SWIB14: Bibliotheken bauen das vertrauenswürdige semantische Web

Bericht über die Fachkonferenz „Semantic Web in Libraries“, Bonn, 1. bis 3. Dezember 2014

Vera Münch

Nach der großen Datenexplosion hat die Informatikforschung um die Jahrtausendwende unter dem Stichwort „Semantic Web“ mit dem Sortieren der Informationsinhalte im World Wide Web begonnen. Dabei hat sie das bibliothekarische Fachwissen mit seinen traditionellen Praktiken strukturierter Katalogisierung, Referenzierung und Informationsverknüpfung entdeckt. Das schlägt sich seither in ungezählten Linked Data und Linked Open Data (LOD) Pilotprojekten nieder. Nun soll die geballte Arbeitskraft der Bibliotheken dieser Welt dazu genutzt werden, das Web aufzuräumen. Die ersten Arbeiten sind im Gange.

„Bibliotheken können als eine Art Glaubwürdigkeitsmaschinen hochwertige Informationen auswählen, sammeln und jederzeit und überall für alle bereitstellen, die irgendwo auf irgendeinen Bildschirm gucken“, so Eric Miller, einer der beiden Hauptredner der SWIB14. Gleichzeitig könnten sie sich und ihr „weltweit glaubwürdigstes Informationsangebot“ im Web sichtbar machen. Doch nicht nur in der semantischen Erschließung der Metadaten läge die große Chance: „Es sind die Links, die ein Informationsangebot im Web sichtbar machen. Die Links, die auf die Seiten hinweisen“, betonte Miller immer und immer wieder. Die von Dritten gesetzten Verknüpfungen zur Ursprungsseite der Information brächten Webseiten in den Rang- und Trefferlisten der Suchmaschinen nach oben, so seine Erklärung. Bibliotheken hätten enorme Möglichkeiten, sich durch Interlinking im Web sichtbar zu machen.

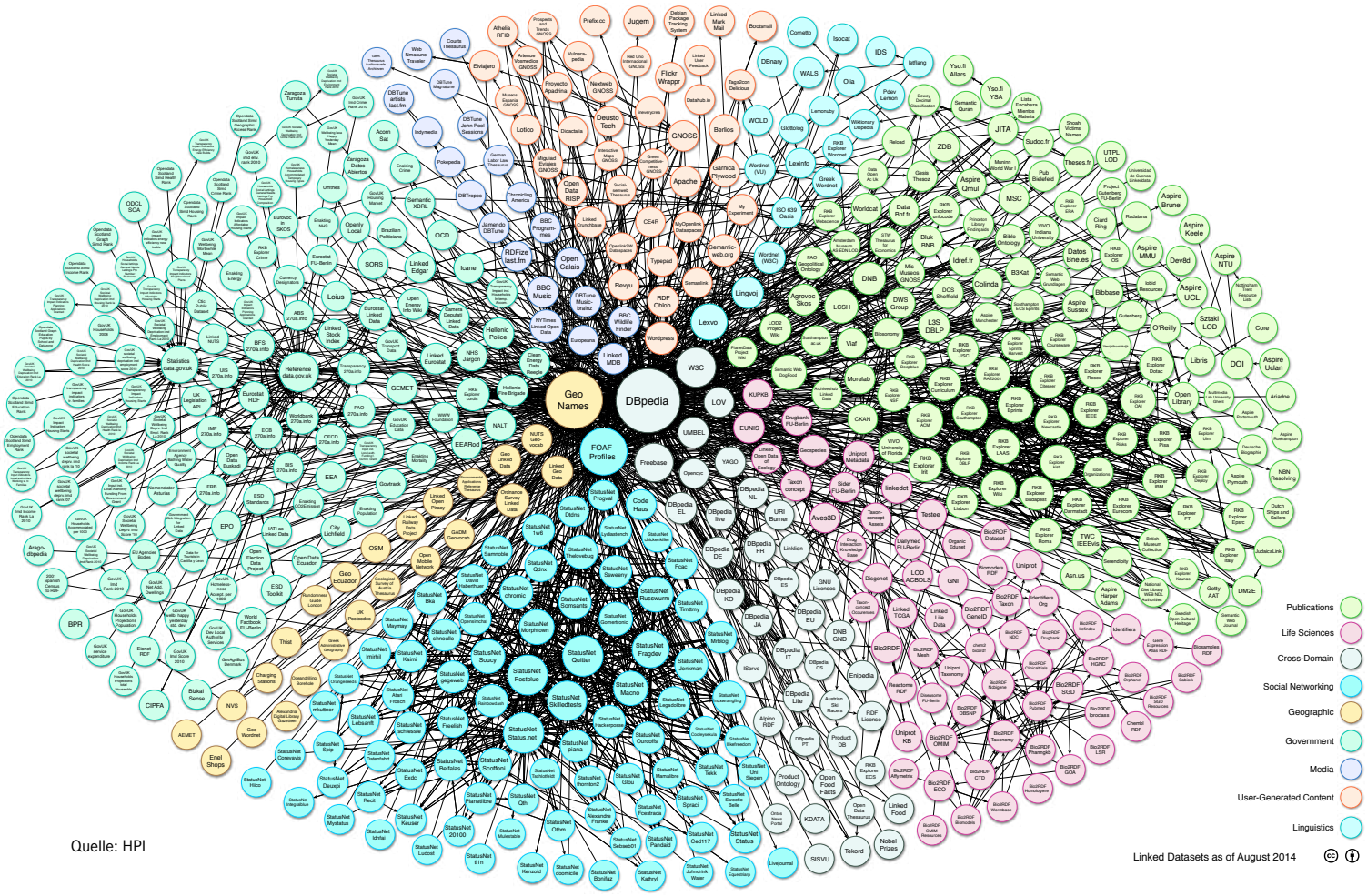
OCLC-Evangelist schlägt entitätenbasierte Softwarearchitektur vor

