

Linked Open Data Big Data

Alles vernetzt

Tout en réseau

Tutto in rete

Ausgabe 4 6. Dezember 2013

[a [r [b | i] d] o]

VSA Verein Schweizerischer Archivarinnen und Archivare
BIS Bibliothek Information Schweiz

AAS Association des Archivistes Suisses
BIS Bibliothèque Information Suisse

AAS Associazione degli Archivisti Svizzeri
BIS Biblioteca Informazione Svizzera

webOpac.net

Belebend für Ihre Bibliothek

Mit dem webOpac.net präsentieren Bibliotheken, Ludotheken und Dokumentationsstellen ihren Katalog im Internet – einzeln oder im Verbund. Die umfangreichen Selbstbedienungsfunktionen (Reservieren von Medien, Verlängern von Ausleihfristen, Abfragen des eigenen Kontoauszuges, u.v.m.) machen die Bibliothek für die Benutzer noch attraktiver und rund um die Uhr zugänglich. Und dabei braucht die Bibliothek keine zusätzliche Infrastruktur!



netbiblio

→ **INTEGRIERTE INFORMATIKLÖSUNG
FÜR BIBLIOTHEKEN, MEDIOTHEKEN,
DOKUMENTATIONSSTELLEN
UND ARCHIVE**

Katalogverwaltung, Ausleihverwaltung, WebOPAC, Statistiken, Erwerbung, Zeitschriftenverwaltung, MARC21-Transformationen, Z39.50 Client, Z39.50 Server, Notiz-Server, News-Server, RFID-Integration, Notausleihe, Volltextsuche für externe Dokumente, WebOPAC/Mobile

→ **SOLUTION INFORMATIQUE INTÉGRÉE
POUR BIBLIOTHÈQUES, MÉDIATHÈQUES,
CENTRES DE DOCUMENTATION
ET ARCHIVES**

Gestion de catalogue, Gestion du prêt, WebOPAC, Statistiques, Acquisitions, Gestion des périodiques, Transformations MARC21, Serveur Z39.50, Client Z39.50, Serveur de notices, Serveur de nouveautés, Intégration RFID, Prêt de secours, Recherche en texte intégral pour les documents externes, WebOPAC/Mobile

AlCoda GmbH
Haslerstrasse 21
3008 Bern

026 48 48 020
info@alcoda.ch
www.alcoda.ch



Inhalt / Sommaire

arbido print 4/13

03 Editorial D und F

29 Impressum

I. Grundlagen und Perspektiven /
Bases et perspectives

04 Dimensionen und Zusammenhänge grosser, verknüpfter, offener und wissenschaftlicher Daten

René Schneider, Haute Ecole de Gestion de Genève

07 Linked Open Data: quelles nouvelles compétences pour les professionnels de l'information?

Jasmin Hügi et Nicolas Prongué, spécialistes en information documentaire

09 «Daten öffnen, weltweit kooperieren» – Grosse Fragen, noch keine abschliessenden Antworten?

Vera Munch, freie Journalistin und PR-Beraterin

12 Les institutions patrimoniales suisses sont favorables à l'open data

Beat Estermann, Haute école spécialisée bernoise

15 Les applications basées sur les LOD en bibliothèque: un tour d'horizon

Jasmin Hügi et Nicolas Prongué, spécialistes en information documentaire

17 Alle Bibliotheken können die Zukunft der Katalogisierung gestalten

Eliane Blumer, Haute Ecole de Gestion de Genève, Karsten Schuldt, HTW Chur

II. Projekte und Initiativen aus Bibliotheken und Archiven /
Projets et initiatives dans le domaine archive et bibliothèque

19 Die Bibliographic Framework Initiative (BIBFRAME) der Library of Congress – Offene, verknüpfte Datensätze

Bruno Wenk, HTW Chur

22 Vom Projekt zum Produkt: Der Linked Data Service der Deutschen Nationalbibliothek

Julia Hauser, Deutsche Nationalbibliothek

25 opendata.admin.ch

Andreas Kellerhals, Schweizerisches Bundesarchiv

28 Inscrire les données du Service de la culture dans le Web pour mieux valoriser le patrimoine

Alain Dubois, Archives de l'Etat du Valais

30 Linked Open Data in archivischen Fachinformationssystemen

Peter Sandner, Hessisches Hauptstaatsarchiv, Wiesbaden

III. Projektberichte aus der Hochschulforschung / *Projets de recherches universitaires*

34 swissmetric.ch – neutrale, transparente und vergleichende Webanalyse nicht nur für Bibliotheken

Wolfgang Semar, SII Schweiz. Institut für Informationswissenschaft

36 Wissensvorsprung statt Datenflut mittels Web Monitoring und Web Intelligence

Anton Weichselbraun, SII Schweiz. Institut für Informationswissenschaft

IV. Rezensionen / *Recensions*

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

Stephan Holländer

39 Big Data für Anfänger – ein Reiseführer für «Nichtinformatiker»

Eliane Blumer

41 Eine interessante Einführung zu einer Vielzahl von Themen der Informationswissenschaft

Stephan Holländer

V. Aktualitäten / *Actualités*

42 Anna Pia Maissen zum Abschied / *Eloge de Anna Pia Maissen*

Gregor Egloff

44 Bibliofreak – Freaks und Leidenschaften für Bibliotheken / *Bibliofreak – Une campagne en faveur des bibliothèques*

Gerhard W. Matter

46 BIS-Kongress in Lugano / *Congrès BIS à Lugano*

Katia Röthlin

47 Umfrage *arbido* / *Sondage arbido*

Herbert Staub

48 Antonio Spadafora (1942–2013)



Strichcode-Etiketten für Bibliotheken und Archive

Beschaffen Sie Ihre Barcode-Etiketten beim Spezialisten für Strichcode-Kennzeichnung. Wählen Sie aus unserem breiten Sortiment an spezialisierten Produkten für Bibliotheken und Archive. Profitieren Sie von der langjährigen Erfahrung und nutzen Sie die persönliche Beratung. Unser Printshop ist schnell und zuverlässig.

