



Linked Data für Bibliotheken: Vorteile eines konzeptionellen Wechsels von bibliotheksspezifischen Datensatzstrukturen hin zu RDF-basierten Datenmodellen

Getaneh Alemu, Brett Stevens

Penny Ross und Jane Chandler

University of Portsmouth
Portsmouth, United Kingdom

*Deutsche Übersetzung:
Susanne Oehlschläger
Deutsche Nationalbibliothek*

Meeting:

92 — New futures for bibliographic data formats: reflections and directions — UNIMARC Core Activity

Abstract:

Die aktuellen Metadatenprinzipien und -standards tendieren dazu, das Dokument und weniger die Daten in den Mittelpunkt zu stellen, d. h. sie führen zu menschenlesbaren anstelle von maschinell zu verarbeitenden Metadaten. Damit Bibliotheken gemeinsam nutzbare, miteinander mischbare und weiterverwendbare Metadaten erstellen und nutzbar machen können, kann durch die Anpassung der derzeitigen Bibliotheksmodelle wie die Resource Description and Access (RDA) und die Functional Requirements for Bibliographic Records (FRBR) an Modelle, die auf den Linked-Data-Prinzipien basieren, ein konzeptioneller Wechsel erreicht werden. In Bezug auf technische Formate können Bibliotheken zu Linked-Data-Formaten wechseln, ohne die aktuellen Bibliotheksmetadaten-Abläufe zu unterbrechen.

Dieser Vortrag liefert sechs wichtige Empfehlungen für Bibliotheken und Standardisierungsagenturen. Dazu gehört, sich der Herausforderung zu stellen und die Möglichkeiten anzunehmen, welche die aktuellen technologischen Trends bieten, die Minimalanforderungen der Linked-Data-Prinzipien anzunehmen, Ontologien zu entwickeln, darüber zu entscheiden, was aus den gegenwärtigen Bibliotheksmodellen beibehalten werden muss, Teil der Linked-Data-Cloud zu werden und gemischte Metadatenansätze (standardbasiert und von der Community erstellt) zu entwickeln. Schließlich endet das Papier mit der Identifizierung und Diskussion von fünf wesentlichen Vorteilen einer solchen Neukonzeption der Metadaten. Zu den Vorteilen gehören die Offenheit und der Austausch der Metadaten, Serendipity-Effekte bei der Recherche nach Informationsressourcen, die Identifikation von dem Zeitgeist entsprechenden und neu entstehenden Metadaten, facettenbasierte Navigation und mit Links angereicherte Metadaten.

Schlagwörter: *Metadaten, Metadatenstandards, MARC, Linked Data, RDF, community-erzeugte Metadaten, gemischte Metadaten-Ansätze*

Einführung

Die Prinzipien, Standards und Protokolle, die derzeit in Bibliotheken verwendet werden, um Metadatenfunktionen abzusichern, haben eine lange Geschichte (Denton, 2007; Dunsire, 2009; IFLA, 2009; Lubetzky, 1953; Wright, 2007). Die Geschichte der Katalogisierung in Bibliotheken spannt sich über viele tausend Jahre und geht zurück zu den Tagen der Bibliothek von Alexandria im 3. Jahrhundert vor Christus, wo der Bibliothekar Callimachus eine systematische Bibliografie der griechischen Literatur erstellte (Day, 2005; Wright, 2007). Morville (2005) geht mit der Geschichte der Katalogisierung sogar noch weiter zurück in die Zeit des Assyrischen Reichs, wo König Assurbanipal 650 v. Chr. eine Bibliothek mit 30.000 Tontafeln aufgebaut hat. Die Geschichte der Standardisierungsbestrebungen beim Katalogisieren beginnt jedoch erst im 19. Jahrhundert (Coyle & Hillmann, 2007; Denton, 2007; Weinberger, 2005, 2007; Wright, 2007). Zu den Hauptpionieren der modernen Katalogisierung gehören Sir Anthony Panizzi, Charles Cutter, Melville Dewey, Paul Otlet, S. R. Ranganathan und Seymour Lubetzky (Denton, 2007; Lubetzky, 1953; Wright, 2007). Unter Bezugnahme auf die Arbeiten dieser Pioniere beschäftigten sich lokale und internationale Konsortien mit Projekten, die darauf abzielten, strenge Katalogisierungsprinzipien und -regeln einzuführen, die folgerichtig in Veröffentlichungen solcher Standards wie der Paris Principles, 1961, der Anglo-American Cataloguing Rules (AACR), 1967, Machine-Readable Cataloguing (MARC), in den späten 1960er Jahren, der International Standard Bibliographic Description (ISBD) for Monographic Publications, 1971, der Functional Requirements for Bibliographic Records (FRBR), 1996, und Resource Description and Access (RDA), 2010, resultierten (Denton, 2007). Mit dem beständigen Wachstum der Bestände in digitalen Bibliotheken werden die Metadatenprinzipien und -standards intensiven Prüfungen unterzogen (Alemu, Stevens, & Ross, 2012; Coyle, 2010; Coyle & Hillmann, 2007; Lagoze, 2010; Mathes, 2004; Shirky, 2005; Veltman, 2001; W3C, 2011; Weinberger, 2005, 2007). Kritiker behaupten, dass die aktuellen Bibliotheksstandards einige der Beschränkungen fortführen, die dem traditionellen Kartenkatalog innewohnen, und deshalb sowohl für das Fehlen der Skalierbarkeit als auch der Interoperabilität mit der heutigen Informationslandschaft verantwortlich sind (Coyle, 2010; Coyle & Hillmann, 2007; Weinberger, 2005, 2007). Die durch die Verwendung der gegenwärtigen Metadatenstandards und technischen Formate erzeugten Metadaten sind hauptsächlich auf die menschliche Ausschöpfung und weniger auf maschinelle Verarbeitung ausgerichtet. Dies ist sowohl auf die konzeptionellen als auch auf die technischen Beschränkungen der Standards und Technologien zurückzuführen, die in Bibliotheken verwendet werden. Coyle (2010) legt dar, dass "der Bibliothekskatalog der einzige Kontext für Bibliotheksdaten" ist, und damit nicht mit externen Informationslieferanten zusammenarbeiten kann. Die Herausforderungen an Prinzipien, Standards und Protokolle könnte man aus zweierlei Perspektiven betrachten: der konzeptionellen und der technischen. Der konzeptionelle Unterbau von gegenwärtigen Metadatenstandards wie den FRBR und den RDA hat wohl zu Metadatensätzen als Dokumenten geführt, die auf die menschliche Benutzung abgestimmt sind und weniger auf die Verarbeitung durch Maschinen (Coyle & Hillmann, 2007). Wenn solche Metadatenprinzipien unter Verwendung technischer Formate wie MARC implementiert werden, weisen die resultierenden Datensätze außerdem Probleme mit der Metadaten-Duplizierung, der Dateninkonsistenz, dem Mangel an Granularität und der Komplexität auf (Coyle, 2010; Coyle & Hillmann, 2007; Day, 2000; Guenther & McCallum, 2003; Tennant, 2002). Und obwohl der Ruf nach einem Ende der Verwendung von MARC als lange überfällig bezeichnet wird (Tennant, 2002), bleibt der Standard die dominante Metadatenstruktur, die bis heute in Bibliotheken verwendet wird. Dies kann direkt mehreren Gründen zugeordnet werden, einschließlich der Tatsache, dass MARC tief in die Bibliothekssysteme und -funktionen eingebettet ist und deshalb jede Veränderung einfach zu schwierig und zu teuer werden würde, oder dass MARC für Bibliotheken trotz allem

„passend“ genug ist und ihren Zwecken dient, oder es kann auch sein, dass alternativen Formaten, einschließlich der eXtensible Markup Language (XML), zusätzliche Funktionalitäten fehlen, die für ihre Bevorzugung erforderlich wären und den Wechsel rechtfertigten. Aber es gibt immer noch starke Zweifel im Hinblick auf die Eignung von MARC, die von mehreren Metadatenexperten vertreten werden, die behaupten, dass der Standard nicht zu maschinell zu verarbeitenden und umsetzbaren Metadaten passt (Coyle, 2010; Coyle & Hillmann, 2007; Wallis, 2011a, 2011b).

Zu den vorgeschlagenen Alternativen gehört die Anwendung von Linked Data. Wie deren Erfinder, Berners-Lee, et al (2001), feststellen, ist das Originalweb konzeptionell gesehen dokumentenzentrisch, d. h. die Links haben keinerlei semantische Bedeutung und die resultierenden Webseiten sind eher für die menschliche Verwendung als für die Maschinenverarbeitung gedacht, wenn sie mit technischen Formaten wie der Hyper-Text Mark Up Language (HTML) implementiert werden.

Andererseits, während das Web für den Austausch von Dokumenten außergewöhnlich effizient ist und Möglichkeiten zur Zusammenarbeit bietet, erfordert ein Dokument menschliches Zutun, um seine Bedeutung zu verstehen, sobald es auf dem Bildschirm des Benutzers angezeigt wird. Einfach gesagt, Maschinen können den Sinn solcher Dokumente nicht erkennen. Um dieser Beschränkung entgegenzuwirken, haben Berners-Lee, et al (Berners-Lee, 1997; Berners-Lee, 1998; Berners-Lee, et al., 2001) die Idee des Linked Data (Berners-Lee, et al., 2001) ins Spiel gebracht. Zu den Technologien, die zum Implementieren von Linked Data benötigt werden, gehören RDF, RDFS, SPARQL und OWL (Allemnag & Hendler, 2008; Berners-Lee, 1998; Decker et al., 2000; W3C, 2004a, 2004b). Nach Allemnag und Hendler (2008) ist „der Hauptgedanke von Linked Data, ein verteiltes Web auf der Grundlage von Daten und weniger auf der Grundlage von deren Präsentation (Dokumenten) zu unterstützen. Anstatt dass eine Webseite auf eine andere hinweist, können die Daten zu anderen verweisen unter Verwendung von globalen Referenzen, die URIs genannt werden.“ Trotz des wachsenden Interesses an Linked Data gibt es eine andauernde, verbreitete Debatte in Bezug auf schrittweise Veränderungen an traditionellen bibliothekszentrierten konzeptionellen Prinzipien wie RDA und FRBR. Dieselbe Frage gilt, wenn es darum geht, ob der Wechsel bei technischen Formaten von MARC zu RDF evolutionär erfolgen soll oder ob es einen zwingenden Grund dafür gibt, einige oder alle diese traditionellen Formate aufzugeben und diese neuen Ansätze zu verfolgen (Coyle, 2010; Coyle & Hillmann, 2007; Marcum, 2011; Styles, 2009; Styles, Ayers, & Shabir, 2008; W3C, 2011; Wallis, 2011a, 2011b).