Richard Wallis, Technologie-Evangelist der Bibliothekskooperative OCLC, tat es Eric Miller gleich: „Links, Links, Links“ predigte er in Bonn. Mit einer Suche nach der Autobiographie von Mahatma Gandhi „The Story of My Experiments With Truth“ führte er vor Augen, wie es mit der Sichtbarkeit von Bibliotheken im Web derzeit aussieht. Der erste auf einen Bibliotheksbestand hinweisende Treffer stand auf der 60. – in Worten: sechzigsten – Seite der Google-Trefferliste. Wallis bezeichnete Linked Data (die Methode, um Daten und Informationen im Web in Sinnzusammenhängen bereitzustellen) „als starke und wertvolle Technologie, aber eben nur eine Technologie“. Um Bibliotheksbestände aus den Tiefen des Webs strukturiert auf die Bildschirme der Informationssuchenden zu heben, müsste den von ihnen bereitgestellten Linked Data bzw. Linked Open Data (LOD) eine „Entity based Data

Die SWIB auf einen Blick

Seit 2009 veranstaltet die ZBW – Deutsche Zentralbibliothek für Wirtschaftswissenschaften zusammen mit dem Hochschulbibliothekszentrum des Landes Nordrhein-Westfalen (hbz) die internationale Konferenz „Semantic Web in Bibliotheken“, kurz SWIB. Die SWIB (http://swib.org/) ist eine Konferenz, die sich in Vorträgen, Diskussionen und Workshops sowohl wissenschaftlich als auch alltagspraktisch mit den übergeordneten Themen Semantic Web, Linked Open Data und webbasierter Wissenschaftskommunikation beschäftigt. hbz und der ZBW organisieren die Konferenz im jährlichen Wechsel. Das Programm der SWIB14 wurde durch ein international besetztes Programmkomitee unter der Leitung von Joachim Neubert (ZBW) und Adrian Pohl (hbz) zusammengestellt.
Zur SWIB14 kamen165 Teilnehmerinnen und Teilnehmer aus 24 Ländern, darunter USA, Kanada, Kenia und Burundi. Die Konferenz wurde per Livestream übertragen, die Vorträge als Videomittschnitt aufgezeichnet. Links zu Folien und Konferenzvideos:
http://swib.org/swib14/programme.php
Zu den Referent/inn/en: *http://swib.org/swib14/speakers.php*
Zu den Tweets: *https://twitter.com/search?f=realtime&q=%23swib14&src=typd*

1 Tim Berners-Lee, http://en.wikipedia.org/wiki/Giant_Global_Graph



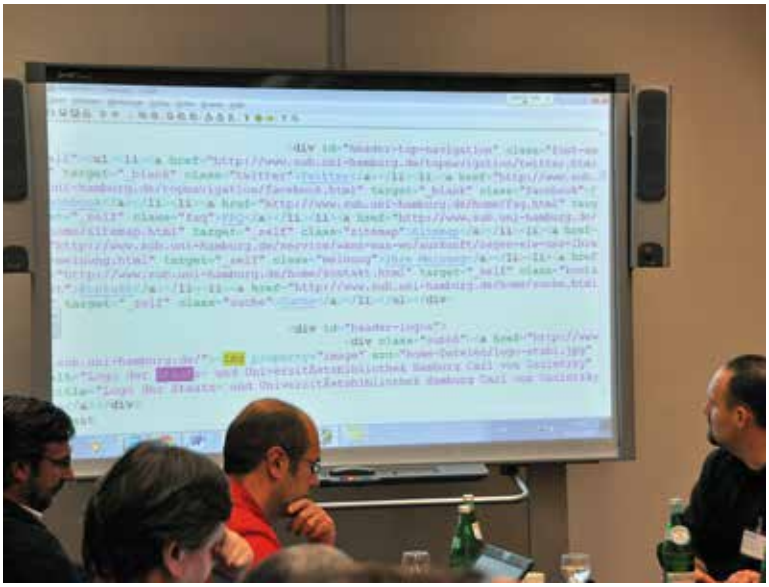
Quelle: HPI

Das „Linked Open Data Cloud Diagram" versucht einen Überblick zu geben, welche Verzeichnisse bereits als LOD-Zugriffspunkte freigeschaltet sind. Bereitgestellt vom Hasso Plattner Institut, Universität Mannheim im Rahmen des EU-Projektes LATC (Linked Open Data Around-The-Clock).

