

# DIE ROLLE DES SEMANTIC WEB FÜR BIBLIOTHEKEN: LINKED OPEN DATA UND MEHR: WELCHE STRATEGIEN KÖNNEN HIER DIE BIBLIOTHEKEN IN DIE ZUKUNFT FÜHREN?

Philipp Zumstein

Universitätsbibliothek Trier / Bibliotheksakademie Bayern

zuphilip@gmail.com

---

## 1. Einleitung

Heutzutage haben wir weniger das Problem, dass Informationen nicht im Internet abrufbar sind, sondern es existiert ein Überfluss an Informationen (sogenannter 'information overload'), z.B. wird die Anzahl der momentan von Suchmaschinen indizierten Webseiten auf mindestens 10 Milliarden geschätzt.<sup>1</sup> Zusätzlich kommt hinzu, dass viele Webauftritte in sich abgeschlossen sind oder nur unspezifische und nur aus dem Kontext interpretierbare Links nach außen haben. Elektronische Bibliothekskataloge (OPACs) bilden hier keine Ausnahme. Daher sind Informationen häufig nur einzeln recherchierbar und kombinierte, aufeinander abhängige Recherchen müssen von Hand jeweils nacheinander angestoßen werden. Bei diesen Problemen soll das Semantic Web helfen.

Das Semantic Web ist die Vision einer Erweiterung des World Wide Webs, so dass die Daten nicht nur für Menschen leicht verständlich dargestellt werden, sondern auch von Maschinen verwertbar sind. Mit einer entsprechenden Ausgestaltung von Links zwischen einzelnen Webressourcen wäre das Web als riesige, globale Datenbank nutzbar. Darin könnten dann Softwareagenten für uns auch komplexe Fragestellungen und Planungen bearbeiten.

Das Semantic Web ist keine 'Spielerei' der Informatik, sondern es hat jetzt schon Anwendungen und ihm wird ein riesiges Potential vorausgesagt. Daher gibt es vielseitige Aktivitäten rund um das Semantic Web wie folgende Auflistung mit einem Schwerpunkt im bibliothekarischen Umfeld belegen soll:

- Das Bundesministerium für Wirtschaft und Technologie (BMWi) hat das Forschungsprogramm THESEUS<sup>2</sup> initiiert und fördert es in Millionenhöhe. Im Fokus des Programms stehen semantische Technologien.
- Das neue Europeana Datamodel (EDM)<sup>3</sup> ermöglicht eine Nutzung von Europeana als Linked Open Data.
- Der Wissenschaftsrat erwähnt in seinen aktuellen Empfehlungen zur Zukunft des bibliothekarischen Verbundsystems in Deutschland<sup>4</sup> auch das Semantic Web und Linked Open Data als gewandelte Nutzungsformen.
- Das GESIS (Leibniz-Institut für Sozialwissenschaften) hat ein Projekt zu Linked Open Data und hat insbesondere schon den Thesaurus Sozialwissenschaften im Semantic Web veröffentlicht.<sup>5</sup>
- Im Bachelor-Studiengang Bibliotheksmanagement an der FH Potsdam ist ein Bestandteil des Moduls Inhaltserschließung das Semantic Web und an der gleichen Hochschule gibt es im Master-Studiengang Informationswissenschaften ein eigenes Modul Semantische Technologien.<sup>6</sup>
- Die Konferenz Semantic Web in Bibliotheken (SWIB) findet dieses Jahr bereits zum dritten Mal statt.<sup>7</sup>
- Die erste DGI-Konferenz der deutschen Gesellschaft für Informationswissenschaft und Informationspraxis e.V. (DGI) fand 2010 unter dem Thema „Semantic Web und Linked Data - Elemente zukünftiger Informationsstrukturen“ statt.<sup>8</sup>
- Es gibt zahlreiche Publikationen und Vorträge zum Thema Semantic Web in der Bibliothekswelt. Ein Blick ins Literaturverzeichnis dieser Arbeit belegt dies ausführlicher.
- Zahlreiche Bibliotheken und Bibliotheksverbände haben begonnen Daten im Semantic Web zu veröffentlichen, unter anderem die schwedische<sup>9</sup>, deutsche<sup>10</sup>, französische<sup>11</sup> und ungarische<sup>12</sup> Nationalbibliothek, die British Library<sup>13</sup> und die Library of Congress<sup>14</sup>, das Hochschulbibliothekszentrum

des Landes Nordrhein-Westfalen (hbz)<sup>15</sup>, die Zentralbibliothek für Wirtschaftswissenschaften (ZBW)<sup>16</sup>, UB Mannheim<sup>17</sup> und UB Münster<sup>18</sup>.

Es gibt also vielfältige Tätigkeiten rund um das Semantic Web auch von Bibliotheken. In dieser Arbeit soll gezeigt werden, dass jede Bibliothek interessante Daten für das Semantic Web hat und umgekehrt von ihm profitieren kann.

Im Kapitel 2 wird das Semantic Web und allgemeine semantische Technologien kurz vorgestellt. Ein Schwerpunkt liegt dabei schon auf möglichen Anwendungsszenarien. Im Kapitel 3 werden Strategien für Bibliotheken mit dem Semantic Web skizziert ausgehend von verschiedenen Daten. Abschließend folgt ein kurzes Fazit und ein kleiner Ausblick im Kapitel 4.

## **2. Das Semantic Web**

### **2.1 Was ist das Semantic Web?**

Das Semantic Web (auch Semantisches Web) ist eine Erweiterung des World Wide Webs (WWW), in dem Informationen eine wohldefinierte Bedeutung besitzen, so dass diese auch maschinell verwertbar sind.<sup>19</sup> Das Ziel des Semantic Webs ist es, Informationen automatisch weiterzuverarbeiten und damit neuartige webbasierte Dienste anzubieten.<sup>20</sup>

Viele Teile davon sind noch in Entwicklung und daher ist das Semantic Web heutzutage noch eine Vision. Es gibt aber vielseitige Entwicklungen: auf der technischen Seite sind durch das Engagement des World Wide Web Consortiums (W3C)<sup>21</sup> Technologien wie etwa XML, RDF, RDFS, OWL, SPARQL standardisiert worden. Auf der Seite der Daten gibt es etliche Organisationen und Projekte, die strukturierte Metadaten im Sinne des Semantic Webs liefern. Daher kann man schon einige Anwendungen des Semantic Web sehen.

Die Daten in Semantic Web sollten als Linked Open Data oder Linked Data publiziert werden<sup>22</sup>, wobei Linked Data eine Menge von bewährten Vorgehensweisen für das Veröffentlichen und Verlinken von strukturierten Daten im Web ist<sup>23</sup>. Kernpunkt sind strukturierte Daten, die dann auch maschinell verarbeitet werden können und deren Verlinkung. Ebenfalls sollte man bemerken, dass es allgemein um Daten geht und nicht unbedingt nur um Dokumente. Diese bewährten

Vorgehensweisen wie etwa URI und RDF werden wir im Abschnitt 2.2 kurz streifen. Bei Linked Open Data kommt hauptsächlich die Veröffentlichung unter einer offenen Lizenz hinzu.<sup>24</sup>

Die Abgrenzung zwischen Semantic Web und Linked Open Data ist dabei fließend. Die beiden Begriffe werden teilweise als Synonyme gebraucht und an anderer Stelle wird aber darauf hingewiesen, dass Linked Open Data eher als Weg hin zum Semantic Web gesehen wird<sup>25</sup>.