Eine weitere wichtige Herausforderung, die Standardisierungsstellen und Bibliotheken im Hinblick auf die aktuelle Metadaten-Neukonzeptionierung im Auge behalten sollten, ist die Konsequenz des Web-2.0-Ansatzes bei der Erzeugung und Verwendung von Metadaten in digitalen Bibliotheken. Obwohl die Rolle des Web 2.0 für Bibliotheken weitgehend anerkannt ist, sind, wie Evans (2009) schreibt, „*seine Implikationen insbesondere im Bibliothekswesen noch nicht vollständig realisiert*“, vielleicht weil es eine konzeptionelle Inkompatibilität zwischen den traditionellen Bibliotheksmodellen und den aufkommenden Web-2.0-Ansätzen gibt, wie Lagoze (2010) behauptet. Es wird kritisiert, dass die Ansätze zu community-erzeugten Metadaten (Web 2.0) flach, eindimensional und inkonsistent sind (Gruber, 2007, 2008), während andererseits die standardbasierten Metadatenansätze für ihre rigiden Hierarchien und ihr Unvermögen, den Wortschatz der Benutzer zu verwenden (Shirky, 2005; Veltman, 2001; Weinberger, 2007), kritisiert werden. Andererseits verfechten Gruber (2007) und Morville (2005) den Standpunkt, dass durch Experten kontrollierte und durch Benutzer erzeugte Metadaten nicht als gegensätzliche Ansätze angesehen werden sollten, die nicht nur nebeneinander bestehen, sondern sich auch gegenseitig beeinflussen.

Im Hinblick auf eine optimale Nutzung der Möglichkeiten von Linked Data (Web 3.0) ist ein gemischter Metadatenansatz ausschlaggebend, der nicht nur aus der physischen Beschreibung von Informationsobjekten (wie zum Beispiel Verfasser, Titel, ISBN, Thema, Format usw.) besteht, sondern auch Elemente einschließt, die deren sozio-kulturelle Facetten (User Tags, Kommentare, Rezensionen, Links, Bewertungen, Empfehlungen) beschreiben. Mit anderen Worten, die Repräsentation des sozialen Bereichs von Metadaten sollte als ebenso wichtig angesehen werden wie das Erfassen von standardisierten und objekt-bezogenen Metadatenelementen (Web 1.0), die bisher verwendet werden, um die physischen Eigenschaften von Informationsobjekten zu charakterisieren. Dies ist besonders bedeutend für die Metadatenfülle, wo sowohl durch den Benutzer (die Community) erzeugte Metadaten als auch auf Standards basierende, von Bibliothekaren erzeugte Metadaten weiterhin nutzbar gemacht werden.

Dieser Vortrag diskutiert die Herausforderungen gegenwärtiger Metadatenprinzipien und -standards und zeigt Vorteile und Folgen eines konzeptionellen Wechsels von dokumentenzentrischen hin zu datenzentrischen Metadatenansätzen auf. Es werden zwei wesentliche breite Klassen, nämlich Metadatenprinzipien (wie RDA und FRBR) auf der einen Seite und fachspezifische Datensatzformate (wie MARC) auf der anderen Seite identifiziert, und es werden Empfehlungen gegeben, welche dieser Klassen modifiziert, angepasst oder beibehalten werden sollten, um den Linked-Data-Prinzipien zu entsprechen.

Linked-Data-Prinzipien

Wie der Name bereits zeigt, versteht man unter Linked Data ein Datenmodell, das strukturierte Datenelemente miteinander verlinkt und in Beziehung setzt, analog zu der Art und Weise wie relationale Datenbanksysteme funktionieren, nur dass dieses im Web funktioniert. Der allgemeine Zweck von Linked Data ist es, die Weiterverwendung, die Verlinkung, die Integration und den Austausch von Daten zu erleichtern (Berners-Lee, 2009; Shadbolt, 2010; W3C, 2011). Berners-Lee (2007) bemerkt, dass "eine Seite hinzuzufügen, Inhalt erzeugt, einen Link hinzuzufügen die Organisation, Struktur und Ergänzung von Informationen im Web bietet, was die Inhalte insgesamt in etwas sehr Wertvolles verwandelt."

Es ist wichtig festzustellen, dass Linked Data ein Meta-Modell darstellen, wobei dieses einen Rahmen bietet, um Schemas und Vokabulare jeder Art und Größe in einer bestimmten Domain zu definieren, zu planen, zu entwickeln und zu pflegen. Das wiederum bedeutet, dass Institutionen wie zum Beispiel Bibliotheken, nicht notwendigerweise ihre bestehenden Metadatenstandards, kontrollierten Vokabulare, Normdateien und vorhandenen Daten aufgeben müssen. Um allerdings mit der Welt der Linked-Data-Cloud zusammenarbeiten zu können, müssen diese Institutionen die darunter liegenden Linked Data übernehmen.

Die Linked-Data-Prinzipien starten beim grundlegendsten Bestandteil von Linked Data, d. h. der Verwendung von weltweit eindeutigen URIs als Bezeichnungen, um solche Dinge wie Informationsobjekte, Personen, Orte und Ereignisse unterscheidbar zu kennzeichnen (Berners-Lee, 2009). Zu den weiteren wesentlichen Prinzipien gehört die Fähigkeit, eine vorhandene URI unter Verwendung von Hypertext Transfer Protocol (HTTP) aufzulösen, um relevante Informationen zu finden, wie z. B. worauf sich diese bestimmte URI bezieht, Datenformate zu verwenden, wie zum Beispiel RDF/XML, Vokabular-Definitionssprachen zu verwenden, wie RDFS und OWL, und eine Suchsprache zu verwenden, die als SPARQL bekannt ist, und das Letzte, aber nicht minder Wichtige ist die Einbindung von Links innerhalb und außerhalb der Datensets, womit die Daten angereichert werden und eine vom Kontext abhängige Bedeutung erhalten.

Linked Data zur Modellierung von Bibliotheksmetadaten

Linked Data und die damit in Verbindung stehenden Technologien spielen eine bedeutende Rolle bei der Metadatenmodellierung, ihrer Enkodierung, ihrer Darstellung und ihrem Tausch. Die Verwendung von URIs für Metadaten-Elementbezeichnungen, Labels und Beziehungen erleichtert sowohl deren Benennung als auch die Identifizierung von Konflikten bei der Verwendung der Elemente. Es wird darauf aufmerksam gemacht, dass das einfache Datenmodell von RDF die Bildung von semantischen Links zwischen Informationsressourcen ermöglicht (Coyle, 2010; Coyle & Hillmann, 2007; Day, 2000, 2003a, 2003b; Helen, 2010; Nilsson, 2010; Rothenberg, 2008; Styles, et al., 2008; W3C, 2011; Wallis, 2011a; Wilson 2010).

Annahme geschlossene versus offene Systeme

Die Prinzipien von Linked Data machen ein offenes, dynamisches und interaktives System möglich und funktionieren in diesem. Bibliotheksstandards allerdings funktionieren zumeist in einer geschlossenen und statischen Umgebung, losgelöst von der allgemeinen Web-Informationslandschaft. Nach Allemnag und Hendler (2008) bietet Linked Data ein verteiltes Datenmodell, in dem „jeder alles über jedes Thema sagen kann“, was zu „Variationen und Widersprüchen“ in der Bedeutung von Entitäten führen kann. Wie die Verfasser bestätigen, „ist eine offene Welt in diesem Sinn eine, in der wir zu jedem Zeitpunkt davon ausgehen müssen, dass neue Informationen an den Tag kommen können und wir zu keinem Zeitpunkt davon ausgehen können, dass die verfügbaren Informationen alle verfügbaren Informationen sind“. Im Kontext digitaler Bibliotheken führt die Annahme der offenen Systeme selbst offen zu neuen Metadatenbeiträgen aus verschiedenen Sichtweisen, wobei immense Vorteile darin bestehen, was als Netzwerkeffekte bekannt ist. Durch die Verwendung von URIs – welche explizite Spezifikationen von Begriffen in einer bestehenden Domain unter Verwendung von Vokabularen und Beziehungen zwischen ihnen sind (Gruber, 1993) – können Ontologien vermischt, gematcht und gemergt werden.

Aktueller Stand der Verwendung von Linked Data in Bibliotheken

Es scheint, insbesondere unter den Organisationen des öffentlichen Bereichs, wie zum Beispiel den Nationalbibliotheken, den Drang zu geben, ihre bibliografischen Daten frei und offen verfügbar zu machen (Wilson, 2009). 2011 wurden zwei Abschlussberichte veröffentlicht, die beide die Bedeutung der Öffnung von Bibliotheksmetadaten in Formaten hervorheben, die einfach zugänglich und wiederverwendbar sind (Library of Congress, 2011; W3C, 2011). Wie Haslhofer and Isaac (2011) hervorheben, sind die Europeana und die mit ihr affilierten National- und Regionalbibliotheken übereingekommen, den Linked-Data-Weg zu gehen. Die British Library hat bereits ein Linked-Data-Modell entwickelt (The British Library, 2011). Ebenso hat die Europeana ihr Metadatenmodell von der Europeana Semantic Elements-Spezifikation zum Europeana Data Model (EDM) überarbeitet, wobei das letztgenannte mehr auf Linked Data angepasst ist (Doerr et al., 2012; Haslhofer & Isaac, 2011).

Wilson (2010) weist darauf hin, dass die Entscheidung der British Library, ihre bibliografischen Daten frei zur Verfügung zu stellen, in Einklang mit dem Bekenntnis der Regierung des Vereinten Königreichs hin zu Transparenz und Verantwortlichkeit steht. Der Verfasser stellt auch fest, dass ein Teil des Schritts zu Linked Data der wachsenden Erwartung der Bibliotheksbenutzer zugerechnet werden kann, die Bibliothek mit neuesten technologischen Entwicklungen Schritt halten zu lassen. Das Datenmodell der British Library enthält und verwendet mehrere bestehende URI-basierte Vokabulare und Ontologien wie den Virtual International Authority File (VIAF), die Library of Congress Subject Headings

(LCSH), Lexvo (eine URI-referenzierte, kontrollierte Liste von Zeichen, Wörtern und Termen), GeoNames (eine geografische Datenbank), die MARC-Länder- und Sprachencodes, Dewey.info (die obersten Klassen der Dewey Dezimalklassifikation) und RDF book mash-up (Informationen über Bücher und ihre Verfasser) (The British Library, 2011).