Architecture" zugrunde gelegt werden; Softwaresysteme, die mit Blick auf die Publikation von Daten im Web konzipiert sind. Das bedeute, bei ihrem Design von der Verarbeitung von Informationsobjekten – Entitäten – auszugehen und diese durch Uniform Ressource Identifiers (URI) zu identifizieren und zu verknüpfen. So können dann beispielsweise zu einem per Websuche gefundenen Buch die geografisch am nächsten liegenden Bibliotheken im Suchergebnis angezeigt werden. Wallis nennt diese Erschließung der Informationsinhalte von Datensätzen „Entification". Er sieht sie als „den Weg zu nützlichen Bibliotheksdaten".

Video-Mitschnitte, Abstracts und Folien gibt es open access

Der aufgrund dynamischer Folienpräsentation auf 204 Einzelseiten-



Bibliotheksarbeit heute: Programmieren, Daten sortieren und Information klassifizieren, um Bibliotheken im Web sichtbar zu machen. Dan Scott von der kanadischen Laurentian University (rechts) zeigte im Workshop vor der SWIB14, wie man Schema.org in den HTML-Code von Webseiten einbaut. Da sieht der Bildschirm dann so aus.

darstellungen angewachsene Foliensatz zum Vortrag von Wallis ist wie alle anderen Folien auf der Programmseite der SWIB14 zum Nachlesen bereitgestellt. Dort finden sich auch Kurzbiografien der Referentinnen und Referenten sowie Videoaufzeichnungen der meisten Vorträge. *(Adresse im Kasten „Die SWIB auf einen Blick".)* 165 Fachleute aus 24 Ländern tauschten sich auf der Konferenz darüber aus, welche Erkenntnisse und Entwicklungen es zur semantischen Informationsverknüpfung im Web heute gibt und was Forschung und Anwendungspraktiker als nächste Schritte planen. Wer sich vor dem Ansehen der Folien bzw. Vorträge mit Linked-Data-Technologie vertrauter machen möchte, dem sei das gut verständliche Video[2] der Europeana empfohlen.

Eric Miller glaubt fest daran, dass die Sichtbarkeit der Bibliotheken und ihrer Informationsangebote im Web allen Menschen nützt. Der Daten- und Semantic-Web-Spezialist war mit seinem Unternehmen Zepheira im Jahr 2011 Partner der amerikanischen Library of Congress bei der Entwicklung von BIBFRAME, einem Modell und Referenzrahmen, um bibliothekarische Daten im Beschreibungsformat MARC in das semantische Web zu übertragen.

Informationsbeschreibung auf kleinster Ebene

Überall auf der Welt laufen Entwicklungsprojekte, die versuchen, den Inhalt von Datensätzen, Datenbanken und Repositorien so zu beschreiben, dass Suchmaschinen ihn interpretieren können. Entitäten werden durch die bereits erwähnten „Uniform Ressource Identifiers", kurz URI, eindeutig identifiziert. Die Anwendungsentwickler erstellen dafür beispielsweise zur Entität „Automobil" als kleinster Informationseinheit Bezugsdreiecke aus Subjekt, Prädikat und Objekt, aus denen maschinell abgeleitet werden kann, dass es sich bei einem Automobil um eine Maschine handelt die fährt und Menschen und Güter transportieren kann. Diese Beschreibung wird in einem Datenformat des Resource Description Framework (RDF) erstellt und heißt RDF-Tripel[3]. Tripel sind die Bausteine der semantischen Erschließung des Webs. Aus vielen Tripeln entsteht ein riesiges Bezugsnetz: der Giant Global Graph, in dem Marken, Modelle, Baujahre, Ingenieure etc. miteinander verknüpft sind. Aus dieser Wissensbasis sollen Suchmaschinen, so jedenfalls wünscht es sich die Semantic-Web-Forschung, irgendwann in Zukunft Sinnzusammenhänge erkennen können. Dann sind die Suchalgorithmen in der Lage, zu einer Suchanfrage tatsächlich die inhaltlich passende Antwort zu liefern. Mit einem Klick auf den in der Trefferanzeige angegebenen URI der ursprünglichen Informationsquelle wird der Nutzer auf die vollständige Information dort weitergeleitet. Je mehr Webseiten nun auf eine URI verlinken, desto höher wird sie in der Trefferliste angezeigt, erklärten Miller und Wallis.