- › Medienetiketten
- › Archivetiketten
- › Inventaretiketten
- › Sicherungsetiketten
- › RFID-Etiketten
- › Ausweiskarten
- › Spendegeräte
- › Etikettendrucker
- › Handscanner

Appelez-nous au 043 377 30 10!



STRICO AG, Allmendstrasse 14, CH-8320 Fehraltorf, Telefon 043 377 30 10, Fax 043 377 30 11, strico@strico.ch, www.strico.ch

[a [r [b [i [d] o]]]

INSERATE/ANNONCES

- **arbido newsletter**
- **arbido print**
- **arbido website**

→ **inserate@staempfli.com**

→ **031 300 63 89**

ABONNEMENTE

- **arbido print**
- **abonnemente@staempfli.com**
- **arbido newsletter**
- **www.arbido.ch**

Editorial

Linked Open Data und Big Data

Annkristin Schlichte, Stephan Holländer

Linked Open Data und Big Data werden ein integraler Bestandteil der zukünftigen Webinfrastruktur sein, in der grosse Datenmengen verfügbar und via Uniform Resource Identifiers (URI) identifiziert und vernetzt sind. Diese Entwicklung wird auch Auswirkungen auf Archive, Bibliotheken und andere Informationseinrichtungen haben, selbst wenn diese bisher erst in Ansätzen erkennbar sind und das Potenzial von Linked Open Data und Big Data hier noch längst nicht ausgeschöpft ist.

«Linked Data» bezeichnet wohlstrukturierte, hoch konnektive und syntaktisch interoperable Datensets, die über mehrere Repositorien innerhalb einer Organisation oder organisationsübergreifend verteilt sind. In den vergangenen Jahren haben diverse kommerzielle und nichtkommerzielle Initiativen begonnen, Linked Open Data für die kostengünstige Akquise, Anreicherung, Disambiguierung (d.h. die Auflösung von sprachlichen Mehrdeutigkeiten) und Veröffentlichung von Daten zu nutzen und so Einsparungen bei der Datenintegration und der Verbesserung der Qualität und Zugänglichkeit von Daten zu erzielen. Dies gibt auch Archiven und Bibliotheken neue Möglichkeiten, ihre Bestände im Kontext des World Wide Webs zu erschliessen und zu katalogisieren.

In einem aktuellen Heft zur Thematik der Linked Open Data darf aber auch das Thema Big Data nicht fehlen. Darunter werden all jene Datensets verstanden, deren Bewirtschaftung aufgrund des grossen Datenvolumens durch konventionelle Softwaretools und Methoden nicht bewältigt werden kann. Es deutet alles darauf hin, dass unter diesem Begriff die zunehmende Bedeutung von Daten als Faktor der Wertschöpfung zusammengefasst wird. Obwohl «Big Data» je nach Institution unterschiedliche Erscheinungsformen haben, wird die Fähigkeit zur Analyse und Verwaltung von sehr grossen Datenmengen zu einer Schlüsselkompetenz im Umgang mit Daten und zu einem Wettbewerbskriterium.

Informieren Sie sich daher im vorliegenden Heft über die Konzepte von Linked Open Data und Big Data und lassen Sie sich inspirieren von den vorgestellten Projekten!

Linked Open Data et Big Data deviendront une partie intégrante de la future infrastructure du Web, de grandes quantités de données seront ainsi disponibles et pourront être identifiées et mises en réseau via Uniform Resource Identifiers (URI). Ce développement aura également des répercussions sur les archives, les bibliothèques et d'autres structures d'informations, même si ces impacts sont pour l'instant encore peu perceptibles et que le potentiel des Linked Open Data et Big Data n'est de loin pas encore vraiment exploité.

Les «Linked Data» désignent des ensembles de données bien structurés, reliés entre eux et présentant une inter-opérabilité syntaxiques; ces ensembles de données sont répartis sur plusieurs répertoires au sein d'une organisation ou entre plusieurs organisations. Diverses initiatives commerciales ou non ont été lancées ces dernières années dans le but d'utiliser les Linked Open Data pour l'acquisition bon marché, l'enrichissement et la publication de données et parvenir ainsi à des économies au niveau de l'intégration de ces mêmes données, d'en améliorer la qualité et l'accessibilité. Voilà qui donne également aux archives et aux bibliothèques de nouvelles possibilités de mettre en valeur et de cataloguer leurs fonds dans le contexte du World Wide Web.

Ce numéro d'*arbido* consacré aux Linked Open Data aborde également le thème des Big Data. Les Big Data sont des ensembles de données qui deviennent tellement volumineux qu'il n'est plus possible de travailler au moyen des outils et des méthodes classiques de gestion des données. Tout indique que ce concept traduit l'importance croissante des données comme facteurs de création de valeur ajoutée. Bien que les «Big Data» prennent des formes différentes selon les institutions, la capacité d'analyser et de gérer de très grandes quantités de données deviendra une compétence clé et un critère de concurrence.