Langsame Annahme von Linked Data für Bibliotheksmetadaten

Der Bericht der W3C Library Linked Data Incubator Group (2011) bestätigt das langsame Tempo, mit dem Bibliotheken Linked Data annehmen, während er den Bibliotheken starke Argumente und Empfehlungen liefert, sich die Linked-Data-Prinzipien zu Eigen zu machen. Der Bericht betont die Bedeutung dessen, bibliografische Bibliotheksdaten offen und frei zugänglich zu machen, in einer Form, die „austauschbar, erweiterbar und leicht weiterverwendbar ist“ (W3C, 2011). Er weist auch darauf hin, dass „Bibliotheksstandards (MARC, Z39.50) nur für die Bibliothekscommunity konzipiert sind“, und es für andere schwierig machen, die Daten mit deren Daten weiterzuverwenden und neu zu kombinieren. Der Bericht stimmt vollständig mit den Einschätzungen überein, die von Coyle (2010) und Styles, et al (2008) abgegeben wurden, die die heutige Bibliotheksmetadatenentwicklung und das Annehmen der Herausforderungen behandeln und sich dafür einsetzen, diesen Herausforderungen durch die Annahme von webbasierten Standards zu begegnen, die die Wiederverwendbarkeit, Verlinkung und den Austausch von Metadaten begünstigen.

Shadbolt und Hall (2006) bemerkten, dass es gilt, zu begreifen, dass „den Punkt zu erreichen, an dem die glückliche Wiederverwendung von eigenen und fremden Daten möglich wird“. Sie bedauern, dass wenige Bemühungen darauf verwendet wurden, eine breitere Verwendung des Semantic Web voranzutreiben, was zum Teil der Betonung zugeschrieben werden kann, die auf „Sprachen, Formalismen, Standards und Bedeutung“ gelegt wurde. Weinberger (2012) glaubt, dass „während das ursprüngliche Semantic Web den Schwerpunkt darauf legte, Ontologien aufzubauen, die ‚Repräsentationen des Wissens‘ der Welt darstellen, es sich zeigt, dass wenn wir geradewegs dazu übergehen, eine Fülle verlinkter, aber unvollkommener Daten freizugeben und sie in standardisierter Form breit verfügbar zu machen, das Netz zu einer grundlegend verbesserten Wissensinfrastruktur wird.“

Angesichts der großen Menge an Investitionen und Bemühungen, die bereits von Organisationen wie der IFLA, der Library of Congress, OCLC und vielen anderen Bibliotheken für die Entwicklung und Pflege des MARC-Formats aufgewendet wurden, ist der Übergang zu Linked Data keine einfache Angelegenheit. Die Tatsache, dass Bibliotheken bedeutende Mengen an wertvollen Metadatensätzen in MARC-Formaten haben, ist unbestritten. Zum Beispiel hat die British Library seit den 1950er Jahren ihre nationalbibliografischen Datensätze Außenstehenden über einen subskriptionsbasierten Zugriff verfügbar gemacht, ursprünglich unter Verwendung von gedruckten Formaten und dann nach der Einführung der Bibliotheksautomatisierung unter Verwendung des Standards MARC (The British Library, 2011; Wilson 2010).

Wallis (2011b) beobachtet, dass obwohl es nun einige nationale und regionale Kulturerbe-Institutionen (wie CENL und die Europeana) gibt, die sich öffentlich dazu bekannt haben, den Open-Data-Weg zu fahren, insgesamt grundsätzliche Herausforderungen bleiben, ihre Daten weiter zu verwenden. Er setzt sich damit auseinander, dass obwohl die Konvertierung vorhandener bibliothekarischer Datensätze in RDF äußerst wichtig ist, die daraus resultierenden Datensets die bibliotheksspezifischen Sprachen und Terminologien beibehalten, die „für jeden, der auf der Suche nach nutzbaren Daten ist, der jedoch kein Bibliothekar ist, unzugänglich bleiben.“

Herausforderungen, Linked Data in Bibliotheken anzuwenden

Der Weg hin zur Anwendung von Linked Data in Bibliotheken ist nicht ohne Herausforderungen. Wenn es um Bibliotheken geht, gibt es drei Herausforderungen. Erstens die weitreichende Verwendung des Standards MARC als Grundlage der aktuellen Bibliotheksverwaltungssysteme und die vorhandenen Metadaten, die seit vielen Jahren erzeugt werden. Obwohl es vorherrschend ist, wird das MARC-Format als datensatz- und dokumentenzentrische Metadatenstruktur und weniger als ein einsetzbares datenzentrisches Format angesehen (Coyle, 2010; Coyle & Hillmann, 2007; Styles, 2009; Styles, et al., 2008). Es gibt buchstäblich Milliarden Datensätze in MARC-Formaten; ein Versuch, die kleinste Bewegung davon weg zu machen, hätte umfassende Implikationen im Hinblick auf Ressourcen. Obwohl die Beschränkungen des MARC-Standards seit Anfang des einundzwanzigsten Jahrhunderts dargelegt sind, verwenden Bibliotheken es bis heute (Tennant, 2002). Das Thema heute ist nicht, dass den Bibliotheken und Standardisierungsgremien die Beschränkungen von MARC nicht bewusst sind, sondern vielleicht die Tatsache, dass alternative Formate wie XML als Ersatz nicht besser geeignet sind.

Eine zweite Herausforderung, die die W3C Library Linked Data Incubator Group (W3C, 2011) herausgestellt hat, ist die terminologische Verschiedenheit, die zwischen bibliotheks- und webbasierten Standards besteht. Im Hinblick darauf führt Styles (2009) an, dass Zweideutigkeiten aus der Anwendung des FRBR-Modells entstehen können. Styles (2009) behauptet, dass "niemand über Werke, Expressionen und Manifestationen [die den Grundaufbau des FRBR-Modells darstellen] spricht, also warum sollen wir unsere Daten auf diese Art und Weise beschreiben?" Wallis (2011a), der eine ähnliche Ansicht vertritt, schlägt ein sehr vereinfachtes Modell vor, das sich auf das Objekt (das Exemplar) zum Beispiel ein bestimmtes Buch konzentriert, und weniger auf seine Abstraktionen als Werk, Expression oder Manifestation. Beide Verfasser stimmen darin überein, dass solche Abstraktionen nicht im Alltagsgebrauch von Verlagen, Katalogisierern und Bibliotheksnutzern vorkommen. Insbesondere Wallis empfiehlt, dass die Bibliotheks- und die Linked-Data-Community zusammenarbeiten und solche Unterschiede überbrücken sollen, um die Weiterverwendbarkeit und Erweiterbarkeit von Bibliotheksdaten für außerhalb dieser Community stehende zu erleichtern. Wallis erörtert, dass die neuen Initiativen zur Entwicklung von Bibliotheksstandards wie RDA und FRBR auf Einfachheit abzielen sollten, während sie den Metadatenreichtum ausnutzen, der durch die Verwendung von Linked Data möglich wäre.

Die dritte und wichtige Herausforderung, die potentielle Anwender konfrontiert, ist die Komplexität von Linked-Data-Technologien wie RDF/XML, RDFS, OWL und SPARQL. Es gibt einen offensichtlichen Mangel an Tools und Applikationen zur Erzeugung von Linked Data in Bibliotheken. Berners-Lee hat bemerkt, dass „das [aktuelle] Web gewachsen ist, weil es einfach ist, eine Webseite zu schreiben und mit anderen Seiten zu verlinken“ (Berners-Lee, 2007). Im Hinblick darauf ist es zwingend notwendig, dass die besagten Technologien relativ einfach erlernbar und anwendbar gemacht werden, analog zu der Einfachheit, HTML-Seiten in den frühen Jahren des Web zu erstellen. Wie die Dinge heute stehen, sind diese Technologien generell für Leute außerhalb der Linked-Data-Community zu kompliziert. Damit eine breitere Erlernbarkeit erfolgen kann, sollte jeder mit Basiskenntnissen von Webdesign in der Lage sein, eine Seite zu erstellen, die auf diesen Linked-Data-Technologien basiert. Last but not least konzentrieren sich die aktuellen Bemühungen zur Annahme von Linked-Data-Prinzipien und bei der Veröffentlichung von bibliografischen Daten auf standardbasierte, vorhandene Metadaten und tendieren dazu, innerhalb der Community erzeugte (benutzererzeugte) Metadatenansätze zu ignorieren.

Empfehlungen für einen konzeptionellen Wechsel

Im Folgenden sind sechs Hauptempfehlungen für Bibliotheken und Standardisierungsstellen aufgeführt, damit diese sich mit den drei oben identifizierten hauptsächlichsten Herausforderungen beschäftigen können. Zu diesen Empfehlungen gehören:

- Bibliotheken sollen sich den Herausforderungen durch Linked Data stellen
- Übernahme der Minimalanforderungen der Linked-Data-Prinzipien
- Entwicklung von Ontologien: Entitäten (Klassen), Elemente (Propertyts) und Werte (Instanzen)
- Entscheiden, was beibehalten werden muss
- Teil des Linked-Data-Webs werden
- Übernahme eines Ansatzes gemischter Metadaten

Bibliotheken sollen sich den Herausforderungen durch Linked Data stellen

Im Licht der aktuellen Paradigmen des Web 2.0 und des Web 3.0 und mit dem Ziel, von den Mechanismen der weltweit eindeutigen Identifizierung, der Verlinkung und den Möglichkeiten zur Weiterverwendung zu profitieren, die die Linked-Data-Prinzipien ermöglichen, sollten sich Bibliotheken, Standardisierungsstellen und mit Bibliotheken in Verbindung stehende Stellen wie die IFLA den Herausforderungen stellen und die Möglichkeiten annehmen, die die aktuellen technologischen Trends mit sich bringen. In diesem Zusammenhang schlägt Coyle vor, dass „der Wandel, dem wir begegnen müssen, ist dass das Web zunehmend die Informationsquelle für Suchende und Forscher ist, und dass die Bibliothek mit diesem Datennetz gekoppelt sein muss.“ Der bereits zitierte Bericht der Library Linked Data Incubator Group des W3C (W3C, 2011) erkannte ebenfalls die Notwendigkeit für einen Wechsel von den domänenspezifischen Bibliotheksstandards und Datenformaten (wie MARC, Z39.50) hin zu Linked-Data-Standards und Datenformaten wie RDF. Der Transfer von vorhandenen Bibliotheksmetadaten aus dem OPAC in den Web-Kontext wird dennoch eine massive Herausforderung sein, deshalb ist die enge Zusammenarbeit der verschiedenen Beteiligten wie Bibliotheken, Archive, Museen, Verlage und Standardisierungsstellen zwangsläufig äußerst wichtig.