Gibt es die richtige Antwort auf eine Suche?

Weil es aber im Datenozean des World Wide Web auf eine Anfrage immer mehr als eine Antwort gibt, bewerten die Forscher nun das Verlinken auf vertrauenswürdige Quellen als so immens wichtig. Bibliotheken haben hier mehr als ein Pfund in die Waagschale zu werfen, um der immer stärker von Werbeinteressen geprägten Suchmaschinenoptimierung Paroli zu bieten. Alexander Vogt, Mitglied des Landtags von Nordrhein-Westfalen und medienpolitischer Sprecher der SPD, betonte bei der Eröffnung der SWIB14 die große Bedeutung eines frei zugänglichen Web: „Die freie, demokratische Verfügbarkeit von Information war immer ein Kampf um demokratische Gesellschaften." Sammeln, strukturieren und publizieren von Wissen und Information trage zur Demokratisierung der Gesellschaften bei, was bedeute: „Sie tragen hier bei, die Demokratie zu formen", so der Politiker. Vogt befürwortet Open Access: „Daten, die mit öffentlicher Förderung entstanden sind, gehören der Öffentlichkeit." Er kündigte an, dass Daten aus der Administration des Landes Nordrhein-Westfalen über das Portal „Open.NRW"[4] zum Download bereitgestellt werden. Die Entwicklung läuft.

2 http://vimeo.com/36752317.

3 http://en.wikipedia.org/wiki/Resource_Description_Framework, http://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/

4 http://www.nrw.de/opennrw/opennrw-1/

Linked-Open-Data-Cloud-Diagramm zeigt LOD-Quellen

Die Datenaufbereiter für das semantische Web beschreiben tatsächlich jede Entität, jedes bedeutungstragende Datum, jeden Informationswert. Das geschieht allerdings nicht immer von Grund auf neu. Man nutzt zunehmend vorhandene Klassifikationen, Taxonomien, Thesauri und Ontologien; also strukturierte Verzeichnisse, wie sie das Bibliotheks- und Dokumentationswesen seit Urzeiten aufbaut und besitzt. Immer mehr Einrichtungen stellen ihre Verzeichnisse als LOD-Referenzquellen zur Nachnutzung und Verlinkung zur Verfügung. Einen Überblick, welche Verzeichnisse bereits als LOD-Zugriffspunkte freigeschaltet sind, versucht das „Linked Open Data Cloud Diagram" zu geben.[5] Die dynamische Grafik (Abb. S. 56) ähnelt einem aus unzähligen großen und kleinen, kreuz und quer vernetzten Kreisen zusammengesetzten Feuerwerk. Im Zentrum stehen die Datenbanken GeoNames[6] und die DBpedia[7], strukturierte Informationen aus Wikipedia, und die FOAF[8]-Personenprofile. Im Orbit finden sich bereits viele bibliothekarische Verzeichnisse, Normdateien und Register, die als LOD-Satelliten den Cyberspace mit zuverlässigen Referenzdaten bedienen.

hbz-Leiterin sieht große Chancen für Bibliotheken

Silke Schomburg, die als Leiterin des hbz mit ihrem Team um Adrian Pohl für die lokale Organisation der SWIB14 in Bonn zuständig war, freut sich über die Entwicklungen: „Es tun sich so viele Chancen auf, die Sichtbarkeit von Bibliotheken im Web zu erhöhen. Was die Referentinnen und Referenten auf der Konferenz präsentieren, ist spannend. Ich bin sehr neugierig, was die internationalen Praktiker und Forscher aus den vielfältigen Möglichkeiten machen."

„Die freie, demokratische Verfügbarkeit von Information war immer ein Kampf um demokratische Gesellschaften", so Alexander Vogt, Mitglied des Landtags von Nordrhein-Westfalen und medienpolitischer Sprecher der SPD, bei der Eröffnung der SWIB14.

BBC öffnet alle Ontologien

Tom Grahame, der mit seiner Keynote den Vortragsteil der Konferenz eröffnet hatte, erzählte, wie die BBC aus ihrer statisch präsentierten Sportberichterstattung im Web ein dynamisches Portal gebaut hat. „Der Natur des Sportes entsprechend", stellt es Informationen schnell und unter Nutzung aller Medienarten bereit und bezieht soziale Kommunikationsfunktionen ein. Graham sagte, die BBC hätte die Art, wie sie Inhalte publiziere, dabei völlig verändert und erklärte: „Solange Sie Ihr Informationsangebot nicht semantisch aufbereiten, sind Ihre Inhalte in die Webseite eingebacken." Nach dem Sport hat die BBC im nächsten Schritt auch ihre weitere Berichterstattung im Web dynamisiert. Dauerhaft kann sie, so Graham, aber nicht alle dafür notwendigen Ontologien selbst aufbauen und pflegen. Hier hofft sie auf Unterstützung durch Bibliotheken. Im Gegenzug gibt sie alle bei der bisherigen Entwicklung ihrer Informationsangebote aufgebauten Ontologien zur allgemeinen Nutzung frei[9], darunter neben Sport Business, Wildlife, Food, Politics u.a.