Familiarisez-vous donc, en lisant ce numéro, avec les concepts de Linked Open Data et de Big Data; peut-être que les projets qui y sont présentés vous inspireront!

I. Grundlagen und Perspektiven

Dimensionen und Zusammenhänge grosser, verknüpfter, offener und wissenschaftlicher Daten

René Schneider, Haute Ecole de Gestion de Genève

Nachdem die Flut der Diskussionen, Kommentare und Prophezeiungen das Social Web betreffend nun verebbt ist, bricht bereits die neue Flutwelle über die Dämme und überspült uns mit dem, was im Social Web und anderswo in Massen ohne Unterlass und Überlegung produziert wird: Daten!

Einleitung

Man gewinnt nicht nur den Eindruck, dass die verhältnismässig alte Bezeichnung der elektronischen Datenverarbeitung eine ganz neue Bedeutung erhalten hat, sondern dass die Daten gleichsam selbst zu Information geworden sind, auch wenn dies auf den ersten Blick als widersinnig erscheinen mag. Dieser paradoxe Eindruck verstärkt sich jedoch, wenn in einflussreichen Kreisen nicht mehr vom Informationszeitalter gesprochen wird, sondern vom Zeitalter der Daten. So wird bereits ein neues und viertes wissenschaftliches Paradigma¹ heraufbeschworen und nach den wissenschaftstheoretischen Vorläufern der empirischen Beschreibung von Naturphänomenen (1. Paradigma), der theoretischen Generalisierung (2. Paradigma) und der computerbasierten Simulation (3. Paradigma) nun das Zeitalter der daten-

intensiven wissenschaftlichen Entdeckung verkündet.

Aber was heisst das eigentlich? Sind die übrigen Paradigmen der Wissenschaftsgeschichte damit bedeutungslos geworden? Und wie verhält es sich mit dem Kontext der Daten, durch den sie normalerweise zu Information werden? Oder sind Daten nicht mehr gleich Daten? Und was heisst dies für Bibliotheken, Archive und Dokumentationszentren? Sind sie oder ihre Daten betroffen? Ergeben sich neue Arbeitsaufträge für sie oder werden sie von einem weiteren Bedeutungsverlust betroffen sein?

Im Folgenden soll versucht werden, diese Fragen einzugrenzen und entsprechend dem derzeitigen Stand der Dinge zu beantworten. Dies soll anhand dreier unterschiedlicher Datentypen illustriert werden, die derzeit unterschiedlich stark diskutiert werden, sich trotz ihrer Unterschiede aber an einzelnen Punkten berühren bzw. in unterschiedlicher Beziehung zueinander stehen: Big Data, Linked Open Data und Research Data (d.h. grosse Datenmengen, verknüpfte und offene Datenmengen sowie Forschungsdaten). Diese dreifache Auffächerung von Datenverarbeitungsinitiativen beantwortet bereits unmittelbar eine der weiter oben formulierten Fragen und zeigt,

Die Informationswissenschaft und alle damit verbundenen Institutionen stehen damit auch vor der Herausforderung, sich auf diese Daten und die damit verbundenen Aufgaben einzustellen.

dass Daten wirklich nicht mehr gleich Daten sind. Die Informationswissenschaft und alle damit verbundenen Institutionen stehen damit auch vor der Herausforderung, sich auf diese Daten

und die damit verbundenen Aufgaben einzustellen.

Big Data

Den grössten Block (im doppelten Sinn) machen dabei die gemeinhin unter Big Data subsumierten Datensätze aus, von deren Existenz und Konsequenz mittlerweile auch die Normalbürger erfahren haben, was insofern gut ist, da sie selbst zu den Hauptproduzenten, bedauerlicherweise aber nicht zu den Nutzern gehören. Big Data wurde dabei letztlich durch zwei Phänomene möglich: erstens die weiterhin exponentiell steigende Leistungsfähigkeit der Prozessoren sowie die damit parallel verlaufende Steigerung der Speicherkapazitäten der Rechner, zweitens durch den Anstieg der täglich bzw. sekundlich produzierten und aufgezeichneten Datenmengen, die allerorts durch die Vernetzung der Endgeräte (vom Grossrechner bis zum Smartphone und den darauf befindlichen Apps) ins Web eingespeist und von dort weiterverarbeitet werden. Gespeichert wird dabei alles, was geht: Daten, mit oder ohne Metadaten, oder die Metadaten alleine, aus denen sich ja streng genommen der eigentliche Kontext der Daten bzw. ihr Verständnis ergibt. Entscheidend ist aber, dass aus diesen Daten ein neuer Kontext erzeugt wird und zwar mithilfe spezieller Algorithmen, genauer mit sogenannten Inferenzmechanismen, also Schlussfolgerungsregeln der Maschinen, die sie verwalten und bearbeiten. In den Schlussfolgerungsregeln des Big Data werden orthogonale, d.h. prinzipiell voneinander unabhängige Merkmale miteinander verknüpft, um neue oder bis dahin unbekannte Zusammenhänge aufzuzeigen². Mit dieser an und für sich recht einfachen Vorgehensweise, die von Menschen täglich bewusst oder unbewusst ausgeführt wird, stellt sich

1 Jim Gray on e-Science: «A transformed Scientific Method», in: Hey, Tony, Tansley, Stewart, Tolle, Kristin: *The Fourth Paradigm: Data Intensive Scientific Discovery*. Microsoft Corporation. 2009, S. xviii http://research.microsoft.com/en-us/collaboration/fourthparadigm/4th_paradigm_book_complete_lr.pdf

2 Dietmar Dath, Ranga Yogeshwar (2013): «Rechnen Sie damit, lebenslang ein Verdächtiger zu sein», FAZ vom 12.7.2013.

der einfache, und von daher sehr erfolgsversprechende Business Case von Big Data dar, der nun aber von den Maschinen ausgeführt wird. Damit sollte auch klar sein, dass hinter Big Data zu einem sehr grossen Teil wirtschaftliche oder andere Interessen stehen, die nicht primär im Interesse der Produzenten sind, auch wenn die ethischen Implikationen an dieser Stelle nicht ausgeführt werden sollen.