Übernahme der Minimalanforderungen der Linked-Data-Prinzipien

Für Bibliotheken und Standardisierungsstellen besteht die Notwendigkeit, die Minimalanforderungen der Linked-Data-Prinzipien zu übernehmen. Deshalb wird es bei der Anpassung von vorhandenen Bibliotheksmetadaten an Linked-Data-Strukturen erforderlich sein, das RDF/XML-Serialisierungsformat zu verwenden. Der Bericht der Library Linked Data Incubator Group (2011) des W3C weist darauf hin, dass Bibliotheken, Standardisierungsstellen und Händler davon profitieren würden, wenn sie „ihren Bereich erweitern oder mit Linked-Data-Initiativen zusammenarbeiten würden“ (W3C, 2011). Die Verwendung von URIs, HTTP-URIs, RDF und die Verlinkung innerhalb der Datensets und aus ihnen heraus würden zu Übereinstimmungen mit diesen Minimalanforderungen der Linked-Data-Prinzipien führen.

Entwicklung von Ontologien: Entitäten (Klassen), Elemente (Property's) und Werte (Instanzen)

Ein konzeptioneller Wandel in der Wahrnehmung von Bibliotheksmetadaten als ein Dokument oder ein Datensatz hin zu dem, was Coyle (2010) als verwertbare Metadaten (actionable Metadata) bezeichnet, d. h. solche, die maschinenlesbare, vermischbare und neu kombinierbare Metadaten sind, ist notwendig. Um dies zu erreichen, ist es erforderlich, dass vorher Metadatenmodelle (Ontologien) entwickelt werden. Zu diesem Zweck sollten FRBR und RDA ein zentrales Gerüst bilden und als hochwertige Bibliotheksdatenmodelle dienen. Nach Horridge, Knublauch, Rector, Stevens und Wroe (2004) helfen Ontologien dabei, das Wissen eines bestimmten Bereichs mit allen Begriffen und den zwischen ihnen bestehenden Beziehungen zu erfassen. Wie Noy and McGuinness (2000) beschreiben, ermöglichen Ontologien ein gemeinsames Verständnis der Struktur von Informationen zwischen Mensch und Software, erleichtern die Weiterverwendung von Wissen innerhalb eines Bereichs, indem sie dessen Annahmen explizit beschreiben und das Wissen des Bereichs von dem operativen Wissen trennt und bei der Analyse des Bereichswissens helfen.

Nach Berners-Lee, et al (2001), bezieht sich das Wort Ontologie nicht auf "eine Theorie über die Art der Existenz, also darüber welche Art von Dingen existiert", sondern die Definition ist aus ihrer Verwendung innerhalb der Künstliche-Intelligenz-Community übernommen. Deswegen ist der Term „Ontologie“ definiert als „ein Dokument oder eine Datei, das/die die Beziehungen zwischen Termen formal definiert“ (Berners-Lee, et al., 2001). Nach Horridge, et al (2004) werden "Ontologien dafür verwendet, das Wissen über einen Interessensbereich zu erfassen". Die am häufigsten zitierte Definition von Ontologie stammt allerdings von Gruber (1993) als „eine detaillierte Spezifikation einer Konzeptionalisierung“ (Gruber, 1993).

Nach Noy und McGuinness (2000) gehört zur Entwicklung einer Ontologie auch der Prozess, Klassen, Unterklassen, Property's und Instanzen (Werte) zu definieren. Es wird dringend empfohlen, dass bei der Entwicklung von Bibliotheksmetadaten-Ontologien die Konstrukte verwendet werden, die durch die Semantic-Web-Technologien bei der Spezifikation der Definitionen von Vokabularen auf der oberen Ebene angenommen wurden. Außerdem wird auch vorgeschlagen, dass die Entwicklung von Ontologien zentralisiert (normiert) wird und vorzugsweise etablierten Institutionen wie der IFLA, dem Joint Steering Committee for Development of RDA, der British Library, der Library of Congress und OCLC überlassen werden, obwohl die Beiträge von dezentralisierten, aber kollaborativen Initiativen von Einzelpersonen und Communitys nicht ausgeschlossen werden können.

Entscheiden, was beibehalten werden muss

Es muss zwischen Bibliotheksmodellen und Datensatzformaten unterschieden werden, weil Bibliotheken in hohem Maße davon profitieren würden, die Prinzipien der zugrunde liegenden Bibliotheksmetadaten zu übernehmen, wie sie in existierenden, bibliothekszentrischen Modellen wie den RDA und FRBR manifestiert sind. Bibliotheken könnten allerdings auch davon profitieren, obsolete Formate sowie Such- und Retrievalprotokolle (z. B. MARC und entsprechend auch Z39.50) aufzugeben. Im Hinblick darauf sind die Verfasser dieses Vortrags davon überzeugt, dass der Ruf danach, existierende Metadatenmodelle (wie RDA und FRBR) aufzugeben, gegen das Ethos des Semantic Web steht. Desweiteren wird empfohlen, dass internationale Institutionen wie die International Federation of Library Associations and Institutions (IFLA) ebenso wie regionale Standardisierungsstellen und Bibliotheken im Allgemeinen die Führungsrolle übernehmen sollten und gezielte Bemühungen in Initiativen setzen sollten, die darauf abzielen, aktuelle Modelle und Standards mit den Linked-Data-Modellen kompatibel zu machen. Um von

diesem Prozess optimal profitieren zu können, sollten Bibliotheken das beste aus ihrem Wissen und ihrer Expertise im Bereich von kontrollierten Vokabularen wie LCSH, Medical Subject Headings (MeSH) und Normlisten machen. Von dem Obenstehenden ausgehend hat dieses Papier zwei Vorbedingungen identifiziert, nämlich den Nutzen der Neukonzeptionalisierung bestehender Modelle wie RDA und FRBR, um diese mit den Linked-Data-Prinzipien kompatibel zu machen. Und Bibliotheken sollten bestehende MARC-kompatible Metadatensätze in RDF/XML-Formate konvertieren und sich damit langsam von MARC und seinen Derivaten wegbewegen. In anderen Worten: es ist notwendig, einen konzeptionellen Wechsel zu machen, weg von Datensatz-zentrierten Formaten hin zu RDF-basierten Datenformaten.

Teil des Linked-Data-Web werden

Wenn Datensets nach dem Linked-Open-Data-Prinzip veröffentlicht werden, erzeugt dieser Prozess ein global miteinander verbundenes semantisches Netz, innerhalb dessen die Benutzer und Anwendungen identifiziert, selektiert, adaptiert, verwendet und wiederverwendet werden können.

Die Data-Cloud, so wie sie derzeit dasteht, sieht ziemlich lückenhaft aus, weil die Daten von Regierungen, Geschäften und Bibliotheken meist in getrennten, abgeschotteten Silos gespeichert und verwaltet werden (W3C, 2011). Wie der Bericht des W3C (2011) feststellt, „ist viel vom Inhalt der heutigen Linked-Data-Cloud das Ergebnis von spontanen, einmaligen Konversionen von öffentlich verfügbaren Datensets nach RDF und wird nicht regelmäßig auf Genauigkeit geprüft oder aktualisiert.“ Das bedeutet, dass viel mehr gemeinsame Bemühungen erforderlich wären, um dauerhafte und verlässliche Namespaces, Vokabulare (Ontologien) und Datensets zu entwickeln und zu verwalten. Mit einer richtigen Annahme der Linked-Data-Prinzipien würden Bibliotheken nicht nur Teil dieser Data Cloud, sondern auch die dominierenden Akteure sein, weil sie enorme Mengen an bibliografischen Daten und Normdaten haben, obwohl der Großteil dieses Reichtums derzeit in Silos und bibliotheksspezifischen Formaten eingesperrt ist. Einer der Wege, die Persistenz und Verlässlichkeit der Data Cloud sicherzustellen, ist es, bestehende Ontologien (Vokabulare) und URIs nachzunutzen. Einige der bereits bestehenden Namespaces sind RDFS, OWL, Dublin Core, VIAF, Friend of a Friend (FOAF), Schema.org, BLT (British Library Terms), Lexvo, GeoNames, MARC-Länder- und Sprachencodes, Dewey.info und RDF Book Mashup.