Ein semantisches Netz ist der Oberbegriff von Semantic Web. Ein semantisches Netz kann auch bei abgeschlossenen Netzwerken wie etwa einem Firmennetz zum Einsatz kommen als System für das innerbetriebliche Wissensmanagement. Man denke zum Beispiel an ein elektronisches Ablagesystem. In Gegensatz zur gewohnten Ordnerstruktur, welche nur die Unterscheidung nach einem Kriterium pro Ebene zulässt, kann man in einem semantischen Netzwerk mehrere Kriterien gleichzeitig berücksichtigen.<sup>26</sup>

## 2.2 Technologien im Semantic Web

### *URI*

Ein wichtiger Unterschied des Semantic Web zum heute herkömmlichen Web besteht darin, dass es nicht mehr (nur) um Dokumente geht, die miteinander verlinkt sind, sondern vielmehr geht es um ein 'Web of Data'. Darum ist es auch notwendig, dass einzelnen Daten (Ressourcen) eindeutig identifiziert werden können. Wenn ein Restauranttester von Krebs spricht, dann meint er höchstwahrscheinlich das Krebstier und nicht die Krankheit oder das Sternzeichen. Um solche Verwechslungen zu verhindern, braucht man einen eindeutigen Identifikator und genau dafür ist die URI gedacht.

URI steht für uniform resource identifier und ist ein einheitlicher Identifikator für eine Ressource, welcher dem Standard RFC 3986<sup>27</sup> folgt. In der Praxis verwendet man entweder eine existierende URI oder kreiert eine neue URI als URL in der eigenen Domäne. Dies hat den Vorteil, dass man unter der URL gleich Informationen zu dem zu identifizierendem Objekt ablegen kann. Beispielsweise sollte man den Identifikator <http://dbpedia.org/resource/Crustacean> oder ähnliche

verwenden um Aussagen im Semantic Web über Krebstiere zu machen.<sup>28</sup> Solche Aussagen sind für Menschen etwas umständlich zu lesen, aber der große Vorteil ist, dass es für Maschinen klar verständlich ist und diese können uns dann wiederum eine besser lesbare Variante anbieten.

### *XML*

Die Markupsprache XML ist hervorragend geeignet für maschinenlesbare Dokumente und besitzt viele Werkzeuge für die Erstellung und Bearbeitung. Ein XML-Dokument besteht hauptsächlich aus Start- und Endtags sowie darin eingeschlossenem Inhalt, wie zum Beispiel einfacher Text. Die Tags müssen sich jeweils geordnet schließen<sup>29</sup>, so wird die Struktur eines XML-Dokuments baumförmig. Da die Verwendungen von Tag-Namen sehr flexibel ist, handelt es sich bei XML eigentlich um eine Metasprache<sup>30</sup>.

### *RDF, RDFS und OWL*

Das Resource Description Framework RDF ist eine formale Sprache für die Beschreibung strukturierter Informationen.<sup>31</sup> In dieser Allgemeinheit kann man auch sagen, dass man in RDF zu allgemeinen Ressourcen Metadaten darstellen kann. In RDF werden Tripel bestehend aus jeweils einem Subjekt, Prädikat und Objekt abgespeichert. Die Bestandteile solcher Tripel sollen jeweils entweder URIs oder Literale wie zum Beispiel Text sein. Wir können zum Beispiel die Aussage „Der Restauranttester isst Krebse“ in RDF modellieren, wobei das Subjekt dann „der Restauranttester“, das Prädikat „isst“ oder „essen“ und das Objekt „Krebse“ sind. Aber in RDF müssen Subjekt und Prädikat sowie gegebenenfalls das Objekt ein URI sein. Anstatt Krebse schreibt man also <http://dbpedia.org/resource/Crustacean> etc.

Eine Menge von solchen Tripeln kann man als Graph (im Sinne der Graphentheorie) darstellen: Subjekte und Objekte sind Knoten und Kanten werden über die Prädikate definiert. Damit bekommt man viel mehr Möglichkeiten als wenn man, wie etwa bei Taxonomien, nur hierarchische Strukturen modelliert.

Formal kann man Instanzen (z.B. einzelne Individuen wie der Restauranttester) und Klassen/Konzepte (wie etwa Krebstiere) unterscheiden. Menschen haben zu solchen Konzepten jeweils ein Hintergrundwissen. Ein Beispiel

für solches Hintergrundwissen wäre, dass Krebstiere eine Gruppe der Gliederfüßer sind. Ebenfalls kann man aus dem Prädikat „isst“ schließen, dass der Restauranttester ein Mensch ist<sup>32</sup> und Krebse Nahrungsmittel sind. Allgemein spricht man in der Informatik von 'Ontologien', wobei „eine Ontologie einen Wissensbereich mit Hilfe einer standardisierenden Terminologie sowie Beziehungen und ggf. Ableitungsregeln zwischen den dort definierten Begriffen beschreibt“<sup>33</sup>. Um solches Wissen maschinell auch zu erfassen und zu verarbeiten, werden Ontologiesprachen wie etwa RDFS und OWL benutzt.

Damit hat man ein mehrstufiges System: zu unterst hat man die Objekte und deren Beziehungen. Objekte aber auch Beziehungen sind Instanzen von allgemeineren Konzepten, die wiederum untereinander Beziehungen haben. Dies kann nach oben beliebig fortgesetzt werden. Der Vorteil liegt darin, dass man nicht ein Metadatenformat mit einem dicken Handbuch hat. Es ist vielmehr so, dass man die Daten strukturiert speichert und das Datenformat selbst wiederum als strukturierte Daten speichert.<sup>34</sup>

Abstrakte Konzepte, wie Krebstiere, werden maschinell verwertbar gemacht, indem man sie in einem Kontext stellt. Beispielsweise wäre eine Computeranwendung dann in der Lage, die Frage nach Gliederfüßer, die auch Nahrungsmittel sind, mit Krebsen zu beantworten.

Es gibt schon zahlreiche standardisierten Ontologien in RDFS/OWL für verschiedene Gebiete und diese sollte man nach Möglichkeit auch verwenden. Dabei kann man die einzelnen Vokabulare beliebig miteinander mischen.<sup>35</sup>

#### *Abfragesprache SPARQL*

SPARQL ist eine Abfragesprache für RDF-Daten, die stark an SQL angelehnt ist. Man kann mit einem select-Befehl alle Tripel abfragen, bei denen ein oder zwei Werte vorgegeben sind. Also beispielsweise kann man etwa abfragen „Was isst der Restauranttester?“. Dabei kann man die Bedingungen an die Variablen auch verschachteln und so zu viel komplexeren Abfragen kommen wie etwa „Welche Städtepartnerschaft zwischen deutschen und Schweizer Städten mit jeweils mehr als 30'000 Einwohnern gibt es?“ oder „Welche europäischen Länder haben mindestens drei offizielle Amtssprachen?“<sup>36</sup>

### 2.3 Allgemeine Anwendungsszenarien

Im Folgenden sollen allgemeine Anwendungsszenarien beschrieben werden und insbesondere der Frage nachgegangen werden, welche Vorteile das Semantic Web mit sich bringt. Allgemein hört man in der Semantic Web Community, dass eine 'Killerapplikation' noch nicht gefunden sei. Es gibt aber kleinere Anwendungen und auch eher visionäre Beispiele, die man sich auf Basis des Semantic Webs überlegen kann.

Allgemein gibt es Linked Data Browser<sup>37</sup> und Suchmaschinen<sup>38</sup>, die die Informationen zu einer URI oder allgemein zu einem Stichwort aus dem Semantic Web anzeigen und über die Links aus den RDF-Tripeln verknüpfen. Dies ist insbesondere wichtig, da nicht alle Tripel zu einer Ressource an genau einer Stelle sich befinden. Bestimmte Prädikate wie etwa `rdfs:seeAlso` oder `owl:sameAs` werden teilweise automatisch mit ausgewertet.