Linked Open Data (LOD)

Eher unabhängig davon verläuft die Linked-Open-Data-Initiative³. Unabhängig in dem Sinne, dass die hier verknüpften (Linked) und veröffentlichten (Open) Datensätze nicht notwendigerweise eine Teilmenge der grossen (Big) Daten sind, im Gegenteil: Der Umstand, dass LOD die gemeinnützige Publikation sowie eine Verknüpfung zur *Conditio sine qua non* macht, schliesst den orthogonalen Charakter der grossen Datensätze aus und eröffnet eine ganz andere Sichtweise. Im Gegensatz zu Big Data beschränkt sich LOD auf Daten, die im Web zur Verfügung gestellt werden, und nicht in irgendwelchen nichtöffentlichen Datenbanken oder Clouds gespeichert werden. Des Weiteren ist ihr Umfang, trotz der an und für sich unvorstellbaren Menge der Daten, die die LOD-Cloud mittlerweile enthält, vergleichsweise klein.

Im Rahmen von Linked Open Data soll ein Grundproblem gelöst werden, das darin besteht, dass der grösste Teil des Wissens, das im Web publiziert wird, in relationalen Datenbanken gespeichert, aber nur einem sehr flachen Format (genauer: HTML, der Hypertext Markup Language) repräsentiert wird. Dies führt dazu, dass die ursprünglich gesetzten Relationen verloren gehen und eine direkte Kommunikation (im Sinne eines reziproken Verständnisses) zwischen den Datenbanken nicht möglich ist. Dieses Problem wird nun dadurch gelöst, dass eine neue Datenebene eingeführt wird, die Auskunft über die semantischen Relationen gibt. Der eigentliche Mehrwert von Linked Open Data besteht also darin, dass sie die Relationen aller Datenbanken in ein einheitliches Format überführen, was – genau wie im Fall von Big Data – dazu führen kann, dass die Maschine über

sie rasonieren kann. Wiederum geht es also nicht um kognitive Prozesse von Menschen, sondern darum, diese in Rechnern zu repräsentieren und zu simulieren und den Maschinen zu erlauben, selbstständig Rückschlüsse aus den verknüpften Daten zu ziehen.

Forschungsdaten

Etwas im Abseits dieser grossen Bewegungen findet – auf den Bereich der Wissenschaften beschränkt – seit einigen Jahren eine weitere Diskussion statt, die die sogenannten Forschungsdaten zum Hauptgegenstand des Interesses gemacht hat.

Die gesamten Anstrengungen lassen sich sehr gut an zwei grundlegenden Modellen illustrieren: dem Modell des Datenlebenszyklus⁴, das von ersten Projektideen über die Ablage und dauerhafte Speicherung von Datensätzen bis hin zu deren Integration in Publikationen und deren Zitation reicht, und dem Modell für das Datenkontinuum⁵, das eine kontinuierliche Weitergabe und den Austausch der Daten garantieren soll und zwar unabhängig davon, ob es sich um Rohdaten, strukturierte Daten, Metadaten oder die in späteren Publikationen eingebundenen Daten handelt. Der Mehrwert für den Forscher und die Öffentlichkeit ergibt sich daraus, dass beide quasi uneingeschränkt und über alle Soft- und Hardwarewechsel Zugang zu diesen Daten

Der eigentliche Mehrwert von Linked Open Data besteht also darin, dass sie die Relationen aller Datenbanken in ein einheitliches Format überführen, was dazu führen kann, dass die Maschine über sie rasonieren kann.

für eine Nachnutzung oder Überprüfung haben. Auch wenn man davon ausgehen kann, dass sich das Interesse dafür sowohl bei den meisten Forschern als auch der Gesellschaft, die diese Forschung finanziert, in Grenzen hält, wird eine Sicherung der Forschungsdaten mittlerweile zu den guten und festen Regeln für eine transparente Forschung gezählt. Der grösste Vorteil für den Forscher ergibt sich daraus, dass die Datensätze selbst, d.h. die Primärdaten, zitier- und publizier-

bar werden. In einigen Disziplinen gibt es parallel dazu bereits «Peer Reviewed Data Paper», die allein die Datensätze enthalten.

Was das spezielle Verhältnis von Forschungsdaten zu Big und Linked Open Data betrifft, lässt sich sagen, dass sie – etwa im Fall des Large Hadron Colliders – sehr gross sein können. Sie können auch mit LOD verknüpft und angereichert werden, müssen es aber nicht.