BIBFRAME bringt MARC-Katalogdaten ins Web

Mit jedem Vortrag ging es tiefer in die Informatik und Informationswissenschaft. In den technisch ausgerichteten Vorträgen tauchten immer mehr Abkürzungen auf, mit denen die Fachleute Standards, Vokabulare und Datenprojekte beschreiben. *(Siehe dazu auch Kastentext „SWIB Fachchinesisch" auf S. 59.)* Der eingangs bereits zitierte zweite Keynotspeaker der SWIB 14, Eric Miller, arbeitet „seit einem Vierteljahrhundert an Metadatenprojekten, die uns da hin gebracht haben, wo wir heute sind", so die Moderatorin des Konferenzblocks, Karen Coyle[10], die als Bibliothekarin 30 Jahre Erfahrung mit digitaler Bibliothekstechnologie und selbst zahlreiche interessante Publikationen[11] dazu geschrieben hat. Eric Miller war mit seinem Unternehmen

5 http://lod-cloud.net/
6 http://www.geonames.org/
7 http://de.dbpedia.org/
8 http://www.foaf-project.org/
9 http://www.bbc.co.uk/ontologies
10 http://www.kcoyle.net/infopeople/index.html
11 http://orcid.org/0000-0002-1085-7089

Zepheira[12] im Jahr 2011 Partner der amerikanischen Library of Congress bei der Entwicklung von BIBFRAME, einem vielbeachteten Modell und Referenzrahmen, um bibliothekarische Daten im Beschreibungsformat MARC in das semantische Web zu übertragen. In seinem Vortrag erklärte er, wie man mit Hilfe von BIBFRAME in MARC erfasste Katalogdaten so aufbereiten kann, dass aus Bibliotheksdaten Linked Data werden, die dann in der Lage sind verschiedenste Anwendungen und Aktivitäten im Web im großen Stil zu bedienen. BIBFRAME entspricht laut Miller sowohl den funktionalen Erfordernissen von Katalogdaten, als auch denen der semantischen Beschreibung von Entitäten auf Basis des RDA-Regelwerks.

Schema.org in die eigene Webseite einbinden

Im weiteren Konferenzverlauf präsentierten Referentinnen und Referenten Werkzeuge, Verfahren und Plattformen zur semantischen Erschließung und Anreicherung von Inhalten sowie Referenzdatenbanken und Normdateien (Authority Files) für den automatischen Abgleich und die automatische Ergänzung von Metadaten und sie erklärten, wie man diese in HTML-Webseiten ergänzt. Normdaten werden zunehmend von Einrichtungen wie der Library of Congress oder der Deutschen Nationalbibliothek (DNB) bereitgestellt, so dass langsam ein immer größeres Netz an vertrauenswürdigen Referenz- und Normdatenquellen entsteht. Schema.org, das vom W3C mit Google, Bing und Yahoo! entwickelte Rahmenwerk, dem sich auch Yandex angeschlossen hat, spielte sowohl in einem der bereits Wochen vor der Tagung ausgebuchten Workshop der SWIB14 eine große Rolle als auch in der Konferenz selbst. Dan Scott von der kanadischen Laurentian University erläuterte „maschinenlesbares Katalogisieren für das offene Web“ mit Hilfe von Schema.org. Im fünfstündigen Workshop „RDFa with Schema.org“ brachte er den Teilnehmerinnen und Teilnehmern bei, wie man Schema.org in den HTML-Code von Webseiten einbaut.

Entity Facts – DNB bietet Faktenblätter zu Entitäten der GND

Als jüngste Entwicklung der DNB stellten Michael Büchner und Christoph Böhme „Entity Facts“ vor, einen Datendienst, der maschinenlesbare „Faktenblätter“ zu Entitäten der Gemeinsamen Normdatei (GND) bereitstellt. Sie können in externe Anwendungen eingebunden werden. Die DNB bezeichnet den Dienst, der Entitäten und Quellenhinweise bietet, als „leichtgewichtigen Authority Data Service“. Das Ziel sei ein Service, der es ermöglicht, aus der GND maschinell Informationen zu einer Entität aggregieren zu können, indem sowohl ausgewählte Informationen zur Entität als auch Links zu anderen Datenquellen bereitgestellt werden. Ein Beispiel, wie ein solches Referenzdatenblatt aussieht, kann man sich auf der Culturgraph-Webseite bei den Projekten[13] ansehen. Die Deutsche Digitale Bibliothek (DDB) arbeitet schon mit dem DNB Dienst. Sie bietet Personenseiten auf Basis von Entity Facts an.