Die Zukunft von Suchmaschinen könnte dahin gehen, dass komplexe Suchanfragen in Form von Sätzen oder Fragen eingegeben werden können und vom System verstanden und entsprechend ausgewertet werden.<sup>39</sup>

#### *Anreicherungen der eigenen Information*

Wenn man die eigenen Daten als Linked Open Data veröffentlicht und auch auf andere Daten verlinkt, dann kann man diese zusätzlichen Informationen zur Anreicherung benutzen. Häufig werden die Adressdaten von Organisationen mit einer Karte von Google ergänzt. In Bibliothekskatalogen werden oft die Bücher mit dem entsprechenden Cover und Inhaltsverzeichnisse angereichert.

Diese Anwendungen existieren heute schon und sind nicht wirklich Neuerungen, die uns das Semantic Web bieten kann. Aber das Semantic Web kann eine Weiterentwicklung für die Verknüpfung von solchen Anreicherungen sein und somit als nächster Schritt in diese Richtung gesehen werden.

Im Katalog des österreichischen Bibliotheksverbundes wird bei einer Anzeige eines Titels jeweils ein kleiner Ausschnitt aus der DBpedia<sup>40</sup> zum entsprechenden Autor angezeigt.<sup>41</sup>

### *Neue Abfragemöglichkeiten*

Im Semantic Web werden Informationen so annotiert, dass diese auch maschinell verarbeitet werden können. Dadurch und durch die Verknüpfung von verschiedenen Datensets steigen die Abfragemöglichkeiten enorm an. Es wird dann auch möglich sehr komplexe und verschachtelte Abfragen zu beantworten. Beispielsweise kann man in der DBpedia folgende Anfrage stellen: erstelle eine Liste aller Fußballspieler, die als Torhüter für einen Klub mit einem Stadion mit mindestens 40'000 Sitzen gespielt haben und welche in einem Land mit mehr als 10 Millionen Einwohnern geboren wurden.<sup>42</sup> Für diese Abfrage müssen nicht einmal andere Datenquellen verwendet werden.

Der Leipziger Professorenkatalog (Catalogus Professorum Lipsiensis)<sup>43</sup> ist eine weitere Anwendung mit Technologien des Semantic Webs. Die Daten werden in einem normierten Vokabular eingetragen und sind maschinenlesbar. Eine Abfrage mit SPARQL ist für Universitätsangehörige möglich. Dadurch ist es beispielsweise möglich nach Professoren zu suchen, die in der Zeit 1933-1945 berufen wurden und Mitglied der NSDAP waren.<sup>44</sup>

### *Zusammenstellungen*

Man kann die Inhalte zweier oder mehrerer Webseiten auf einer neuen Webseite zusammenstellen oder allgemeiner miteinander verschmelzen (sogenannte 'Mashups'). Einfache Beispiele sind etwa persönliche Startseiten<sup>45</sup>, Zusammenstellungen von Nachrichten<sup>46</sup> aus verschiedensten Quellen oder Zusammenstellungen von Produkten<sup>47</sup> bei verschiedensten Anbietern. Eine wichtige Aufgabe bei den letzten zwei Beispielen ist die Zusammenführung von gleichen oder ähnlichen Nachrichten bzw. Produkten. Mit dem Semantic Web sind solche Zusammenführungen zuverlässig und sehr genau möglich.

Eine richtige Verschmelzung von Webinhalten kann teilweise durch öffentlich nutzbare freigegebene APIs ermöglicht werden, d.h. die Anbieter geben gewisse Funktionen frei zur Benutzung.<sup>48</sup> Offene APIs und offene Schnittstellen sind sehr schön, aber die größte Freiheit bekommt man durch Linked Open Data.

### *Automatische Verwertung von Informationen*

Wenn man Adressdaten oder Kontaktdaten auf seiner Webseite zusätzlich maschinenlesbar angibt, dann können diese Daten beim Besuchen der Webseite automatisch<sup>49</sup> ins elektronische Adressbuch des Benutzers übernommen werden. Eine Routenplanung vom momentanen Standort des Benutzers wäre auch denkbar. Ebenfalls können damit Veranstaltungen kodiert werden, sodass ein Eintrag im elektronischen Kalender übernommen wird.

Die zusätzlichen maschinenlesbaren Informationen werden in Mikroformaten als Teil der HTML- oder XHTML-Seite angegeben. Mikroformate werden auch als 'Lower Case Semantic Web' oder auch das 'Semantic Web von unten' angesehen.<sup>50</sup>

### *Softwareagenten für komplexe Aufgaben*

Die Vision von sehr mächtigen Softwareagenten für die Erledigung von komplexen Aufgaben im Semantic Web soll mit folgendem Beispiel beschrieben werden<sup>51</sup>: die Mutter von zwei Geschwistern muss nach einem Arztbesuch einen Physiotherapeuten zweimal wöchentlich aufsuchen, zu dem sie die Geschwister abwechslungsweise fahren werden. Der Softwareagent nimmt daraufhin die Daten des Arztes entgegen und verarbeitet diese mit Daten von Physiotherapeuten aus dem Semantic Web, insbesondere auch deren Sprechzeiten aber auch einzelne Bewertungen von anderen Menschen. Aus diesen Daten und den persönlichen Daten der Familie, dazu zählen Standortinformationen, Versicherungsdaten aber auch die Daten aus den Terminkalendern, kann solch ein zukünftiger Softwareagent eine komplexe Planung vorbereiten und Vorschläge für die Umsetzung geben. Andere Szenarien sind im Bereich der computergestützten Freizeitplanung<sup>52</sup> und im Bibliotheksumfeld<sup>53</sup> gegeben worden.

## **2.4 Wege ins Semantic Web**

Wie kommen die Daten ins Semantic Web? Es ist klar, dass das Umwandeln einiger Millionen Daten automatisch erfolgen soll. Dann kommt es darauf an, in welcher Form die Daten vorliegen.

Es gibt Softwareprogramme für die Umwandlung der Daten aus einer relationalen Datenbank in RDF-Tripel. Alternativ ist es möglich SPARQL-Anfragen über RDF-Tripel auf einer relationalen Datenbank zu simulieren. Zu Umwandlungen von anderen strukturierten Formaten (auch aus Excel) gibt es zahlreiche Programme und Programmbibliotheken.<sup>54</sup> Dabei muss man jeweils ein Daten-Mapping angeben, d.h. welche Felder und Beziehungen worauf abgebildet werden sollen. Falls solche Daten-Mappings nicht möglich sind, benötigt man semantische Technologien.<sup>55</sup>

Bei sogenannter semi-strukturierten Webseiten kann man Methoden der Informationsextraktion benutzen. Obwohl die Quellseiten in HTML oder XML sind und damit eigentlich eine klare Strukturierung vorherrscht, können die Strukturelemente auch nur zur Darstellung der Daten dienen (z.B. `<b></b>` für fett, `<i></i>` für kursiv). Falls die Seiten immer die gleiche Darstellung besitzen, kann man aus ihnen die Informationen mit einem sogenannten Wrapper automatisch extrahieren und in strukturierter Form abspeichern. Wrapper besitzen einfache Regeln, wie etwa 'der Titel ist immer fett' oder 'der Autor ist kursiv', aber sie können auch viel komplexere Regeln wie zum Beispiel 'der Titel ist in einer Aufzählung beim zweiten Punkt und fett geschrieben'. Diese Technik wird auch außerhalb des Semantic Webs angewendet, zum Beispiel bei der bibliographischen Datenbank DBLP<sup>56</sup> oder auch beim Literaturverwaltungsprogramm zotero<sup>57</sup>. Als Semantic Web Anwendung gibt es zum Beispiel SparqPlug<sup>58</sup>.