Dabei sollte nicht ausser Acht gelassen werden, dass bei den Forschungsdaten eine ganz andere grundsätzliche Frage im Vordergrund steht: wie bewahren wir (Roh-)Daten langfristig auf, sodass sie uns zu einem späteren Zeitpunkt zur Nachnutzung wieder

Der grösste Vorteil für den Forscher ergibt sich daraus, dass die Datensätze selbst, d.h. die Primärdaten, zitier- und publizierbar werden.

problemlos zur Verfügung stehen? Es geht also in erster Linie um ganz grundsätzliche Fragen der Archivierung, der Kompatibilität und der Interoperabilität, im Grunde also um Fragen, die auch die Bereiche der grossen und der verknüpften Datenmengen irgendwann betreffen könnten. Man kann sich aber auch nicht des Eindrucks erwehren, dass gerade in der Debatte um die Forschungsdaten so prinzipielle Fragen angegangen werden, dass den Beteiligten noch gar nicht genau klar ist, wie diese gelöst werden sollen, auch wenn es zumindest in einzelnen Disziplinen bereits Lösungsansätze bzw. respektable Lösungsvorschläge gibt.

3 Bizer, Christian, Tom Heath, and Tim Berners-Lee (2009). «Linked data-the story so far.» *International Journal on Semantic Web and Information Systems (IJSWIS)* 5:3, 1–22.

4 Sarah Higgins (2008): «The DCC Curation Lifecycle Model». *International Journal of Digital Curation* 3(1), 134–148.

5 Andrew Treloar (2011): *Private Research, Shared Research, Publication, and the Boundary Transitions*. http://andrew.treloar.net/research/diagrams/data_curation_continuum.pdf

Bibliotheken und Daten: eine neue und notwendige Herausforderung

Bleibt die Frage nach der Rolle der Institutionen, die traditionell mit diesen Fragen beschäftigt waren. Zuvorderst handelt es sich dabei um die Bibliotheken, auch wenn die Archive und Dokumentationszentren in einzelnen Bereichen ein Wort mitsprechen könnten, sofern ihre Expertise noch gefragt ist.

Die grossen Datenmengen (Big Data) können die Bibliotheken, Archive und Dokumentationszentren schnell ausser Acht lassen, auch wenn sie traditionell damit beauftragt waren, grössere Informationsmengen aufzubewahren. An Big Data zeigt sich vielmehr, wie schnelllebig die Zeitläufte geworden sind. Wähten sich die Bibliotheken (zumindest einige erlauchte unter ihnen) bis vor Kurzem noch an vorderster Front in der grössten Digitalisierungsinitiative und durften sie sich noch als strategisch wichtige Datenlieferanten der erfolgreichsten Suchmaschine ansehen, müssen sie nunmehr erkennen, dass sie vom Gehilfen des globalen Players zum staunenden Betrachter noch weitaus grösserer Datenverwalter geworden sind.

Anders verhält es sich bei der Linked-Open-Data-Initiative. Sollten die Bibliotheken und ihre Katalogisierungsabteilungen sowie die Verbundzentren dazu bereit sein, würde sich Linked Open Data als der notwendige, wenn nicht sogar zwingende kommende Schritt darstellen, um die Kataloge in ein Format zu bringen, von dem deren Anbieter selbst, aber auch viele andere profitieren können. Schlussendlich könnten damit – neben vielen anderen positiven Seiteneffekten und neuen Nutzungsformen der Katalogdaten – vor allen Dingen die Probleme der Konvergenz unterschiedlicher Verbunddaten sowie die komplizierten Bemühungen für die Erstellung standardisierter Metakataloge ein Ende haben.

6 Augustinus, *Confessiones*, XI, 20, 26, <http://www9.georgetown.edu/faculty/jjod/latinconf/11.html>

7 Jakob Voss (2013): «Was sind eigentlich Daten?» *Libreas. Library Ideas*, 23, <http://libreas.eu/ausgabe23/02voss>

8 John Freely (2009): *Aladdin's Lamp*, Alfred E. Knopf, New York.

Ähnlich verhält es sich mit den Forschungsdaten, auch wenn hierbei weniger die Katalogisierenden als vielmehr die Kollegen aus dem Bereich des E-Publishing sowie die Betreiber der Open-Access-Plattformen eine Rolle spielen werden. Ihre Aufgabe wird es sein, neben den Langzeitarchivaren und im Verbund mit den Datenzentren die Datensätze, so gross oder (eher) klein sie auch sein mögen, mit einer persistenten Adressierbarkeit für eine Nachnutzung, Nanopublikation oder -zitation zur Verfügung zu stellen.

Erinnerung, Betrachtung, Erwartung

Festzuhalten bleibt, dass es insgesamt sehr unterschiedliche Vorstellungen vom Umgang mit Daten gibt, die derzeit in drei unterschiedlich grossen Initiativen angegangen und diskutiert werden, wobei durchaus Berührungspunkte bestehen. Das gegenwärtige Bild der Dateninitiativen zeigt sich so dynamisch, dass derzeit nicht antizi-

Aus dieser Perspektive betrachtet, lassen sich kurioserweise Parallelen zu den drei Dimensionen der Zeit erkennen, die Augustinus im elften Buch der *Confessiones* in seiner Reflektion über das Wesen der Zeit entwickelte.

pierbar ist, ob die Dateninitiativen weiter zersplittern, konvergieren oder etwa zu einem Schichtenmodell zueinander finden können.