Wikidata, eine strukturierte Sammlung weltweiten Wissens

„Wikidata – eine freie, kollaborative Wissensbasis“, über die Markus Krötzsch in Bonn berichtete, ist das Produkt internationaler Zusammenarbeit ungezählter tausender Beteiligter, die in bis zu 400 Sprachen Informationen ebendort bereitstellen. Große Teile der eingetragenen Daten und Fakten sind mit weltweit verfügbaren Normdateien, Katalogen und Referenzdatenbanken verlinkt. Durch die gemeinsame technische Basis ist, so Krötzsch, „eine der größten von Menschen hergestellten Sammlungen strukturierter Information“ entstanden. Die als vertrauenswürdig eingestuften Referenzdaten sind unter CC-0 Lizenz publiziert. Sie können in allen Wikimedia-Projekten eingebaut und darüber hinaus von jedermann verwendet werden, der sie in sein Angebot einbauen möchte.

Globaler Thesaurus für die Landwirtschaft und SPARQL-Skosmos

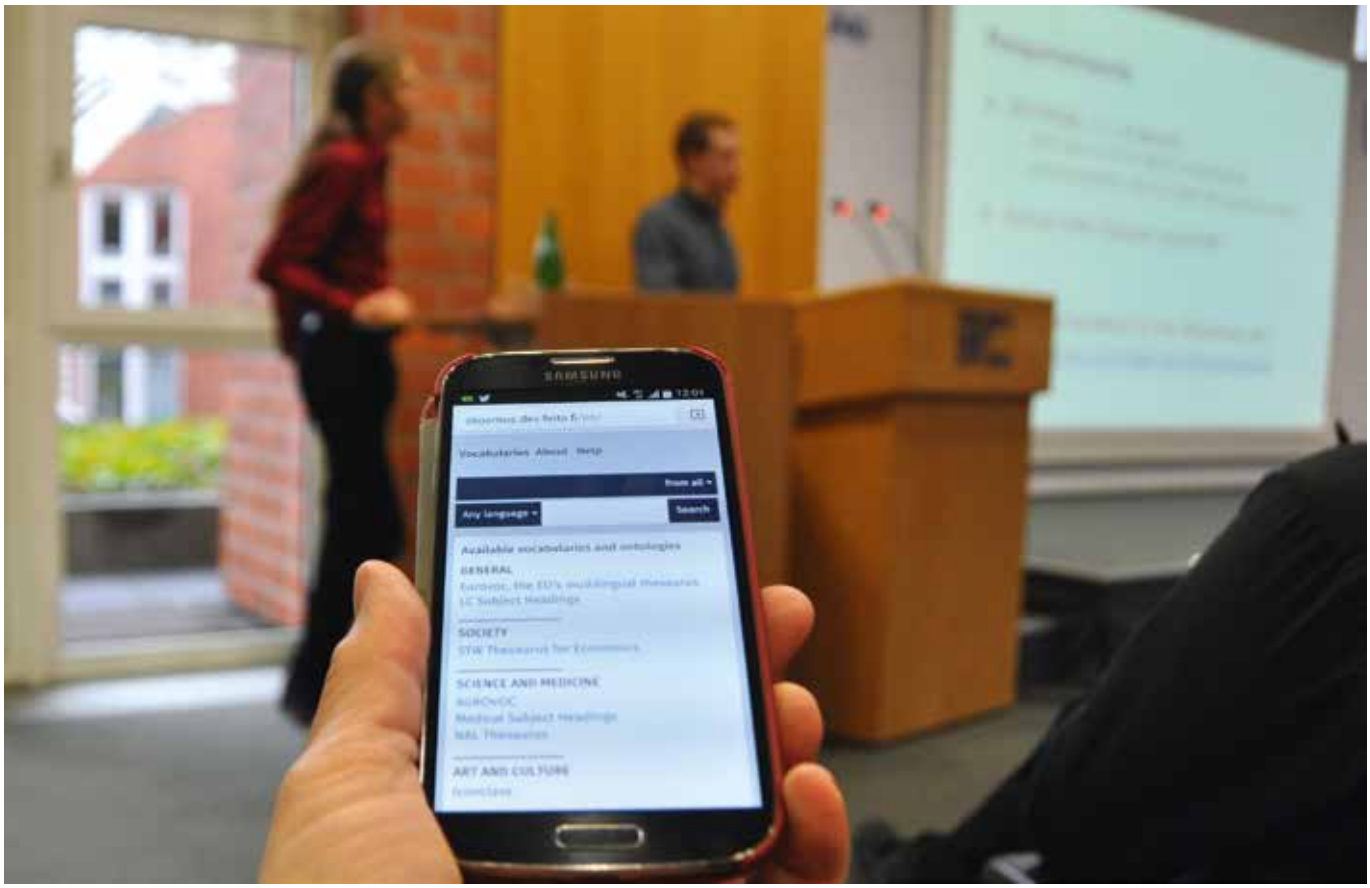
Eine weit fortgeschrittene Nutzung von Thesauri, Taxonomien und Ontologien zum Aufbau fachlicher Wissensgraphen und globaler fachbezogener Konzeptschemata präsentierte Osma Suominen, Forscher an der Nationalbibliothek von Finnland. Mit Kollegen in Finnland und dem Forscher Thomas Baker von der Koreanischen Sungkyunkwan University hat er drei überlappende Thesauri aus der Agrarindustrie, der Land-, Forst- und Fischereiwirtschaft zu