Unstrukturierten Daten, wie sie etwa in Foreneinträgen zu finden sind, sind am schwierigsten ins Semantic Web umzuwandeln. Da sie keinerlei Struktur besitzen, muss man mit anderen Methoden den Kontext herausfinden. Dazu kann man linguistische und statistische Methoden aber auch Techniken der Künstlichen Intelligenz hinzunehmen.<sup>59</sup>

### **3. Semantic Web und Bibliotheken**

Es sollen verschiedene Daten, die in Bibliotheken anfallen, analysiert werden auf ihre Verwendbarkeit im Semantic Web und auf den Nutzen der Bibliothek bei der Veröffentlichung dieser Daten im Semantic Web. Daraus soll dann auch die mögliche Rolle des Semantic Webs in Bibliotheken hervorgehen und auch umgekehrt die Rolle der Bibliotheken im Semantic Web.

Bei Bibliotheksdaten denkt man normalerweise an Bibliothekskataloge, in denen die vorhandenen Bücher, Zeitschriften, elektronischen Medien recherchierbar sind. Ein Katalogeintrag ist dann in verschiedene Felder unterteilt, z.B. Titel, Autor, Schlagwörter, Formatangaben. Die Katalogdaten sind dabei untereinander verlinkt zum Beispiel bei mehrbändigen Werken werden die einzelnen Bände verlinkt oder die Verlinkung zwischen der gedruckten und elektronischen Ausgabe einer Zeitschrift. Ebenfalls werden die Autoren und Schlagwörter durch Links zu den Normdaten angegeben. In den Normdaten sind insbesondere auch Synonyme gespeichert und werden bei einer Recherche automatisch mitgesucht. Ebenfalls unterscheidet man Homonyme, d.h. gleichlautende Wörter mit unterschiedlichen Bedeutungen (z.B. Krebs).

Wieso sind dann Bibliotheksdaten nicht schon Linked Open Data in der jetzigen Form? Ein großes Problem ist, dass man von außen nur schwer wenn überhaupt auf die Bibliotheksdaten verlinken kann. Das ist der Grund warum z.B. in Google die Katalogdaten nicht indiziert werden. Häufig wird dafür der Begriff 'Datensilo' verwendet. Es gibt Links innerhalb des Bibliothekskataloges und Verknüpfungen zu den Normdaten. Aber keine Links außerhalb des Bibliotheksumfeldes. Ein zweites Problem sind die Bibliotheksstandards wie etwa die Datenformate MAB oder MARC und die Regelwerke wie etwa RAK-WB (über 600 Seiten) oder AACR. Bevor eine Softwareentwicklung basierend auf Bibliotheksdaten gemacht werden kann, muss man sich mit den verwendeten Datenformaten und Regelwerken auseinandersetzen. Dagegen sind Softwareentwicklungen im Semantic Web dann automatisch mit Linked Open Data aus Bibliotheken kompatibel.

### **3.1 Bibliographische Daten**

Es gibt eine ganze Menge von bibliographischen Daten in Bibliotheken und diese kann man in verschiedene Entitäten aufteilen (z.B. Personen, Themen, Werke, Ereignisse, Geographika) und dazwischen gibt es Verknüpfungen. Bibliographische Daten aus Bibliotheken sind somit interessant für das Semantic Web. Umgekehrt ist es aus verschiedenen Gründen für Bibliotheken auch interessant ihre bibliographischen Daten als Linked Open Data zu veröffentlichen.

Durch eine Veröffentlichung der bibliographischen Daten können diese auch von außen angereichert werden und zum Beispiel im OPAC angezeigt werden, vergleiche auch Abschnitt 2.3.1. Diese Anreicherungen funktionieren aber auch in die andere Richtung. Nach einer Veröffentlichung der bibliographischen Daten als Linked Open Data kann man diese auch in anderen Webseiten zur Anreicherung verwenden. Zum Beispiel gibt es in Wikipedia bei einzelnen Artikeln unter weiterführende Literatur jeweils Links zu Bibliothekskatalogen. Natürlich sind solche Links zur Anreicherung auch von anderen Seiten her möglich. Es ist zum Beispiel vorstellbar, dass ein Pizza-Bäcker auf seiner Homepage sein verwendetes Kochbuch verlinkt oder Rezensionen zu einem Buch werden gleich mit dem Buch selbst verknüpft. Für einen Softwareagenten wird es ein Leichtes sein, dieses Buch in der nächsten Bibliothek für einen Anwender zu lokalisieren.

Heutige Bibliothekskataloge sind besonders darauf ausgelegt ein bestimmtes Medium zu finden, und weniger gut darin, den Benutzern Unterstützung bei der Recherche zu einem ihnen wenig vertrauten Thema zu geben. Eine Möglichkeit ist es, im Bibliothekskatalog beispielsweise auch Autorensseiten oder Themenseiten zu erstellen. Die Deutsche Nationalbibliothek (DNB) hat bei ihrem Katalog auch die Normdaten hineingenommen. Noch etwas weiter geht das Projekt openlibrary, bei dem auf den Themenseiten die Literatur zum Thema aufgelistet und analysiert wird, beispielsweise werden die Autoren mit den meisten Publikationen bei diesem Thema angezeigt.<sup>60</sup> Zusammenfassend kann man nach Coyle<sup>61</sup> sagen, dass es weniger um einen Katalog mit einem Fokus auf den Dokumenten („document-centric catalog“) geht als um Katalog, in dessen Zentrum einzelne Informationen sind („information-centric catalogue“).

### **3.2 Thesauri, Klassifikationen**

Häufig werden die Medien bei der Sacherschließung in Bibliotheken auch in eine Klassifikation eingeordnet. Typischerweise ist dies die DDC oder RVK, aber es gibt auch eigene Systematiken. In Klassifikationen gibt es Beziehungen zueinander wie etwa 'Unterklasse von', 'Oberklasse von' oder 'verwandte Klasse'. Des Weiteren gibt es bei zwei verschiedenen Klassifikationen sicherlich teilweise Überschneidungen. Somit sind Thesauri und Klassifikationen ideale Daten für das Semantic Web.

Es gibt gute Gründe für eine Veröffentlichung von Klassifikationsschemas als Linked Open Data. Man kann seine Klassifikationsgruppen mit anderen Konzepten im Semantic Web verknüpfen, beispielsweise mit der DBpedia. Dies erleichtert sicherlich auch die Anwendung bzw. das Erlernen des Klassifikationsschemas. Über die Verlinkung ist es auch möglich die sachliche Erschließung zu automatisieren. Denkbar ist, dass man automatisch die Erschließung von anderen Internetbenutzern oder auch nur ausgewählten vertrauenswürdigen Quellen mit dem gleichen Klassifikationsschema übernimmt im Sinne einer kollaborativen Erschließung. Ebenfalls möglich ist es, die Daten aus anderen Klassifikationen zu nutzen um die eigene Klassifikation automatisiert zu erstellen.<sup>62</sup>

Eine möglichst vollständige Klassifizierung ist auch in Hinblick auf Suchen mit Facettierungs- und Filtermöglichkeiten äußerst relevant.<sup>63</sup>

### **3.3 Ausleihdaten**

Bei einer Bestellung eines Mediums im OPAC werden die Ausleihbedingungen mit den persönlichen Daten und Rechten des Benutzers verglichen. Als Ergebnis wird das Buch bestellt oder die Bestellung abgewiesen. Dabei sind eine Menge an Informationen im Spiel. Welche dieser Informationen könnten im Semantic Web von Bedeutung sein? Zunächst ist es wegen dem Datenschutz klar, dass keinerlei personenbezogene Daten nach außen weitergegeben werden können. Nützlich wäre nach Bestellung die automatische Übernahme des Ausleihendes in den elektronischen Kalender der Bibliothekskunden.