Die reine Betrachtung der Gegenwart liefert aber noch keine Antworten auf die zwei grundlegenden Fragen: Warum heben wir die Daten eigentlich auf? Und: Warum interessieren wir uns so sehr für sie? Diese zwei Fragen lassen sich letztlich mit ganz grundsätzlichen menschlichen Bedürfnissen beantworten: erstens einem eher rückwärtsgewandten Bedürfnis, nämlich dem Sammeln und dem daraus resultierenden Bedürfnis nach Erinnerung, zweitens mit der Organisation bzw. der Repräsentation von Wissen, um es immer, d.h. gegenwärtig, zur Verfügung zu haben, und drittens dem eher vorwärtsgewandten Wunsch, Zukünftiges vorhersagen zu können. Aus dieser Perspektive betrachtet, lassen sich kurio-

serweise Parallelen zu den drei Dimensionen der Zeit erkennen, die Augustinus im elften Buch der *Confessiones* in seiner Reflektion über das Wesen der Zeit entwickelte. Interessanterweise sprach er nicht allein von Praeteritum, Praesens und Futurum, sondern von «praesens de praeteritis memoria, praesens de praesentibus contuitus, praesens de futuris expectatio»⁶, d.h., er assoziierte die Erfahrung der Zeit mit drei zentralen Begriffen: Memoria (Erinnerung), Contuitus (Betrachtung) und Expectatio (Erwartung).

So mag es für ein grundlegendes Verständnis hilfreich sein, die unterschiedlichen Konzepte zum Umgang mit Daten schlussendlich mithilfe dieser drei Dimensionen zu begreifen: der Memoria entsprechen die Bemühungen um die Archivierung und Nachnutzung von Daten; die LOD-Initiative zeigt das Bemühen, eine Datenschicht für die gegenwärtige Betrachtung, also den Contuitus, aufzubauen; die Bestrebungen von Big Data entsprechen schliesslich dem Wunsch, die Handlungs- und Gedankenwege anderer vorherzusagen.

Nicht umsonst begann die Wissenschaft der Daten mit Euklids Geometrie, in denen zum ersten Mal von «Dedomena» (dem Gegebenen)⁷ die Rede ist, die in der ersten Übersetzung ins Lateinische eben mit dem Wort Data wiedergegeben wurden. Euklid sah als Dedomena jenes Gegebene an, aus dem sich gesuchte, d.h. nicht bekannt geometrische Zusammenhänge erschliessen lassen. Nichts anderes ist – wenn wir den geometrischen Zusammenhang beiseitelassen – bei Big Data und Linked Open Data der Fall. Hier ist der Zusammenhang so einfach wie offensichtlich.

Wir sollten dabei auch nicht vergessen, dass seit Anbeginn der Wissenschaftsgeschichte praktisch alle namhaften Forscher darum bemüht waren, die Zukunft vorherzusagen, mit welchen Mitteln auch immer. Auch wenn wir es heute nicht mehr wahrhaben wollen, schrieb Newton, der gerne als Vater der modernen Wissenschaften angesehen wird, und darauf verwies «auf den Schultern von Giganten» zu stehen, mehrere Bücher zur Prophetie und der Alchimie⁸.

Ces deux domaines d'intérêt (Prophétie et Alchimie) ont finalement conduit à l'émergence de disciplines telles que l'Astronomie et l'Astrologie, ainsi que de la Chimie. Théoriquement, toute cette discussion sur les données est en fait une recherche de sens au-delà de la simple recherche rationnelle et calculable.

Contact: rene.schneider@hesge.ch

ABSTRACT

Les grandes quantités de données ouvertes, scientifiques et reliées entre elles: dimensions et contextes

Mais que peuvent donc signifier ces énormes quantités de données ouvertes et reliées entre elles? Les autres paradigmes de l'histoire de la science sont-ils donc devenus obsolètes? OU bien les données ne sont-elles plus des données? Et qu'est-ce que cela signifie pour les bibliothèques, les archives et les centres de documentation? Sont-ils concernés ou est-ce leurs données qui le sont? En résulte-t-il de nouvelles tâches pour ces institutions ou leur importance s'en trouvera-t-elle amoindrie?

Dans cet article, l'auteur tente de cerner ces questions et d'y répondre. Il le fait en discutant de trois types de données qui, malgré leurs différences, ont quelques points communs qui les relient d'une certaine manière, à savoir: big data, linked open data et research data (c.à.d. de grandes quantités de données ouvertes et reliées entre elles ainsi que des données de recherche). Ce triptyque d'initiatives de traitement des données répond d'ores et déjà à l'une des questions formulées ci-dessus et montre que les données ne sont plus tout à fait des données. Les sciences de l'information et toutes les institutions qui y sont liées se trouvent donc face à un défi: maîtriser ces données et les tâches qu'elles induisent.

(traduction: sg)

Linked Open Data: quelles nouvelles compétences pour les professionnels de l'information?

Jasmin Hügi et Nicolas Prongué, spécialistes en information documentaire

Les LOD ne représentent pas la première thématique qui remet en question les compétences voire la formation des professionnels de l'information. Ces dernières années, la littérature scientifique a proposé des réflexions par rapport à des programmes d'études en lien avec les bibliothèques numériques¹, la «Digital Curation» (gestion des actifs numériques)² ainsi qu'avec la gestion de données de recherche³.

Des études concernant les nouvelles compétences des bibliothécaires

Les LOD ne représentent pas la première thématique qui remet en question les compétences voire la formation des professionnels de l'information. Ces dernières années, la littérature scientifique a proposé des réflexions par rapport à des programmes d'études en lien avec les bibliothèques numériques, la «Digital Curation» (gestion des actifs numériques) ainsi qu'avec la gestion de données de recherche.