Übernahme des Ansatzes gemischter Metadaten

In Übereinstimmung mit der Behauptung, dass Ontologien (Taxonomien) und die Ansätze des Web 2.0 sich eher ergänzen als sich zu widersprechen (Gruber, 2007; Morville, 2005; Wright, 2007), ist es wichtig, dass jede Neukonzeptionalisierung eines Metadatenmodells das Beste aus diesen verschiedenen Paradigmen zu ihrem Vorteil nutzt. Beispielsweise macht die Verankerung von Kollaboration und Nutzerbeteiligung die OPACs zu einem Web-2.0-Dienst. Ganz ähnlich ergänzen sich die Technologien des Web 2.0 und 3.0 und könnten deshalb wirksam eingesetzt werden, um zufriedenstellende Benutzererfahrungen zu bieten - sowohl durch die bedeutende Menge an Daten, die aus dem erstgenannten gesammelt wurden, als auch durch die Struktur (der Modellierung) und die technologischen Möglichkeiten des letztgenannten (Gruber, 2007, 2008). Gruber (2007) argumentiert, dass sozial-erstellte Metadaten-Ansätze eindimensional und deshalb inkonsistent und unorganisiert sind, während standardbasierte Metadaten-Ansätze kritisiert werden, weil sie nicht das Benutzervokabular repräsentieren. Ebenso erklärt Morville (2005), dass das Social Web und das Semantic Web nebeneinander bestehen und sich gegenseitig formen. Diesen Punkt weiter ausführend stellt

Morville (2005) die Theorie des "Pace Layering" aus "How Buildings Learn" von Stewart Brand (Brand, 1994; Brand, 1999) in den Kontext und erörtert, dass „Taxonomien und Ontologien ein solides semantisches Netz bilden, das die Oberfläche mit der Infrastruktur verknüpft“, wohingegen Folksonomien über die taxonomische Metadaten-Infrastruktur gelegt sind und diese mit dem sich schnell verändernden und unberechenbaren Benutzervokabular beliefern (siehe auch Campbell & Fast, 2006; Smith, 2008). Indem sie sich von einem Paradigma hin zu dem anderen bewegen, würden Bibliotheken selektiv entscheiden, welche der Prinzipien, Standards und Formate in das nächste Paradigma übertragen werden und welche aufgegeben werden sollen. Beim Übergang von einem durch Bibliothekare kontrollierten (standardbasierten) Metadatenansatz hin zu einem community-erzeugten wird es als wichtig befunden, dass Bibliothekare die Barrieren für neue Benutzer abbauen, um diesen die Beteiligung an der Erzeugung von Metadaten zu ermöglichen, was auch mit einschließt, die strengen Metadatenkontrollen aufzugeben.

Vorteile von verlinkten Bibliotheksmetadaten für digitale Bibliotheken

Es gibt viele Vorteile einer Übernahme von Linked-Data-Prinzipien in die Bibliotheksstandards, von denen fünf wichtige nachfolgend angegeben und erläutert werden.

Dazu gehören:

- Offenheit der Metadaten und Freigabe
- Erleichterung des zufälligen Entdeckens von Informationsressourcen
- Verlinkte Metadaten zur Identifizierung von Ressourcennutzungsmustern, Zeitgeist und aufkommenden Metadaten
- Facetten-basierte Navigation
- Mit Links angereicherte Metadaten

Offenheit der Metadaten und Freigabe

Eine der Hauptkompetenzen der Web-2.0-Technologien ist deren Architektur, die die Mitwirkung erleichtert (O'Reilly, 2005), wodurch die Barrieren für Beitragende abgebaut wurden. Zentral bei dieser Beteiligungsarchitektur ist die Bedeutung des Sich-zueigen-machens von Freigabe und Zusammenarbeit, basierend auf gegenseitigem Vertrauen (Alexander, 2006; Anderson, 2006, 2010; Shirky, 2005; Tapscott & Williams, 2010; Udell, 2004; Weinberger, 2005, 2007). Das Web 2.0 hat deutlich mehr mit innerer Einstellung und Kultur zu tun als mit Technologie (Miller, 2005). In anderen Worten: das Web 2.0 als solches ist keine Neuerfindung; vielmehr ist es durch die Kultur des Teilhabens und Zusammenarbeitens charakterisiert, die um die Technologien des Web 1.0 herum gebaut wurde und die die Benutzer in die Lage versetzt, proaktive Inhaltsproduzenten und -konsumenten zu werden. In ihrem Bestseller „Wikinomics“ versichern Tapscott und Williams (2010), dass "wir aufgrund der tiefgreifenden Veränderungen der Technologie, der Demografie, der Wirtschaft und der Welt, in ein neues Zeitalter gelangen, in dem die Leute sich an der Wirtschaft beteiligen wie niemals zuvor." Sie identifizieren die Offenheit als eines der Hauptprinzipien, das der Zusammenarbeit der Masse zugrunde liegt, zusammen mit Peering (selbst organisierte Peer-Netzwerke der Beitragenden), Freigabe und globalem Handeln, wobei jedes die Möglichkeiten vergrößert, einen noch viel größeren Talentpool anzuzapfen. Sie bemerken auch, dass „Offenheit mit Aufrichtigkeit, Transparenz, Freiheit, Flexibilität, Ausdehnung, Engagement und Zugang in Verbindung steht“. Nach ihnen deuten

die aktuellen ökonomischen, sozialen und technologischen Trends darauf hin, dass Offenheit nicht notwendigerweise die Verletzung des geistigen Eigentums bedeutet.

Weiterhin verfechten sie die These, dass die Kultur der Offenheit und der fortgesetzten Anerkennung ihrer potentiellen Vorteile „die konventionelle Weisheit [beeinträchtigt hat], die besagt, dass Firmen im Wettbewerb stehen, indem sie ihre begehrtesten Ressourcen nahe bei der Brust tragen“ (Tapscott & Williams, 2010).

Alexander (2006) behauptet, dass der multi-direktionale Informationsfluss zwischen Produzenten und Konsumenten und über Domänen, Server und Maschinen hinweg, es erforderlich macht, die Informationssilos zu öffnen und verteilte Dienstleistungen zu fördern. Der Verfasser behauptet, dass sogar kommerzielle Sites, wie zum Beispiel Amazon.com ihren Nutzern erlauben, die „ISBN-Nummern aus ihren Listen zu harvesten“. Er bestätigt, dass „Offenheit ein Kennzeichen dieser aufkommenden Bewegung ist, sowohl ideologisch als auch technologisch“. Auf Bibliotheken übertragen, argumentiert Miller (2005), dass die Prinzipien des Web 2.0 sich beim Gedanken an die durchdachten Daten behaupten, was im Gegenzug den Daten erlaubt, auf vielerlei Arten freigelegt, entdeckt und manipuliert zu werden und dabei unvorstellbare Möglichkeiten für neue Zwecke und zur Weiterverwendung der Daten erzeugt. Während Linked Data benutzbar gemacht werden können, ohne dass sie notwendigerweise offen sein müssen (Cobden, Black, Gibbins, Carr, & Shadbolt, 2011; Shadbolt, 2010; W3C, 2011), bringt die Öffnung der Daten der Gesellschaft viele Vorteile, wie Berners-Lee (2010) hervorhebt. In seinem TedTalk wiederholt Berners-Lee (2010) seine Vision von Linked Data, indem er eine Anzahl internationaler, regionaler und community-basierte Initiativen und Projekte zitiert, die die Linked-Data-Prinzipien angenommen haben und deshalb ihre Daten offen verfügbar machen unter Verwendung von URIs und RDF-Technologien. Wie der Verfasser hervorhebt, können offene Daten in einer unvorstellbaren Zahl von Arten wieder verwendet werden. Offene Regierungsdaten ermöglichen es den Steuerzahlern beispielsweise zu prüfen, wo und für welche Zwecke ihr Geld von den durch sie gewählten Repräsentanten ausgegeben wurde. Wie Berners-Lee (2010) betont, hat das Moment zum Öffnen der Daten „gerade erst begonnen“.

Das Konzept Linked Open Data hat weitreichende Konsequenzen für die Art, wie Metadaten in Bibliotheken erzeugt werden, wie auf sie zugegriffen wird, wie sie freigegeben und neu kombiniert werden. Verlinkte Bibliotheksmetadaten können „frei nutzbar, wiederverwendbar und weiter verteilbar sein [...]; durch die Verwendung von weltweit eindeutigen Identifikatoren zur Kennzeichnung von Werken, Orten, Leuten, Ereignissen, Themen und sonstigen Interessensgegenständen oder Konzepten werden Bibliotheken es ermöglichen, dass Ressourcen über ein breites Spektrum von Datenquellen zitiert werden und so ihre Metadaten besser zugänglich machen.“ Die Übernahme von Linked-Data-Prinzipien und das Zur-Verfügung-Stellen von Metadaten für die Weiterverwendung durch verlässliche Quellen würde die unnötige Duplizierung von Metadaten eliminieren, die bereits irgendwo vorhanden sind. Sie ermöglicht mehreren Bibliotheken, ihre Ressourcen miteinander zu verlinken, und erleichtert damit die Zusammenarbeit. Und schließlich lässt Linked Data die Sichtbarkeit von Bibliotheksmetadaten durch Links erwarten, während es gleichzeitig erlaubt, mit nicht-bibliothekarischen Informationsquellen wie Google, Wikipedia, LibraryThing, CiteULike und Amazon verlinkt zu sein.

Erleichterung des zufälligen Entdeckens von Informationsressourcen

Eine bedeutende Anzahl von bemerkenswerten wissenschaftlichen Innovationen sind, wie berichtet, das Ergebnis des zufälligen Entdeckens durch Wissenschaftler (Campa, 2008; Stoskopf, 2005; Zuckerman, 2010). Solche unerwarteten und nicht vorhersehbaren Entdeckungen könnten freudige Überraschungen für den Wissenschaftler darstellen. Wie

Stoskopf (2005) bemerkt, passieren solche Entdeckungen häufiger dem offenen und neugierigen Beobachter als dem Wissenschaftler, der solche Hinweise als Ablenkung betrachtet. Er behauptet, dass "eine Grundneugier und Beobachtung die notwendigen Wegbereiter für wissenschaftliche Entdeckungen sind. Es sollte anerkannt werden, dass Zufallsentdeckungen von signifikantem Wert beim Vorankommen der Wissenschaft sind und oft das Fundament für wichtige intellektuelle Sprünge beim Verständnis darstellen" (Stoskopf, 2005). Der Term Serendipity wurde 1754 vom englischen Historiker Horace Walpole geprägt (Merton & Barber, 2004). In einem seiner sechshundvierzig Briefe an seinen Freund Horace Mann schrieb Walpole:

"Ich sollte mich bemühen, Ihnen zu erklären: Sie werden es besser durch die Ableitung verstehen als durch die Definition. Einst las ich ein albernes Märchen namens „The Three princes of Serendip“. Während ihre Hoheiten reisten, machten sie durch Zufall und Klugheit ständig Entdeckungen von Dingen, nach denen sie nicht forschten. Zum Beispiel entdeckte einer, dass ein auf dem rechten Auge blinder Esel kurz zuvor dieselbe Route gegangen war, weil das Gras nur auf der linken Seite abgefressen war, obwohl es dort schlechter war als auf der rechten Seite" (Merton & Barber, 2004).