Die Ausleihdaten (oder dann auch Daten aus der Recherche) könnten als statistische Daten gesammelt werden und dann zum Beispiel als eine Sortierreihenfolge (Ranking) wiederverwendet werden. Durch Veröffentlichung als Linked Open Data könnten auch andere Bibliotheken (mit ähnlicher fachlicher Ausrichtung) davon profitieren analog zu Empfehlungsdiensten wie etwa BibTip oder bX.<sup>64</sup>

### **3.4 Organisatorische Daten**

Bibliotheken möchten sicherlich auch die Besitznachweise für einzelne Medien speichern. Um dies als Linked Open Data zu machen, muss zuerst jede Bibliothek

eine URI bekommen<sup>65</sup>, die dann durch organisatorische Daten ergänzt werden kann. Ebenfalls kann es von Vorteil sein, wenn einzelne Informationen bei dem Webauftritt der Bibliothek auch maschinenlesbar sind. Somit wäre es möglich die Adressdaten ins elektronische Adressbuch zu speichern oder Termine von Veranstaltungen zu übernehmen. Aber auch komplexere Informationen wie die Öffnungszeiten sind im Semantic Web darstellbar.<sup>66</sup>

Idealerweise würden die Daten bei jeder Bibliothek selbst eingegeben und dann automatisch in ein elektronisches Verzeichnis kopiert. Solch eine dezentrale Organisation hat den Charme, dass die Daten immer aktuell gehalten werden<sup>67</sup>. Interessant wäre auch aus diesen Daten automatisch Statistiken zu ermitteln.

Eine erste Implementation hat das hbz gemacht, wobei die Daten dort (noch) zentral erstellt wurden auf Grundlage des Online-ISIL-Verzeichnisses und der MARC Organization Codes Database.<sup>68</sup> Die Daten können dann auch visuell auf einer Karte dargestellt werden.<sup>69</sup>

### **3.5 Daten aus Digitalisierungsprojekten**

Bei Digitalisierungsprojekten fallen neben den Daten der Digitalisate auch Metadaten an. Dies wird zum einen eine Art Katalogisat sein, welches das Objekt und auch die Technik der Digitalisierung beschreibt. Zusätzlich findet man häufig strukturelle Daten wie etwa die Seitenzählung und eine Art Gliederung. Gerade bei Digitalisierungsprojekten wird häufig neben den reinen technischen Digitalisaten versucht einen Mehrwert zu schaffen. Beispielsweise durch bessere Recherchefunktionen oder Verknüpfungen auf andere Quellen. Genau bei diesen Aufgaben sind Linked Open Data ideal. Solche Projekte können auch gut geeignet sein, um erste Erfahrungen mit dem Semantic Web zu sammeln.

Zum Beispiel möchte man in einer digitalisierten Portrait-Sammlung<sup>70</sup> nach Persönlichkeiten von bestimmten Städten suchen können eventuell mit einer Auswahl auf einer Karte. Eine Möglichkeit wäre eigene Listen mit diesen Daten zu führen und eine entsprechende Suche bereitzustellen. Eine andere Möglichkeit wäre die Daten aus DBpedia über Verlinkung zu beziehen. Falls die Daten in DBpedia bzw. Wikipedia nicht vollständig wären, könnte man die gegebenenfalls dort zentral ergänzen. Damit wäre auch eine optimale Nachhaltigkeit der Arbeit gewährleistet.

Eine relativ einfache Möglichkeit für die Verknüpfung von den Normdaten-Nummern zu passenden Webseiten (z.B. Wikipedia, ADB, NDB) bietet das BEACON-Format<sup>71</sup>.

#### **4. Fazit und Ausblick**

Das Semantic Web kann als technische Neuerung bzw. Vision angesehen werden oder aber auch als eine konsequente Fortführung der bisherigen Strategien von Bibliotheken. Die Ideen der kollaborativen Katalogisierung und Erschließung sowie die Verwendung von Normdaten werden auch im Semantic Web unterstützt und weitergeführt. Die Wissensdatenbanken der Bibliotheken werden damit mit der gesamten Wissensbasis des Semantic Webs verknüpft. Daraus ergeben sich vielfältige Anwendungsmöglichkeiten. Allen voran sehe ich die Erweiterung der Anreicherung von Informationen, die Möglichkeit auch in anderen Systemen nach Literatur zu recherchieren sowie die fast grenzenlose Fülle von neuen Abfragemöglichkeiten.

Für mich ist klar, dass Bibliotheken ihre Daten ins Semantic Web geben sollten. Die Frage stellt sich mehr nach dem 'wie' und 'wer genau'. Vielfach wird bei einer Veröffentlichungs-Strategie zuerst einmal die prinzipielle Durchführbarkeit durch die Entwicklung eines prototypischen Linked Data Services belegt (sogenannter 'Proof of Concept'). Lizenzfragen, die genauen Namensräume der URIs und/oder die beste Vokabular-Modellierung werden dabei teilweise noch außen vor gelassen und erst in einem zweiten Schritt weiterverfolgt. Ebenfalls ist noch kein Bibliographie-Vokabular eindeutig erkennbar, welches alle bibliographischen Informationen aus Bibliothekskatalogen vollständig abbildet.<sup>72</sup>

Die dezentrale Datenpflege ist eine gute Möglichkeit um die Aktualität von Daten zu gewährleisten, aber hat den Nachteil, dass eine mögliche Abfrage länger dauert.<sup>73</sup>

Ich habe versucht in dieser Arbeit aufzuzeigen, dass jede Bibliothek interessante Daten für das Semantic Web hat. Man kann zum Beispiel beginnen Daten auf der Bibliothekswebseite zusätzlich in maschinenlesbarer Form zu speichern oder in Projekten die Techniken des Semantic Webs zu nutzen.

Eine mögliche Dienstleistung von Bibliotheken in Zukunft wäre das Speichern von Inhalten im Semantic Web, z.B. für Forscher eine Wissensbasis

publizieren.<sup>74</sup> Ebenfalls rücken virtuelle Forschungsumgebungen sowie das Speichern von Forschungsdaten immer mehr in den Fokus von Bibliotheken. Dabei wird sicherlich auch das Semantic Web eine wichtige Rolle spielen. Interessant ist hierbei auch die Arbeit der UB Mannheim beim Speichern von Projektdaten mit Hilfe von semantischen Technologien.<sup>75</sup>

Die ZDB<sup>76</sup> und EZB<sup>77</sup> verzeichnen fortlaufende Sammelwerke wie etwa Zeitschriften oder Zeitungen und wären auch Kandidaten für Linked Open Data mit Links zu den klassifikatorischen DDC-Daten. Die Abfragemöglichkeiten würden damit steigen, z.B. suche Zeitschriften über Krebstiere oder welche Zeitungen gab es in Frankfurt während des Zweiten Weltkrieges. Ebenfalls wäre die Webis-Datenbank<sup>78</sup> der Sammelschwerpunkte an deutschen Bibliotheken ein möglicher Kandidat für Linked Open Data. Interessant erscheint mir hier vor allem auch das Register mit den vielen Verweisen zu den Sondersammelgebieten. Denkbar wäre dann, dass die Benutzer passend zu ihrer Recherche im OPAC Hinweise auf Daten in der ZDB, EZB oder Webis bekommen würden. Somit könnte man das Problem der vielen verschiedenen Sucheinstiege für Benutzer etwas entschärfen.