Concernant les LOD en bibliothèque, encore très peu d'études ont été réalisées pour identifier les compétences qui devraient être enseignées aux futurs professionnels de l'information. Deux projets méritent d'être mentionnés dans ce contexte-là. D'une part, un curriculum concernant les LOD en général a été développé pour le projet EUCLID⁴, sans considérer les institutions culturelles. Sur la base de ce curriculum, des matériels d'apprentissage sous forme de webinars, de tutoriels et d'un cahier d'accompagnement ont été élaborés et mis à disposition en ligne. Afin de définir les connaissances nécessaires pour travailler quotidiennement avec les LOD, des experts de différentes industries ont été interrogés. D'autre part, le projet Learning Linked Data⁵ s'est focalisé spécifiquement sur les besoins des professionnels de l'information. Ce projet, qui s'est terminé en septembre 2012, avait pour but de faire un répertoire des sujets d'apprentissage afin de développer une plate-forme soutenant les enseignants et les étudiants dans l'interprétation et la création des données.

Offres d'emploi et avis d'experts

Le Web des données et les LOD représentent un sujet très technique, avec lequel les frontières entre le métier d'informaticien et celui de professionnel de l'information s'estompent. L'étude qui est présentée dans cet article vise à identifier les compétences prioritaires à acquérir pour qu'un professionnel de l'information puisse travailler

- 1 POMERANTZ, Jeffrey, et al. 2006. Curriculum development for digital libraries. Proceedings of the 6th ACM/IEEE-CS joint conference on Digital libraries. New York, NY, USA: ACM. DOI: 10.1145/1141753.1141787.
- 2 YAKEL, Elizabeth, et al. 2011. Digital Curation for Digital Natives. Journal of Education for Library & Information Science. Vol. 52, no. 1, p. 23-31.
- 3 STANTON, Jeffrey M., et al. 2011. Education for eScience Professionals: Job Analysis, Curriculum Guidance, and Program Considerations. Journal of Education for Library and Information Science. Vol. 52, no. 2, p. 79.
- 4 www.euclid-project.eu/
- 5 <http://lld.ischool.uw.edu/wp/>

avec des LOD en bibliothèque. Pour ceci, des offres d'emploi trouvées en Suisse, en Allemagne, en Autriche et en France ainsi que des avis d'experts internationaux ont été analysés afin de dresser un premier constat. Seules les offres d'emploi exigeant un diplôme en science de l'information, en plus des compétences liées aux LOD, ont été retenues. Quant aux experts, ils devaient, pour être sélectionnés, avoir étudié les sciences de l'information ou avoir travaillé en tant que bibliothécaire, en plus de leur expérience en LOD⁶. Au final, huit offres d'emploi et douze réponses d'experts ont été prises en compte.

Huit compétences-clés

La nature technique des LOD avait une certaine influence sur les réponses. Ainsi, ces dernières mentionnent beaucoup de détails techniques, tels que des langages, des formats ou encore des logiciels. L'analyse du contenu des offres d'emploi et des réponses a permis d'identifier plus de 40 éléments de compétence différents. La figure 1 présente les compétences qui apparaissent dans le nombre de cas le plus élevé⁷.

Ontologies et métadonnées

Dans cette première catégorie, il s'agit de connaître des schémas de métadonnées tels que Dublin Core, savoir évaluer et choisir un élément de métadonnées pour l'appliquer à ses propres données, savoir créer des vocabulaires contrôlés ainsi que des ontologies, et savoir effectuer un alignement de deux vocabulaires différents.

Développement web

Cet élément englobe des connaissances concernant le fonctionnement des protocoles web et des serveurs, ainsi que les principes du Web design. La maîtrise de

ces connaissances est nécessaire afin de savoir créer et attribuer des URIs aux triplets, les publier en ligne ainsi que créer des interfaces web pour que les utilisateurs puissent interagir avec.

Triplets

En troisième position arrivent les compétences en lien avec les triplets, c'est-à-dire la connaissance des spécifications RDF émises par le W3C, qui représentent la base pour la création de triplets. Il est également important de savoir développer et modéliser un graphe qui peut être appliqué aux données. Puisque les triplets peuvent être exprimés à travers différentes syntaxes (RDF/XML, Notation 3, etc.), il est utile de comprendre ces dernières ou de connaître des outils qui peuvent les produire.

Web sémantique et Linked Data

Il s'agit de connaître les principes émis par le W3C et par Tim Berners-Lee, ainsi que leurs objectifs. Cette catégorie inclut également la compréhension basique de l'inférence ou du raisonnement, par exemple.

Standards de catalogage

Les compétences concernant les standards de catalogage sont essentielles dans un projet LOD, car pour transformer des données en triplets, il faut bien connaître les données de base. Il s'agit donc de comprendre les règles de saisie des données (telles que les AACR2), le format d'enregistrement (tel que MARC), ainsi que le fonctionnement du SIGB utilisé (tel que Virtua).

Programmation

Certains experts interrogés ont évoqué la connaissance basique d'au moins un langage de programmation (Java, Ruby, PERL, etc.) comme compétence nécessaire. Cela servirait avant tout de moyen de communication avec les informaticiens. Ces connaissances peuvent également être utiles pour la programmation de codes simples qui permettent la conversion en masse de données en LOD.

Technologies du Web sémantique

Les données organisées en triplets nécessitent une base de données spécifique (un «triple store»), un langage d'interrogation spécifique (SPARQL) et une interface d'interrogation spécifique (SPARQL Endpoint). La maîtrise de ces technologies est fortement souhaitée pour travailler avec les LOD.

XML

Enfin, dernier élément mais non des moindres, des connaissances de XML et des langages annexes comme XSLT ont été mentionnées dans les cas analysés. XML est à la base de nombreux standards du Web sémantique, et il est donc indispensable de comprendre et de savoir exploiter des données structurées en XML.