Das zufällige Entdecken von Informationen ist also auch relevant als Forschungsthema in der Bibliotheks- und Informationswissenschaft, obwohl die Forschung über dieses Thema sehr ungenügend ist (Foster & Ford, 2003). Wie Foster und Ford (2003) feststellen, enthalten die aktuellen Modelle zum Benutzerverhalten bei der Informationssuche (wie zum Beispiel Ingwersen, 1996; Wilson und Walsh, 1996; Wilson, 1997; Kuhlthau, 1993; Saracevic, 1996 & Spink, 1997) keinerlei Betrachtungen von Serendipität. Wie Foster und Ford (2003) berichten, ist der naheste Bezug auf den Begriff der von Wilson und Walsh, 1996, und Wilson, 1997, die beide das Phänomen als „passive Aufmerksamkeit“ und „passive Suche“ behandeln. Foster und Ford (2003) untersuchen das Thema detaillierter und behaupten, dass zufälliger Zugriff auf Informationen „die bestehende Problemwahrnehmung oder -lösung des Forschers“ bestärkt oder verfestigt und ihm dazu verhilft, neue Territorien auf der Suche nach Erforschung und Entdeckung auszuwerten. Während sie die Bedeutung für das vollständige Verständnis des Informationssuchverhaltens von Benutzern hervorheben, erwähnen Foster und Ford (2003) die Probleme beim Untersuchen von Serendipität, die zum Teil den Schwierigkeiten im Hinblick auf eine systematische Kontrolle und zum Teil als Vorhersage dessen, wie sie die Benutzererfahrungen verbessert, angelastet werden können. In dem Maße wie Bibliotheksdienste online gehen, verlassen sich die Benutzer darauf, Datenbanken zu durchsuchen, um Zugang auf vorher festgelegte Informationsquellen zu erhalten, und sie verlieren die Möglichkeit für eine Zufallsentdeckung einer Informationsquelle, die sie gefunden hätten, wenn sie ein physisch vorhandenes Bibliotheksregal durchstöbert hätten (Foster & Ford, 2003; Massis, 2011). Es muss festgestellt werden, dass Bibliotheksdatenbanken mit voreingestellter Suchoberfläche sehr wahrscheinlich keine Möglichkeit für zufällige, sachbezogene Suchergebnisse anbieten, da das nicht die Art ist, wie sie konzipiert sind.

Der freie Zugang zu physisch vorhandenen Bibliotheksregalen macht das zufällige Entdecken eines Buchs oder einer Zeitschrift möglich. Die Bibliotheken, die keinen Zugang zu den Bücheregalen erlauben, bieten sehr wenig im Hinblick auf solch zufällige Entdeckungen von Büchern in den Regalen. OPACs, die dafür konzipiert sind, vorher festgelegte Suchanfragen zu beantworten (z. B. unter Verwendung des Verfassernamens oder des Titels) sind analog zu den Praktiken des verwehrten Zugriffs auf physisch vorhandene Bücherregale, weil sich beide darauf beziehen, den Suchanfragen der Benutzer mit Datensätzen zu begegnen, die in den Beständen vorhanden sind, wenn sie auf das Streben eines Benutzers nach einer Informationsquelle antworten. Um es einfach zu sagen, das Element der erfreulicherweise

überraschenden Entdeckung, das durch physisch vorhandene Bücherregale realisiert werden könnte, könnte komplett verschwinden, wenn digitale Bibliotheken sich nur auf passende Benutzeranfragen konzentrieren würden.

Wenn das so ist, stellt sich die Frage „wie kann die Serendipität in die Konzeption von digitalen Bibliotheken einfließen?“ Da der Zugang zu Informationsobjekten in digitalen Bibliotheken von der Verwendung von Metadaten abhängig ist, liegt ein Teil der Lösung in der Wahl des Metadatenansatzes. Es ist daher wichtig zu untersuchen, wie gemeinschaftlich-gebildete Metadatenansätze gemeinsam mit vorhandenen Bibliotheksmetadaten verwendet werden können, um Benutzererfahrungen durch zufällige Entdeckungen zu verbessern. Ebenso muss die Rolle von Linked Data bei der Unterstützung von Serendipität untersucht werden. Durch die Veröffentlichung von Linked Data würden Bibliotheken von der Möglichkeit profitieren, zu Unmengen relevanter Datenbanken und Quellen zu verlinken und so das Zusammentragen von verschiedenartigen Informationsobjekten aus ganz verschiedenen Quellen zu erlauben. Die Verwendung von URIs hilft beim Zusammenfügen solcher Ressourcen aus heterogenen Quellen. Dem Benutzer zu erlauben, nahtlos von einer Quelle zu einer anderen zu springen, führt folglich den zufälligen Zugang zu digitalen Bibliotheken wieder ein, in der gleichen Art und Weise wie frei zugänglich Bibliotheksregale dies den Benutzern bieten.

Verlinkte Metadaten zur Identifizierung von Ressourcennutzungsmustern, Zeitgeist und aufkommenden Metadaten

Je mehr ein Metadatenystem die Variationen und Prägungen in der Nomenklatur von Objekten reflektiert (z. B. Tsunami, Tschechische Republik für die frühere Tschechoslowakei oder Russland für die frühere Sowjetunion oder Mac für Macintosh), desto besser unterstützt es die Auffindbarkeit von Informationsobjekten. Das ist so, weil „das Vokabular des Informationsfachmanns oder des Bibliothekars im Konflikt mit der Sprache des Inhalts des Verfassers und der Benutzer stehen kann und damit die eigentliche Information verschleiert, welche die Taxonomie offenlegen sollte“ (Barbosa, 2008). Da es die Hauptziele des Anbringens von Metadaten an Informationsobjekte sind, deren Auffindbarkeit zu ermöglichen, sollten Metadaten nach der Relevanz für die Erreichung dieses Ziels und weniger nach der Art und Weise beurteilt werden, wie sie ein Informationsobjekt auf objektive und ontologische Art beschreiben. Terminologien ändern sich im Laufe der Zeit, deshalb sollten die Metadaten fähig sein, diese Änderungen nachzuvollziehen und sie widerzuspiegeln. Im Zeitalter der gedruckten Kataloge gaben die Bibliotheken ihr Bestes, um die Aktualität ihres Zettelkatalogs zu aktualisieren und zu verwalten. Die Implikationen im Hinblick auf Ressourcen für die Aktualisierung von Datensätzen bedeuteten, dass selbst wenn neue Terminologien und Nomenklaturen eingeführt und den Bibliothekaren bewusst gemacht wurden, die notwendigen Aktualisierungen eine gewisse Zeitspanne lang verzögert sein mussten. Dieses Problem haben nicht nur die Bibliothekare, weil dasselbe auch gilt, wenn ein sachlicher Fehler in einem gedruckten Enzyklopädie-Eintrag korrigiert wird.

Ein digitales Bibliothekssystem, das auf gemischten Metadatenansätzen (standardbasierte und gemeinschaftlich erzeugte Metadaten) aufbaut und durch zusätzliche Semantik und technische Fähigkeiten von Linked Data (URI, RDF, OWL, SPARQL) angereichert wird, wäre besser aufgestellt, um reichhaltigere und vollständigere Metadatenbeschreibungen von Informationsobjekten zu liefern und genauer die verschiedenen Interpretationen und Terminologien widerzuspiegeln, die von Benutzern verwendet werden. Zusätzlich würde Linked Data dazu beitragen, die Möglichkeit einer Analyse der Verwendungsmuster von Ressourcen (wie zum Beispiel Anzahl der Klicks, Summe der Downloads, häufig verwendete Tags (Tag Clouds), durchschnittliche Ratings und oft empfohlene Ressourcen) zu erleichtern.

LibraryThing.com bietet zurzeit beispielsweise Funktionen wie Zeitgeist „der Geist der Zeit“ und präsentiert Analysen der beliebtesten Bücher aufgrund der Durchschnittsbewertungen, der Tag Clouds und der am häufigsten rezensierten Bücher.

Linked Data als Datenmodell und Set an technologischen Rahmen steht nicht für sich alleine und hat ohne die zugrundeliegenden Daten keinen Sinn. Deshalb ist die Einbindung sowohl von vorhandenen Daten, die von Bibliothekaren erzeugt werden, und gemeinschaftlich produzierten Metadaten, die von den Benutzern generiert werden, entscheidend, weil beide gleichermaßen wichtig sind. Die Einbindung der letztgenannten wird besonders relevant im Hinblick auf die Notwendigkeit, Verwendungsmuster und Zeitgeist zu identifizieren. Weinberger (2005) behauptet, dass „eine Autorin eine Autorität ist, wenn es darum geht, zu sagen, welche Intention ihr Werk hat, aber nicht, wenn es darum geht, was es anderen bedeutet. Wenn es um die Suche geht, ist es weitaus wichtiger, was das Werk dem Suchenden bedeutet als welche Intentionen die Verfasserin hatte.“ Die Modellierung und die Technologie von Linked Data würden substanziell zum Erfolg von Aktivitäten beitragen, die darauf abzielen, Ressourcennutzungsmuster, den Zeitgeist und neu aufkommendes Benutzervokabular zu identifizieren. Dies könnte den Bibliotheken helfen, strategisch richtige Entscheidungen beim Erwerb von Ressourcen (Bücherkauf, Zeitschriftenabonnements) zu treffen, was wiederum diese Investitionen rechtfertigen würde.