Eines der Probleme, dem sich das Semantic Web noch stellen muss, ist Vertrauen. Wenn man Daten aus dem Semantic Web weiterverwenden möchte, dann muss man wissen, ob die Daten vertrauenswürdig sind. In diesem Bereich müssen sich noch Techniken für die Praxis herausstellen. Man kann dies aber auch als Chance für Bibliotheken sehen: Bibliotheken und ihre Daten sind vertrauenswürdig und könnten mit 'Linked Open Bibliotheksdaten' zum Rückgrat des Semantic Webs werden.<sup>79</sup>

## Endnoten

<sup>1</sup> Eine tagesaktuelle Schätzung bekommt man unter <http://www.worldwidewebsize.com/>, wobei zu beachten ist, dass 'billion' im Englischen einer Milliarde im Deutschen entspricht. Dazu kommt noch dass bei Weitem nicht alle Webseiten indexiert werden, vgl. Sack (2009, S. 14): „Diese Zahl betrifft dabei nur einen Teil der insgesamt über das WWW erreichbaren Dokumente und lässt zugriffsbeschränkte Informationen, die meisten dynamisch generierten Inhalte und die Informationsbestände der zahlreichen Intranets unberücksichtigt.“

<sup>2</sup> <http://www.theseus-programm.de/>, [http://de.wikipedia.org/wiki/Theseus\\_\(Forschungsprogramm\)](http://de.wikipedia.org/wiki/Theseus_(Forschungsprogramm))

<sup>3</sup> <http://www.europeana.eu/schemas/edm/>.

<sup>4</sup> <http://www.wissenschaftsrat.de/download/archiv/10463-11.pdf>, S. 10-11.

<sup>5</sup> <http://www.gesis.org/unser-angebot/recherchieren/thesauri-und-klassifikationen/thesaurus-sozialwissenschaften/>.

<sup>6</sup> <http://informationswissenschaften.fh-potsdam.de/iw-moduluebersicht.html> sowie Neher (2010).

<sup>7</sup> <http://www.swib09.de>, <http://swib.org/swib10> und <http://swib.org/swib11/>.

- <sup>8</sup> <http://www.dgi-konferenz.de/index.html> und Tagungsband Ockenfeld (2010).
- <sup>9</sup> <http://libris.kb.se/>.
- <sup>10</sup> [http://www.dnb.de/DE/Service/DigitaleDienste/LinkedData/linkeddata\\_node.html](http://www.dnb.de/DE/Service/DigitaleDienste/LinkedData/linkeddata_node.html).
- <sup>11</sup> <http://data.bnf.fr/>.
- <sup>12</sup> [http://nektar.oszk.hu/wiki/Semantic\\_web](http://nektar.oszk.hu/wiki/Semantic_web).
- <sup>13</sup> <http://www.bl.uk/bibliographic/datafree.html>.
- <sup>14</sup> <http://id.loc.gov/>.
- <sup>15</sup> [http://www.hbz-nrw.de/projekte/linked\\_open\\_data/](http://www.hbz-nrw.de/projekte/linked_open_data/).
- <sup>16</sup> <http://zbw.eu/stw/versions/changes/about>.
- <sup>17</sup> <http://data.bib.uni-mannheim.de/>.
- <sup>18</sup> <http://data.uni-muenster.de/>.
- <sup>19</sup> Übersetzung aus Berners-Lee, Hendler & Lassila (2001): „The Semantic Web is not a separate Web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation.“
- <sup>20</sup> <http://www.w3.org/2001/sw/EO/points>, [http://de.wikipedia.org/wiki/Semantic\\_Web](http://de.wikipedia.org/wiki/Semantic_Web).
- <sup>21</sup> Die Webseite des W3C lautet <http://www.w3.org/> und für die Aktivitäten über Semantic Web siehe <http://www.w3.org/2001/sw/>.
- <sup>22</sup> Tom Heath, ein Koautor von Tim Berners-Lee, beschreibt auf der Webseite <http://linkeddata.org/faq>, dass Tim Berners-Lee 'linked data' häufig als „the Semantic Web done right“ beschreibt und verweist dazu als Beispiel auf die Folien eines Vortrages von Tim Berners-Lee: [http://www.w3.org/2008/Talks/0617-lod-tbl/#\(3\)](http://www.w3.org/2008/Talks/0617-lod-tbl/#(3)).
- <sup>23</sup> Übersetzung aus Bizer, Heath & Berners-Lee (2009, S. 1): „The term Linked Data refers to a set of best practices for publishing and connecting structured data on the Web.“
- <sup>24</sup> Siehe <http://opendefinition.org/> für eine ausführliche Darstellung wie offenes Wissen zu verstehen ist. Zu den rechtlichen Aspekten gibt es einen aktuellen Leitfaden vom Juristen Till Kreuzer im Auftrag des hzb, siehe Kreuzer (2011).
- <sup>25</sup> Siehe Auer (2010, Folie 3): „Linked Data as a deployment path for the Semantic Web“.
- <sup>26</sup> Siehe Reichenberger (2010).
- <sup>27</sup> <http://tools.ietf.org/html/rfc3986>.
- <sup>28</sup> Berners-Lee (2006, Folie 4): „Don't say 'colour' say <<http://example.com/2002/std6#col>>.“
- <sup>29</sup> Dies ist vergleichbar mit wohlgeformten Klammerausdrücken, bei denen sich die öffnenden und schließenden Klammern auch jeweils geordnet verhalten müssen, z.B. ist `()` ein wohlgeformter Klammerausdruck aber `()(` (natürlich nicht).
- <sup>30</sup> Geisler (2009, S. 17).
- <sup>31</sup> Eine detaillierte Ausführung über RDF mit auch vielen technischen Aspekten findet man in Hitzler et al. (2008, Kapitel 3).
- <sup>32</sup> Für ein Tier hätte man das Prädikat „frisst“ verwendet.
- <sup>33</sup> <http://www.gi.de/service/informatiklexikon/detailansicht/article/ontologien.html>.
- <sup>34</sup> Bizer, Heath & Berners-Lee (2009, S.4): „Data is self-describing. If an application consuming Linked Data encounters data described with an unfamiliar vocabulary, the application can dereference the URIs that identify vocabulary terms in order to find their definition.“
- <sup>35</sup> Siehe Bizer, Heath & Berners-Lee (2009, S.7) bzw. Bizer, Cyganiak & Heath (2007): „Only if these vocabularies do not provide the required terms should data publishers define new, data source-specific terminology“.
- <sup>36</sup> Folgende Abfrage auf <http://dbpedia.org/sparql>
- ```
SELECT ?country ?lang1 ?lang2 ?lang3 WHERE {
  ?country <http://purl.org/dc/terms/subject>
  <http://dbpedia.org/resource/Category:European_countries> .
  ?country <http://dbpedia.org/property/officialLanguages> ?lang1 .
  ?country <http://dbpedia.org/property/officialLanguages> ?lang2
  FILTER(?lang1!=?lang2)
  ?country <http://dbpedia.org/property/officialLanguages> ?lang3
  FILTER(?lang1!=?lang3 && ?lang2!=?lang3)
}
```
- liefert das Ergebnis Belgien, Bosnien und Herzegowina, Polen und Schweiz (ohne Gewähr auf Richtigkeit).

