch mit Gewinn zur Seite, da er die Mög-lichkeiten von (Open) Linked Data erahnt und sich der Schwierigkeiten bei deren Umsetzung bewusst wird.

Big Data für Anfänger – ein Reiseführer für «Nichtinformatiker»

Eliane Blumer

Rudi Klausnitzer: Das Ende des Zufalls – Wie Big Data uns und unser Leben vorhersagbar macht (Ecowin Verlag, Salzburg, 2013)

Das Buch «Big Data. Das Ende des Zufalls» ist der Blick eines «alten Hasen» aus der Kommunikations- und Medienwelt, der es sich zur Aufgabe gemacht hat, tiefer in die Big-Data-Welt einzutauchen und seine Entdeckungsreise mittels eines Buches festzuhalten. Dabei will er seiner Auffassung, dass Big Data eines der Jahrhunderthemen ist, jedoch noch nicht die nötige Aufmerksamkeit erhalten hat, Nachdruck verleihen.

Die Reise von Rudi Klausnitzer beginnt ganz vorne – nämlich bei der Fragestellung, warum denn Big Data? Hier seine Erklärung zusammengefasst: Weil der Mensch, neurologisch bewiesen, den Zufall nicht mag und ihn deshalb so oft als möglich vermeiden will. Und Big Data kann uns dabei helfen, denn durch die maschinelle Kontextualisierung der richtigen Datensets können Vorhersagen getroffen werden, die denn Zufall quasi ausschalten. So kann beispielsweise die Verlinkung von Fahndungsdaten mit bereits existierenden kriminellen Mustern dazu dienen, mögliche Verbrechensorte vorauszusagen und den effizienten Einsatz von Polizisten