Facetten-basierte Navigation

Facettierte Browsen bezieht sich auf die Verwendung unterschiedlicher Dimensionen und Propertyts bei der Navigation in einem Informationssystem mit dem Ziel, die Auffindbarkeit zu verbessern (Morville, 2005). Nach Morville (2005) wurde die Idee, Facetten in Bibliotheken zu verwenden, von S. R. Ranganathan populär gemacht. Die Verwendung von Facetten steht der Verwendung von Taxonomien gegenüber, wobei ein Stück nur zu einem einzigen Ort in der Hierarchie gehen kann und damit alternative Wege der Navigation beschränkt. In Übereinstimmung mit Ted Nelsons Begriff der „Intertwingularity“ (Verzweigtheit) argumentiert Weinberger (2007) gegen die Vorstellung, eine einzelne, vordefinierte und perfekte Navigationsstruktur anzustreben. Er schreibt: „die Leute behaupten ständig, dass sie Dinge zutiefst hierarchisch, kategorisierbar und fortlaufend machen können, wenn sie es gar nicht können. Alles ist weit miteinander verzweigt.“ Der von Ted Nelson (dem man auch den Term Hypertext zuschreibt) geprägte Term Intertwingularity bezieht sich auf das Miteinander-in-Beziehung-Stehen und das Voneinander-in-Abhängigkeit-Sein von Ideen (Begriffen), der folglich behauptet, dass jeglicher Versuch einer Kategorisierung künstlich ist (Für weiteres über Ted Nelsons Intertwingularity siehe Weinberger, 2007; Wright, 2007). Digitale Informationsobjekte, so disputiert Weinberger, könnten sortiert, kategorisiert und in einer Vielfalt von Arten, die auf den Benutzeranforderungen und dem Kontext basieren, präsentiert werden (Weinberger, 2007). Um Weinberger (2007) vollständig zu zitieren:

“In der dritten Ordnung der Ordnung [in Volltext-Bibliotheken] kann ein Blatt [ein Informationsobjekt] an vielen Ästen hängen, es kann für unterschiedliche Leute an unterschiedlichen Ästen hängen, und es kann die Äste wechseln für dieselbe Person, wenn diese entscheidet, anders auf das Thema zu schauen. Es ist nicht so, dass unser Wissen der Welt eine Form annimmt, die kein Baum ist oder dass es zu einem unvorstellbaren, vierdimensionalen Baum wird. In der dritten Ordnung der Ordnung hat Wissen keine Form. Es gibt einfach zu viele nützliche, machtvolle und schöne Arten, sich unsere Welt zusammenzureimen.“

In seinem Buch “Too Big to Know” (2012) behauptet Weinberger, dass es die “Lösung des Problems des Informationsüberflusses ist, mehr Metadaten zu erzeugen”, während er die

Bedeutung dessen anzeigt, relevante Ergebnisse basierend auf unterschiedlichen Facetten herauszufiltern. In seinem Buch "The Laws of Simplicity" behauptet Maeda (2006), dass "Einfachheit bedeutet, das Offensichtliche zu subtrahieren und das Bedeutsame zu addieren." Er schreibt "mehr erscheint weniger, indem man es einfach sehr weit weg bewegt. Demzufolge ist eine Erfahrung einfacher gemacht, wenn man das Ergebnis lokal hält und das tatsächliche Werk zu einem weit entfernten Ort bewegt."

Im Prinzip sind digitale Bibliotheken frei von den Beschränkungen, die das physisch existierende Regal in konventionellen Bibliotheken hat, deshalb ermöglichen sie den Benutzern, durch die Verwendung von Stichwörtern und Kategorien wie Facetten beim Retrieval von Informationsobjekten die Informationen in vielfältiger Weise zu filtern. Durch die Mächtigkeit der Verlinkungsstruktur von Linked Data und dem Metadatenreichtum, der durch die Verwendung sowohl der standardbasierten als auch der sozial-erstellten Metadatenansätze erreicht wird, würde eine Vielfalt von Facetten zum Filtern des Inhalts, der den Benutzern präsentiert wird, verwendet werden. Ein solcher Metadatenreichtum kann auch Benutzerprofile verwenden, so dass digitale Bibliothekssysteme in der Lage wären, Suchergebnisse in einen Kontext zu stellen und auf die Benutzer zuzuschneiden. Aktuell sind Last.fm, LibraryThing und Amazon.com Beispiele dafür, wo facettierte Navigation und Metadatenfilter zur Anwendung kommen.

Mit Links angereicherte Metadaten

Derzeit können die Links zu Metadaten in einer OPAC-Anzeige nur bis zu einem bestimmten Punkt weiter verfolgt werden und enden bald in einer Sackgasse, wenn bestimmte Daten nicht weiter verlinkt sind. Die aktuellen Metadatenoberflächen erlauben es einem Benutzer beispielsweise nicht, nach einem bestimmten Buch zu suchen, dann auf den Verfassernamen zu klicken und dessen Publikationen zu sehen sowie vielleicht auch ähnliche Bücher zu finden. Normalerweise bieten sie keinen nahtlosen Link zur biografischen Seite des Verfassers oder von hier zu einer anderen Seite, wie zum Beispiel zu Wikipedia, Google Books oder zurück zum Katalog usw. Ein umfangreich beschriebenes Buch in einer Linked-Data-Umgebung könnte mit allem verlinkt sein, mit dem es in Beziehung steht; es würde am Benutzer liegen, zu entscheiden, entweder die Navigation zu beenden, wenn sein Informationsbedürfnis befriedigt ist, oder ob er Anwendungen möchte, die Kontext und andere relevante Informationen liefern. In anderen Worten, jedes einzelne Metadatenelement wird mit einem Link angereichert, außer wenn solche Metadaten aus Literalen bestehen, wie das Alter der Person, wenn es nicht notwendig ist, dies mit einer anderen Seite zu verlinken.

Metadaten, die mit Kontext und anderen relevanten Links angereichert sind, würden es den Benutzern erlauben, nahtlos zwischen verschiedenen Bibliotheksdatenbanken und externen Informationsanbietern wie anderen Bibliotheken und Suchmaschinen zu navigieren. Durch die weltweit einheitlich identifizierten Entitäten (wie Werke, Personen, Orte und Ereignisse), Metadatenelemente bzw. Propertyts (Autor, Titel, Thema, Beziehungen) und die entsprechenden Werte (Instanzen) bietet Linked Data eine Vielfalt von Wegen, Informationsobjekte mit Metadaten anzureichern, was das Informationsretrieval erleichtern und die Benutzererfahrungen in digitalen Bibliotheken verbessern würde.

Schlußfolgerungen

Aktuelle Bibliotheksstandards und Modelle haben einige der Beschränkungen fortgeführt, die aus dem traditionellen System des Kartenkatalogs geerbt wurden. Die Metadaten, die unter Verwendung dieser Modelle generiert werden, sind hauptsächlich auf menschliche und weniger auf die maschinelle Verarbeitung abgestimmt. Linked Data wird als Lösung

angesehen, um datenzentrierte und mit Maschinen zu verarbeitende Metadaten zu ermöglichen. Allerdings steckt die Einführung von Linked Data bei der derzeitigen Erzeugung und Verwendung von Metadaten noch in den Kinderschuhen.

Es wird vorgeschlagen, dass bestehende Bibliotheksmodelle wie die RDA und FRBR im Hinblick darauf, sie zu den Linked-Data-Prinzipien kompatibel zu machen, neu konzeptioniert werden. Und Bibliotheken sollten bestehende, MARC-kompatible Metadatensätze in RDF/XML-Formate konvertieren und dabei langsam von MARC und seinen Derivaten wegkommen. Im Hinblick darauf wird erläutert, dass der Ruf nach einem Aufgeben von Metadatenmodellen (wie RDA und FRBR) gegen das Ethos von Linked Data ist, das es unterschiedlichen Standards und kontrollierten Vokabularen erlaubt, „glücklich“ nebeneinander im selben Bereich zu existieren, vorausgesetzt dass diese Ressourcen (Objekte) unter Verwendung von weltweit eindeutigen URIs referenziert sind und dass ein High-Level-Datenmodell angewendet wird, das als RDF bekannt ist. In Verbindung damit sollten in jeder Initiative zur Übernahme der Linked-Data-Prinzipien Jahrhunderte alte Best Practices und Erfahrungen, die von Bibliotheken erlernt wurden, während sie kontrollierte Vokabulare (wie zum Beispiel Normdateien für Namen und Schlagworte) entwickelt, gepflegt und bewahrt haben, zum Vorteil genutzt werden. Die richtige Übernahme von Linked Data für Bibliotheksmetadaten, so wird vorgeschlagen, würde der Bibliothek helfen, neue Bibliotheksanwendungen und -dienste wie facettiertes Browsen, zufälliges Browsen, Metadaten-Zeitgeist und die Identifizierung von neu aufkommenden Trends und Benutzervokabular zu entwickeln.

References

- Alemu, G., Stevens, B., & Ross, P. (2012). Towards a Conceptual Framework for User-Driven Semantic Metadata Interoperability in Digital Libraries: A Social Constructivist Approach. *New Library World*, 113(1/2), 38-54.
- Alexander, B. (2006). Web 2.0: A New Wave of Innovation for Teaching and Learning? *EDUCAUSE Review*, 42(2), 32-44.
- Allemnag, D., & Hendler, J. (2008). *Semantic Web for the Working Ontologist: Effective Modeling in RDFS and OWL*. Amsterdam: Morgan Kaufmann.
- Anderson, C. (2006). *The Long Tail: How Endless Choice is Creating Unlimited Demand*. London: Random House Business Books.
- Anderson, C. (2010). *Free : How Today's Smartest Businesses Profit by Giving Something for Nothing*. London: Random House Business Books.
- Barbosa, D. (2008). Taxonomy Folksonomy Cookbook. Retrieved from http://solutions.dowjones.com/cookbook/ebook_sla2008/cookbookebook.pdf
- Berners-Lee, T. (1997). Metadata Architecture. Retrieved from <http://www.w3.org/DesignIssues/Metadata.html>
- Berners-Lee, T. (1998). What the Semantic Web can Represent. Retrieved from <http://www.w3.org/DesignIssues/RDFnot.html>
- Berners-Lee, T. (2007). Testimony of Sir Timothy Berners-Lee Before the United States House of Representatives Committee on Energy and Commerce Subcommittee on Telecommunications and the Internet. Retrieved from <http://dig.csail.mit.edu/2007/03/01-ushouse-future-of-the-web.html>
- Berners-Lee, T. (2009). Linked Data. Retrieved from <http://www.w3.org/DesignIssues/LinkedData.html>
- Berners-Lee, T. (2010). Tim Berners-Lee: The Year Open Data Went Worldwide. from http://www.ted.com/talks/tim_berners_lee_the_year_open_data_went_worldwide.html
- Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The Semantic Web: A New form of Web Content that is Meaningful to Computers will Unleash a Revolution of New Possibilities. *The Scientific American*, 284(5), 34-43