SWIB-Fachchinesisch

Fachleuten sagen die im Nachfolgenden ohne Wertung, Erklärung und Anspruch auf Vollständigkeit aufgeführten Abkürzungen etwas. Anderen Interessierten bringen Webrecherchen Aufschluss. Auf der SWIB14 ging es um Linked Data, LOD, RDF, RDFa, RDF/XML, HTML, JSON-LD, Turtle, Triples, N-Triples, Schema.org, VIAF, FOAF, DBpedia, GND, Entity Facts, Wikimedia, Dewey.info, FAST, BIBFRAME, GitHub, LibHub, lobid, LibreCat, Catmandu, MongoDB, Elasticsearch, d:swarm, SVOT, FRBR, Open Metadata Registry, SKOS, SKOSMOS, EDM, URIs, Aliada.eu, Bartoc, VIVO, Graphity, Metafacture, regal ... und mehr.

12 http://zepheira.com/

13 http://hub.culturegraph.org/entity-facts/118540238



Ein Blick in die Zukunft und doch schon Gegenwart: Skosmos auf dem Handy ausprobieren, während der Vortrag von Thomas Baker, Sungkyunkwan University, Korea, und Osma Suominen, Forscher an der Nationalbibliothek von Finnland, noch läuft.

einem „Global Agriculture“ Konzeptschema zusammengeführt. Die Finnen entwickeln auch interessante Anwendungen für den W3C-Standard „Simple Knowledge Organisation System“ (SKOS[14]). So haben sie eine Software zur Veröffentlichung und zum Browsing von SKOS-Vokabularien entwickelt. Die von Suominen vorgestellte Software mit dem schönen Namen Skosmos begeisterte das SWIB-Auditorium.

Daten aufräumen mit d:swarm, SVOT, VIVO, Bartoc, lobid, regal ...

Beim Beobachten der SWIB14 gewann man mit fortschreitender Zeit immer stärker den Eindruck, dass Technik zum Datenaufräumen ganz groß im Kommen ist und immer leistungsfähiger wird, gleichzeitig aber mit den Entwicklungen ein enormer Bedarf einhergeht, die Produkte selbst auch aufzuräumen. Es gab auf der Konferenz noch viele davon. Jens Mittelbach trug zu d:swarm[15] vor, ein von der SLUB Dresden und dem Dresdner Unternehmen „Avantgarde Labs“ entwickeltes Datenintegrations- und -modellierungswerkzeug. Es hilft bei der „verlustfreien Überführung von Daten aus heterogenen Datenquellen in ein flexibles graphenbasiertes Datenmodell, welches sich explizit in den (Linked) Open Data Webgraphen einbettet“. Nach Aussage der Forscher kanalisiert d:swarm „alle individuellen bibliothekssystemischen Datenflüsse und erzeugt zu jeder bibliothekarischen Ressource einen den Bedürfnissen der jeweiligen Institution entsprechenden Datensatz im Sinne einer Single Version of the truth (SVOT)“. Punkt.

In Lightning Talks, Kurzvorträgen, die bis wenige Minuten vor dem dafür vorgesehenen Zeitfenster angemeldet werden können, kamen dann mit „Vivo“, „Bartoc“ und „regal“ weitere Produktnamen hinzu, die sich jeder merken sollten, der in Bibliotheken mit Linked Data zu tun hat. Vivo[16], erklärt von Ina Blümel, nutzt die Open Source Software VIVO[17] für den Aufbau einer neuartigen Plattform zur Open Science 2.0 Initiative der Leibniz-Einrichtungen. Die Plattform managt Linked Open Data, verlinkt zum Beispiel zu Profilen von Wissenschaftlerinnen und Wissenschaftlern. Die Profile können von Dritten zur Anreicherung eigener Texte genutzt werden. Die vivo-Plattform befindet sich noch im Beta-Stadium. Mit Bartoc[18] führt die Bibliothek der Uni-