- <sup>37</sup> Z.B. Disco hyperdata browser <http://www4.wiwiss.fu-berlin.de/bizer/ng4j/disco/> oder Marbles <http://www5.wiwiss.fu-berlin.de/marbles>.
- <sup>38</sup> z.B. SWSE <http://swse.deri.org/>, VisiNav <http://visinav.deri.org/>, Falcons <http://ws.nju.edu.cn/falcons/objectsearch/index.jsp> oder Sindice <http://sindice.com/>.
- <sup>39</sup> Ein Beispiel das mittels linguistischer Verfahren und vielen verknüpften Daten in diese Richtung geht ist <http://www.wolframalpha.com/>. Beispiele für Suchanfragen sind „Who is the president of Germany?“, „100 Swiss francs in euros“, „ $x^2+(y-\sqrt{abs(x)})^2=3$ “, „where is google?“, „U235“, „compare stegosaurus with tyranosaurus rex“.
- <sup>40</sup> DBpedia speichert die strukturierten Informationen aus Wikipedia als Linked Open Data und verlinkt diese mit anderen Datenquellen, siehe <http://dbpedia.org/>.
- <sup>41</sup> Siehe Krabo & Knittel (2011, S. 26) sowie die Suchmaschine des Österreichischen Bibliotheksverbundes: <http://search.obvsg.at/OBV>
- <sup>42</sup> <http://wiki.dbpedia.org/OnlineAccess#h28-5>.
- <sup>43</sup> <http://www.uni-leipzig.de/unigeschichte/professorenkatalog/>.
- <sup>44</sup> Siehe Auer (2010, ab 38min 25s im Videomitschnitt).
- <sup>45</sup> Z.B. iGoogle <http://www.google.de/ig>.
- <sup>46</sup> Z.B. Google News <http://news.google.de/>.
- <sup>47</sup> Z.B. Google Shopping (Produktsuche) [www.google.de/prdhp](http://www.google.de/prdhp).
- <sup>48</sup> Nowak, Heller & Korzen (2010).
- <sup>49</sup> Z.B. <https://addons.mozilla.org/de/firefox/addon/operator/>.
- <sup>50</sup> Schulze (2010).
- <sup>51</sup> Das Beispiel stammt aus dem Artikel von Berners-Lee, Hendler & Lassila (2001).
- <sup>52</sup> Computergestützte Freizeitplanung ist die Vision einer automatischen Tagesplanung basierend auf geographischen Daten sowie Empfehlungen anderer Personen und einem Thema, siehe Bücken & Macit (2010).
- <sup>53</sup> In Gillitzer (2011) wird ein Szenario eines Softwareagenten beschrieben, welcher bei einer Suche für eine seltene Handschrift, die Öffnungszeiten und Ausleihbedingungen der besitzenden Bibliothek recherchiert und auf dieser Basis eine Reiseverbindung vorschlägt.
- <sup>54</sup> <http://www.w3.org/2001/sw/wiki/SemanticWebTools>.
- <sup>55</sup> Blumauer & Pellegrini (2009, S. 20): „Semantische Technologien, wie z. B. automatische Texterschließung, der Einsatz von semantischen Netzen, Ontologien oder Informationsextraktion mit Hilfe von Wrapper-Technologien, kommen in vielfältiger Weise bereits heute, vor allem im Kontext wissensintensiver, interdisziplinärer und kommunikationsintensiver Arbeitsabläufe als Mittel im 'Kampf gegen die Informationsflut' zum Einsatz.“
- <sup>56</sup> Die DBLP Informatik Bibliographie findet man unter <http://dblp.uni-trier.de/db/index.html>. Die eingesetzten Techniken wurden mir von beteiligten Personen des Projektes persönlich erklärt und sind in der Diplomarbeit von Oliver Hoffmann nachzulesen, siehe Hoffmann (2009).
- <sup>57</sup> Siehe <http://www.zotero.org/> und spezieller <http://www.zotero.org/support/dev/technologies>.
- <sup>58</sup> Coetzee, Heath & Motta (2008).
- <sup>59</sup> Z.B. OpenCalais <http://viewer.opencalais.com/>.
- <sup>60</sup> <http://openlibrary.org/> und eine Themenseite z.B. <http://openlibrary.org/subjects/crustacea> und eine Personenseite z.B. [http://openlibrary.org/authors/OL81959A/Frederick\\_II\\_King\\_of\\_Prussia](http://openlibrary.org/authors/OL81959A/Frederick_II_King_of_Prussia).
- <sup>61</sup> Coyle (2010b, 22min 17s im Videomitschnitt).
- <sup>62</sup> Die DDC als Linked Data findet man unter <http://dewey.info/>. Die UB Mannheim hat die RVK in Linked Data überführt, siehe auch [http://wiki.bib.uni-mannheim.de/linked-data/doku.php?id=dokumentation\\_de](http://wiki.bib.uni-mannheim.de/linked-data/doku.php?id=dokumentation_de). Wahrscheinlich werden sich aber sowohl die URI wie auch die genauen Implementierungsdetails noch ändern. Der Standardthesaurus Wirtschaft wurde von der ZBW erstellt und als Linked Data veröffentlicht, siehe <http://zbw.eu/stw/versions/latest/about.de.html>.
- <sup>63</sup> Krabo & Knittel (2011, S. 25f).
- <sup>64</sup> <http://swib.org/swib11/abstracts.html#pfeffer>.
- <sup>65</sup> Pohl (2011b, S. 1).
- <sup>66</sup> Pohl (2011b, S. 67f und S. 79ff).
- <sup>67</sup> Pohl (2011a, S. 21) sowie Knittel & Krabo (2011, S. 27).
- <sup>68</sup> <http://lobid.org/organisation>.
- <sup>69</sup> z.B. <http://lobid.org/organisation/DE-605/about>.

<sup>70</sup> Z.B. <http://www.tripota.uni-trier.de/>.

<sup>71</sup> <http://de.wikipedia.org/wiki/Wikipedia:PND/BEACON> und Voß, Schindler & Thiele (2011).

<sup>72</sup> Schomburg (2010, ab 24min 19s im Videomitschnitt).

<sup>73</sup> Hartmann & Haffner (2010, 44min 09s): „Wenn wir tatsächlich in diese RDA-Welt eintauchen, bedeutet das, wir müssen bereits katalogisierte Daten weltweit nachnutzen. Das sind dezentral organisierte Datenpools. Wir müssen dann – keine Ahnung – auf 100 verschiedenen Bibliothekssystemen gleichzeitig kucken, gibt es das Werk dort schon, gibt es die Person dort schon. Und das wird uns höchstwahrscheinlich auch an Performanzgrenzen bringen. Das kann man erwarten. Und daher auch hier die Diskussion, wo macht es Sinn zentrale Ankerpunkte einzufügen. Aus meiner Sicht ist für Deutschland die GND definitiv ein sehr gelungener zentraler Ankerpunkt, viaf international das Pendant und auch der culuturegraph könnte eine ähnlich Funktion einnehmen.“

<sup>74</sup> Auer (2010, Folie 26).

<sup>75</sup> Eckert & Pfeffer (2010).

<sup>76</sup> Zeitschriftendatenbank, siehe <http://www.zeitschriftendatenbank.de/>.

<sup>77</sup> Elektronische Zeitschriftenbibliothek, siehe <http://rzblx1.uni-regensburg.de/ezeit/>.

<sup>78</sup> <http://webis.sub.uni-hamburg.de/webis/index.php/Hauptseite>.

<sup>79</sup> Altenhöner (2010, Folie 13).

*Anmerkung:* Alle angegebenen Internetquellen wurden zuletzt am 20.06.2012 geprüft.