- Brand, S. (1994). *How Buildings Learn: What Happens after They're Built*. London: Viking.
- Brand, S. (1999). *The Clock of the Long Now: Time and Responsibility* (1st ed.). New York: Basic Books.
- Campa, R. (2008). Making Science by Serendipity. A review of Robert K. Merton and Elinor Barber's *The Travels and Adventures of Serendipity*. *Journal of Evolution and Technology* 17(1), 75-83.
- Campbell, D. G., & Fast, K. V. (2006). *From Pace Layering to Resilience Theory: The Complex Implications of Tagging for Information Architecture*. Paper presented at the The 7th Information Architecture Summit, Vancouver, British Columbia, Canada.
- Cobden, M., Black, J., Gibbins, N., Carr, L., & Shadbolt, N. (2011). *A Research Agenda for Linked Closed Data*. Paper presented at the Second International Workshop on Consuming Linked Data. Retrieved from <http://eprints.ecs.soton.ac.uk/22711/3/position.pdf>
- Coyle, K. (2010). Library Data in a Modern Context. *Library Technology Reports* 46(1), 5-13.
- Coyle, K., & Hillmann, D. (2007). Resource Description and Access (RDA): Cataloging Rules for the 20th Century. *D-Lib Magazine*, 13(1/2).
- Day, M. (2000). Resource Discovery, Interoperability and Digital Preservation: Some Aspects of Current Metadata Research and Development. *VINE*, 36(117), 35-48.
- Day, M. (2003a). *Integrating Metadata Schema Registries with Digital Preservation Systems to Support Interoperability: A Proposal*. Paper presented at the International Conference on Dublin Core and Metadata Applications, Seattle.
- Day, M. (2003b). *Preservation Metadata Initiatives: Practicality, Sustainability, and Interoperability*. Paper presented at the ERPANET Training Seminar on Metadata in Digital Preservation. Retrieved from <http://www.ukoln.ac.uk/preservation/publications/erpanet-marburg/day-paper.pdf>
- Decker, S., Melnik, S., Van Harmelen, F., Fensel, D., Klein, M., Broekstra, J., et al. (2000). The Semantic Web: The Roles of XML and RDF. *IEEE Internet Computing*, 15(3), 63-74.
- Denton, W. (2007). FRBR and the History of Cataloging. In A. G. Taylor (Ed.), *Understanding FRBR: What it is and How it Will affect Our Retrieval* (pp. 35-57). Westport, Connecticut: Libraries Unlimited.
- Doerr, M., Gradmann, S., Hennicke, S., Isaac, A., Meghini, C., & Van de Sompel, H. (2012). *The Europeana Data Model (EDM)*. Paper presented at the World Library and Information Congress: 76th IFLA General Conference and Assembly, Gothenburg, Sweden.
- Dunsire, G. (2009). *UNIMARC, RDA and the Semantic Web* Paper presented at the World Library and Information Congress: 75th IFLA General Conference and Council Retrieved from <http://www.ifla.org/files/hq/papers/ifla75/135-dunsire-en.pdf>
- Evans, W. (2009). *Building Library 3.0: Issues in Creating A Culture Of Participation*. Oxford: Chandos.
- Foster, A., & Ford, N. (2003). Serendipity and Information Seeking: an Empirical Study. *Journal of Documentation*, 59(3), 321-340.
- Gruber, T. (1993). Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal Human-Computer Studies*, 43, 907-928.
- Gruber, T. (2007). Ontology of Folsonomy: A Mash-up of Apples and Oranges. *International Journal on Semantic Web & Information Systems*, 3(2).
- Gruber, T. (2008). Collective Knowledge Systems: Where the Social Web Meets the Semantic Web. *Journal of Web Semantics: Science, Services and Agents on the World Wide Web*, 6(1), 4-13.
- Guenther, R., & McCallum, S. (2003). New metadata standards for digital resources: MODS and METS. *Bulletin of the American Society for Information Science and Technology*, 29(2).
- Haslhofer, B., & Isaac, A. (2011). *data.europeana.eu - The Europeana Linked Open Data Pilot*. Paper presented at the DC-2012 International Conference on Dublin Core and Metadata Applications, The Hague, The Netherlands.
- Helen, W. (2010). Linked Data and Libraries. *Catalogue & Index*(160), 2-5.
- Horridge, M., Knublauch, H., Rector, A., Stevens, R., & Wroe, C. (2004). *A Practical Guide to Building OWL Ontologies Using the Protégé-OWL Plugin and CO-ODE Tools Edition 1.0* Available from <http://www.co-ode.org/resources/tutorials/ProtegeOWLTutorial.pdf>
- IFLA. (2009). *Functional Requirements for Bibliographic Records: Final Report*: International Federation of Library Associations and Institutions.

- Lagoze, C. (2010). *Lost Identity: the Assimilation of Digital Libraries into the Web*. Cornell University, School of Information Science.
- Library of Congress. (2011). *A Bibliographic Framework for the Digital Age*. Washington, D.C.: Library of Congress.
- Lubetzky, S. (1953). Development of Cataloging Rules *Library Trends* 2(2), 179-186.
- Marcum, D. (2011). *A Bibliographic Framework for the Digital Age*. Washington, D.C.: Library of Congress.
- Massis, B. E. (2011). "Serendipitous" browsing versus library space. *New Library World*, 112(3/4), 178-182.
- Mathes, A. (2004). Folksonomies - Cooperative Classification and Communication Through Shared Metadata Retrieved from <http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html>
- Merton, R. K., & Barber, E. (2004). *The Travels and Adventures of Serendipity: A Study in Sociological Semantics and the Sociology of Science*. Princeton: Princeton University Press.
- Miller, P. (2005). Web 2.0: Building the New Library *Ariadne*, 45.
- Morville, P. (2005). *Ambient findability*. Sebastopol, CA: O'Reilly.
- Nilsson, M. (2010). *From Interoperability to Harmonization in Metadata Standardization: Designing an Evolvable Framework for Metadata Harmonization*. Unpublished PhD, School of Computer Science and Communication, KTH, Stockholm.
- Noy, N. F., & McGuinness, D. L. (2000). Ontology Development 101: A Guide to Creating Your First Ontology. Retrieved from http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html
- O'Reilly, T. (2005). What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. Retrieved from <http://oreilly.com/web2/archive/what-is-web-20.html>
- Rothenberg, J. (2008). Interoperability as a Semantic Cross-cutting Concern *Interoperabiliteit: Eerlijk zullen we alles delen*. Den Haag.
- Shadbolt, N. (2010). *data.gov.uk - The Linked Data Revolution*. Paper presented at the Innovating Through Information Lecture Series. Retrieved from <http://eprints.ecs.soton.ac.uk/18787/>
- Shirky, C. (2005). Ontology is Overrated: Categories, Links, and Tags. *Clay Shirky's Writings About the Internet*. Retrieved from http://www.shirky.com/writings/ontology_overrated.html
- Smith, G. (2008). *Tagging : People-Powered Metadata for the Social Web*. Berkeley, CA: New Riders.
- Stoskopf, M. K. (2005). Observation and cogitation: how serendipity provides the building blocks of scientific discovery. *Institute for Laboratory Animal Research*, 46(4), 332-337.
- Styles, R. (2009). Bringing FRBR Down to Earth. <http://dynamicorange.com/2009/11/11/bringing-frbr-down-to-earth/>
- Styles, R., Ayers, D., & Shabir, N. (2008). *Semantic MARC, MARC21 and the Semantic Web*. Paper presented at the Linked Data on the Web (LDOW2008). Retrieved from <http://events.linkeddata.org/ldow2008/papers/02-styles-ayers-semantic-marc.pdf>
- Tapscott, D., & Williams, A. D. (2010). *Wikinomics: How Mass Collaboration Changes Everything* (Expanded ed.). London: Penguin Books.
- Tennant, R. (2002). MARC must die. *Library Journal*. Retrieved from <http://www.libraryjournal.com/article/CA250046.html>
- The British Library. (2011). Free Data Services. Retrieved December 14, 2011, from <http://www.bl.uk/bibliographic/datafree.html>
- Udell, J. (2004). Collaborative Knowledge Gardening with Flickr and del.icio.us, Social Networking goes Beyond Sharing Contacts and Connections. *InfoWorld*. Retrieved from <http://www.infoworld.com/d/developer-world/collaborative-knowledge-gardening-020>
- Veltman, K. H. (2001). Syntactic and Semantic Interoperability: New Approaches to Knowledge and the Semantic Web. *New Review of Information Networking*, 7(1), 159-183.
- W3C. (2004a). OWL Web Ontology Language Overview. from <http://www.w3.org/TR/owl-features/>
- W3C. (2004b). RDF Primer. *W3C Recommendation*. Retrieved from <http://www.w3.org/TR/2004/REC-rdf-primer-20040210/#basicconcepts>
- W3C. (2011). *Library Linked Data Incubator Group Final Report*: W3C.
- Wallis, R. (2011a). Library of Congress to Boldly Voyage to Linked Data Worlds. <http://consulting.talis.com/2011/11/library-of-congress-to-boldly-voyage-to-linked-data->

- [worlds/](#)
- Wallis, R. (2011b). Will Europe's National Libraries Open Data In An Open Way? Retrieved from <http://dataliberate.com/2011/09/will-europes-national-libraries-open-data-in-an-open-way/>
- Weinberger, D. (2005). Tagging and Why it Matters. Retrieved from <http://cyber.law.harvard.edu/sites/cyber.law.harvard.edu/files/07-WhyTaggingMatters.pdf>
- Weinberger, D. (2007). *Everything is Miscellaneous*. New York: Times books.
- Weinberger, D. (2012). *Too Big to Know: Rethinking Knowledge Now That the Facts Aren't the Facts, Experts Are Everywhere, and the Smartest Person in the Room Is the Room*. New York: Basic Books
- Wilson , N. (2010). *Linked Data Prototyping at the British Library*. Paper presented at the Talis Linked Data and Libraries event. Retrieved from <http://blogs.talis.com/nodalities/neil-wilson-lightning-talk-%E2%80%93-linked-data-and-libraries-2010>
- Wright, A. (2007). *Glut: Mastering Information Through the Ages*. Ithaca: Corenell University Press.
- Zuckerman, E. (2010). Desperately Seeking Serendipity. Retrieved from <http://www.chi2011.org/program/plenaries.html>

Über die Autoren

Getaneh Alemu ist Doktorand an der School of Creative Technologies und Dr. Brett Stevens (Principal Lecturer, School of Creative Technologies), Dr. Penny Ross (Senior Lecturer, School of Computing) sowie Jane Chandler (Principal Lecturer, Faculty of Creative and Cultural Industries), die alle an der University of Portsmouth, Portsmouth, UK angesiedelt sind. Getaneh Alemu ist der Korrespondenzautor, der unter der Adresse Getaneh.Alemu@port.ac.uk erreicht werden kann.