14 http://www.w3.org/2004/02/skos/

15 http://www.dswarm.org/datenmanagement/dswarm-hilfe/was-ist-dswarm/

16 https://osl.tib.eu/vivo/

17 http://www.vivoweb.org/about

18 http://www.bartoc.org/de

versität Basel ein neues, multilinguales Verzeichnis digital verfügbarer Thesauri, Ontologien und Klassifikationen. Eine Fundgrube für jeden auf der Suche nach Ordnungskonzepten. Lobid ist der Linked-Open-Data-Dienst des hbz, über dem unter anderem die hbz-Verbunddaten mittels einer Web-API (Schnittstelle für Anwendungsprogrammierung) bereitgestellt werden. Adrian Pohl stellte den neuen Webauftritt der Nordrhein-Westfälischen Bibliographie (NWBib) vor, über den verzeichnete Literatur zum Land Nordrhein-Westfalen ins Web gebracht wird (beta) und der auf der lobid-API basiert.[19]

regal, die dritte im Kurzvortrag vorgestellte Neuentwicklung, ist ein Repository System, das auf LOD-Daten basiert. Felix Ostrowski, Geschäftsführer, des neu gegründeten Unternehmens graphthinking[20] GmbH präsentierte die Software, die für den Archivserver Edoweb[21] des Landes Rheinland-Pfalz entwickelt wird. Ostrowski leitet seit Jahren auf der SWIB einen LOD-Einführungsworkshop.

Kollaboratives Annotieren auf einer Ebene des Web

„Zu sehen zu welchen fantastischen Dingen die Menschheit in ihrem besten Zusammenwirken fähig ist", gibt Dan Whaley als seine größte Motivation an, sich an der Weiterentwicklung des World Wide Web zu beteiligen. Er hat die gemeinnützige Open-Source-Plattform „Hypothes.is" gegründet, die einheitliches und dezentrales kollaboratives Annotieren und kollaborative Begutachtung einer jeden Webressource ermöglichen will. Whaley sieht die Plattform für Open Annotation (beziehungsweise das Zusammenwirken und die Kommunikation, die darüber möglich werden) als Teil einer Entwicklung, aus der völlig Neues entstehen kann.

Hobbyforscher per Crowdsourcing in Wissensaufbau einbinden

Die Computerwissenschaftlerin Cristina Sarasua untersucht in ihrer Doktorarbeit an der Universität Koblenz-Landau wie man Microtask-Crowdsourcing für Data Interlinking benutzen könnte. Ihr Beitrag behandelte das Thema, interessierte Laien bei der Verknüpfung unterschiedlicher Datenquellen einzubinden. „Sie sind nicht alleine. Da draußen im Web gibt es Arbeitskraft und Interesse an ganz vielen Themen", erklärte sie. Die junge Wissenschaftlerin schlägt vor, nach dem Vorbild erfolgreicher Crowdsourcing-Initiativen dieses Wissen und die Arbeitskraft für die große Aufgabe der Verbesserung und Verlinkung von Daten im Web zu nutzen. Sarasua bittet Wissenschaftlerinnen und Wissenschaftler, die sich in Bibliotheken mit dem Thema Interlinking von Datenquellen beschäftigen, ihr mitzuteilen, welche Erfahrungen sie gemacht haben und welchen Problemen sie begegnet sind. Für den direkten Austausch hat sie ein öffentliches Etherpad[22] im Web eingerichtet, das vom twitternden SWIB-Auditorium noch während des Vortrags rege befüllt wurde. Die von Sarasua vorgetragenen Ideen zur Validierung automatisch generierter Links zu Normdateien und Referenzdatenbanken durch Laien und Hobbyforscher fanden große Beachtung und Anerkennung der Fachleute.

Semantisches Web aus Bibliotheken

Richard Wallis stellte am Ende der Konferenz in einer spontan als Einstieg in seinen Vortrag gelieferten Zusammenfassung der SWIB14 ihren Namen in Frage: „Ich habe mich beim Zuhören die ganze Zeit gefragt,

Linked Data verlangen nach Meinung von OCLC-Technologie-Evangelist Richard Wallis eine neue Softwarearchitektur, die das Informationsobjekt in den Mittelpunkt stellt.

ob diese Konferenz den richtigen Titel trägt. Sie heißt „Semantisches Web in Bibliotheken". Aber müsste sie nicht eigentlich heißen: „Semantisches Web aus Bibliotheken"? Recht hat er. Ebenso wie mit seiner ironisch angehauchten Zusammenfassung der „vielen großartigen LOD-Projekte, die jedes Jahr auf der SWIB gezeigt werden" und die er auch von der SWIB im nächsten Jahr erwartet. Die 7. „Semantic Web from Libraries", pardon, „Semantic Web in Libraries" findet vom 23. bis 25. November 2015 in Hamburg statt. #SWIB15 spricht sich auch viel schöner als SWFL.

Vera Münch
Leinkampstrasse 3
31141 Hildesheim
vera-muench@kabelmail.de

19 http://lobid.org/nwbib
20 http://www.graphthinking.com/
21 https://edoweb-rlp.de/resource
22 https://etherpad.mozilla.org/4IfZDaTBle