## Literatur

- Altenhöner, R. (2010). Kollaboration durch das Semantic Web: Strategie und Aktivitäten der Deutschen Nationalbibliothek. Vortrag gehalten auf der SWIB10, Köln, 29.11.2010. doi:[10.4016/27094.01](https://doi.org/10.4016/27094.01)
- Auer, S. (2010). Von offenen Daten zu einem Ökosystem vernetzten Wissens. Vortrag gehalten auf der SWIB10, Köln, 29.11.2010. doi:[10.4016/27091.01](https://doi.org/10.4016/27091.01)
- 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. *Scientific American*, 284(5), S. 34-43. <http://www.scientificamerican.com/article.cfm?id=the-semantic-web>
- Berners-Lee, T. (2006). Semantic Web, Building on Existing Systems. Vortrag gehalten an der MIT ILP Konferenz, Cambridge, 04.04.2006. <http://www.w3.org/2006/Talks/0404-mit-tbl/>
- Bizer, C., Cyganiak, R. & Heath, T. (2008). How to publish Linked Data on the Web. <http://www4.wiwiss.fu-berlin.de/bizer/pub/LinkedDataTutorial/> (Version vom 17.07.2008)
- Bizer, C., Heath & T., Berners-Lee, T. (2009). Linked Data - The Story So Far. *International Journal on Semantic Web and Information Systems*, 5, S. 1-22. doi:[10.4018/jswis.2009081901](https://doi.org/10.4018/jswis.2009081901)
- Blumauer, A. & Pellegrini, T. (Hrsg.). (2009). *Social semantic web : web 2.0, was nun?* Berlin: Springer.
- Bücker, P. & Macit, U. (2010). Computergestützte Freizeitplanung basierend auf Points of Interest. In M. Ockenfeld (Hrsg.), *Semantic Web & Linked Data: Elemente zukünftiger Informationsstrukturen* (S. 103-116). Wiesbaden: Dinges & Frick. (1. DGI-Konferenz ; 62. Jahrestagung der DGI).
- Coetzee, P., Heath, T. & Motta, E. (2008). SparqlPlug: Generating Linked Data from Legacy HTML, SPARQL and the DOM. In C. Bizer et al. (Hrsg.), *Proceedings of the 1st Workshop on Linked Data on the Web* (Paper Nr. 6). <http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-369/paper05.pdf>
- Coyle, K. (2010a). Understanding the Semantic Web: Bibliographic Data and Metadata. *Library Technology Reports*, 46 (1). <http://alatechsource.metapress.com/content/g212v1783607/>
- Coyle, K. (2010b). Yes We Can! Libraries and the Semantic Web. Vortrag gehalten auf der SWIB10, Köln, 30.11.2010. doi:[10.4016/27061.01](https://doi.org/10.4016/27061.01)
- Eckert, K. & Pfeffer, M. (2010). Linked Open Projects - Nachnutzung von Projektergebnissen als Linked Data. Vortrag gehalten auf der SWIB10, Köln, 30.11.2010. doi:[10.4016/27051.01](https://doi.org/10.4016/27051.01)
- Fürste, F. M. (2011). *Linked Open Library Data*. Wiesbaden: Dinges & Frick. (BIT online: Innovativ, 33)
- Geisler, M. (2009). *Semantic Web: schnell + kompakt*. Unterhaching: entwickler.press.
- Gillitzer, B. (2011). Von alten Zöpfen in neuen Schläuchen oder Warum Aristoteles schon alles wusste, aber erst Google damit Geld verdient. Vortrag gehalten auf dem 100. Bibliothekartag, Berlin, 2011.
- Hartmann, S. & Haffner, A. (2010). Linked-RDA-Data in der Praxis. Vortrag gehalten auf der SWIB10, Köln, 30.11.2010. doi:[10.4016/27060.01](https://doi.org/10.4016/27060.01)
- Hengartner, U. (2010). *Web 3.0 & Semantic Web*. Heidelberg: dpunkt-Verl. (HMD: Praxis der Wirtschaftsinformatik, 271)

- Hitzler, P., Krötzsch, M., Rudolph, S. & Sure, Y. (2008). *Semantic Web*. Berlin: Springer, 2008. doi:[10.1007/978-3-540-33994-6](https://doi.org/10.1007/978-3-540-33994-6)
- Hoffmann, O. (2009). Regelbasierte Extraktion und asymmetrische Fusion bibliographischer Informationen. Diplomarbeit, Universität Trier, 2009. <http://dblp.uni-trier.de/papers/DiplomarbeitOliverHoffmann.pdf>
- Krabo, U. (2011). Consuming Library Linked Data: ein Prototyp einer Linked Data Webapplikation. Masterarbeit, FH Technikum Wien, 2011. <http://hdl.handle.net/10760/15991>
- Krabo, U. & Knitel, M. (2011). Library Linked Data - Technologien, Projekte, Potentiale. *Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen und Bibliothekare*, 64 (1), S. 11-31. <http://eprints.rclis.org/handle/10760/16102#.TqrnZXL0kUE>
- Kreutzer, T. (2011). Open Data - Freigabe von Daten aus Bibliothekskatalogen. <http://www.hbz-nrw.de/dokumentencenter/veroeffentlichungen/open-data-leitfaden.pdf>
- Neher, G. (2010). Semantic Web in der bibliothekarischen Ausbildung : Welche Kompetenzen müssen vermittelt werden? Vortrag gehalten auf der SWIB10, Köln, 30.11.2010. doi:[10.4016/27048.01](https://doi.org/10.4016/27048.01)
- Nowak, M., Heller, L. & Korzen, S. (2010). Mashups und Bibliotheken. In J. Bergmann & P. Danowski (Hrsg.), *Handbuch Bibliothek 2.0* (S. 143-160). Berlin: de Gruyter Saur. doi:[10.1515/9783110232103.143](https://doi.org/10.1515/9783110232103.143)
- Ockenfeld, M. (Hrsg.). (2010). *Semantic Web & Linked Data : Elemente zukünftiger Informationsstrukturen*. Wiesbaden: Dinges & Frick. (1. DGI-Konferenz ; 62. Jahrestagung der DGI)
- Pellegrini, T. & Blumauer, A. (Hrsg.). (2006). *Semantic Web: Wege zur vernetzten Wissensgesellschaft*. Berlin : Springer. doi:[10.1007/3-540-29325-6](https://doi.org/10.1007/3-540-29325-6)
- Pohl, A. (2010). Open Data im hbz-Verbund. *ProLibris*, 3, S. 109-113. <http://hdl.handle.net/10760/14838>
- Pohl, A. (2011a). Linked Data und die Bibliothekswelt. [Erscheint im Konferenzband der ODOK 2010] <http://hdl.handle.net/10760/15324>
- Pohl, A. (2011b). lobid vocab: Zur Konzeption und Entwicklung eines RDF-Vokabulars für die Beschreibung bibliothekarischer Organisationen, ihrer Sammlungen und Dienstleistungen. Masterarbeit, FH Köln, 2011. <http://hdl.handle.net/10760/16175>
- Pohl, A. & Ostrowski, F. (2010). Zur Konzeption und Implementierung einer Infrastruktur für freie bibliographische Daten . In M. Ockenfeld (Hrsg.), *Semantic Web & Linked Data: Elemente zukünftiger Informationsstrukturen* (S. 205-216). Wiesbaden: Dinges & Frick (1. DGI-Konferenz, 62. Jahrestagung der DGI). <http://hdl.handle.net/10760/14839>
- Reichenberger, K. (2010). *Kompodium semantische Netze Konzepte, Technologie, Modellierung*. Heidelberg: Springer. doi:[10.1007/978-3-642-04315-4](https://doi.org/10.1007/978-3-642-04315-4)
- Sack, H. (2010). Semantische Suche: Theorie und Praxis am Beispiel der Videosuchmaschine yovisto.com. In U. Hengartner & A. Meier (Hrsg.), *Web 3.0 & Semantic Web* (S. 13-25). Heidelberg: dpunkt Verlag. (HMD - Praxis der Wirtschaftsinformatik, 271). <http://hmd.dpunkt.de/271/02.php>
- Schomburg, S. (2010). Semantic Web - ein Thema für Bibliotheksverbände. Vortrag gehalten auf der SWIB10, Köln, 29.11.2010. doi:[10.4016/27098.01](https://doi.org/10.4016/27098.01)

- Schulze, C. M. (2010). Mikroformate. In J. Bergmann & P. Danowski (Hrsg.), *Handbuch Bibliothek 2.0* (S. 129-142). Berlin: de Gruyter Saur. doi:[10.1515/9783110232103.129](https://doi.org/10.1515/9783110232103.129)
- Voß, J., Schindler, M. & Thiele, C. (2011). Link server aggregation with BEACON. In J. Griesbaum, T. Mandl & C. Womser-Hacker (Hrsg.), *Information und Wissen: global, sozial und frei?* (S. 519-522). Boizenburg: vwh. (Proceedings des 12. Internationalen Symposiums für Informationswissenschaft, ISI 2011) <http://hdl.handle.net/10760/15407>

*Anmerkung:* Alle angegebenen Internetquellen wurden zuletzt am 20.06.2012 geprüft.