Un monde changeant

Le travail avec des LOD en bibliothèque nécessite visiblement des compétences très spécifiques et au premier abord, cela peut sembler intimidant. Néanmoins, il est important de noter que les compétences déjà existantes parmi les

- 6 La méthodologie sera expliquée en détail dans un rapport qui paraîtra à la fin de cette année.
- 7 Pour donner un exemple: si une compétence est mentionnée dans chacun des 20 cas, la compétence a une haute priorité. Si une compétence n'apparaît que dans un seul cas, sa priorité est basse (même si la compétence est mentionnée très souvent à l'intérieur du cas).

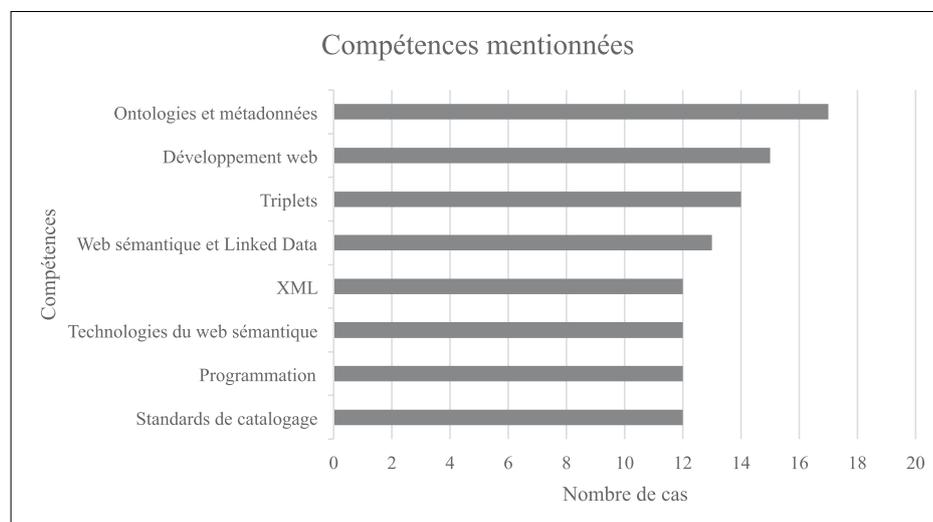


Figure: Compétences les plus mentionnées dans les 20 cas

professionnels de l'information se retrouvent tout à fait dans les éléments cités dans le diagramme. Ainsi, la première compétence touche aux ontologies et aux métadonnées, qui sont utilisées pour la description des ressources, et ceci a toujours fait partie des connaissances de base des professionnels de l'information. Concernant les standards de catalogage, il est inutile d'évoquer que cela appartient à la bibliothéconomie. En outre, les formations de bachelors en information documentaire en Suisse contiennent des modules d'enseignement de développement web et de XML. Par rapport à la programmation, la frontière entre informaticien et professionnel de l'information se montre fluctuante, et l'attribution de cette compétence à l'un des deux métiers représente un sujet de discussion.

Ce qui manque réellement en ce moment, ce sont les compétences stricte-

ment liées au Web sémantique (principes, standards et technologies). Selon nous, l'enseignement de ces compétences devrait être inclus dans le curriculum des professionnels de l'information. Une introduction obligatoire à la thématique ainsi que des cours à choix devraient être dispensés au niveau bachelors, alors que les étudiants au niveau master devraient atteindre une compréhension globale afin d'être capables de prendre des décisions stratégiques.

Des cours de formation continue sont par ailleurs à mettre en place pour que les professionnels sur le terrain puissent eux aussi se former aux LOD. En effet, le Web sémantique est bien plus qu'une tendance, et le monde des bibliothèques ne peut se permettre d'attendre.

Contact: nicolas.prongue@etu.hesge.ch
jasmin.hugi@hesge.ch

ABSTRACT

Linked Open Data – Welche neuen Kenntnisse brauchen I&D-Fachleute auf diesem Gebiet. Die Linked Open Data (LOD) verbreiten sich progressiv in der kulturellen Domäne. Eine wachsende Anzahl an Projekten entstehen in Institutionen und LOD-Applikationen halten Einzug in Bibliotheken. Dieser Einzug stellt gewisse Kompetenzen der I&D-Fachleute infrage, stärkt andere und erfordert neue Kenntnisse und neues Know-how. Dieser Artikel präsentiert die Resultate einer Studie, die zum Ziel hat, diese neuen Kompetenzen, die sich I&D-Fachleute angesichts LOD aneignen müssen, zu bestimmen.

«Daten öffnen, weltweit kooperieren» Grosse Fragen, noch keine abschliessenden Antworten

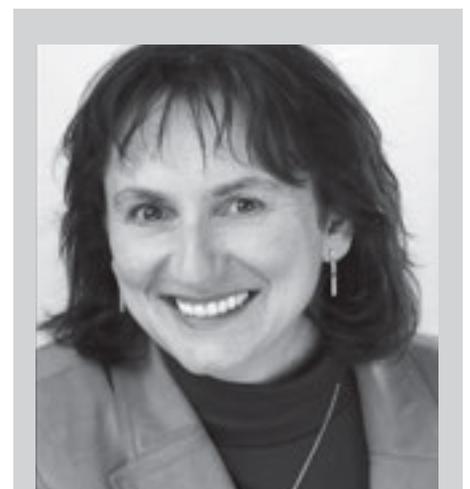
Vera Münch,
freie Journalistin und PR-Beraterin

Die neuen Möglichkeiten der nahezu unbegrenzten Verlinkung von Daten, die Aggregation und Analyse riesiger Datenmengen – diskutiert unter den Überbegriffen Big Data, Linked Data und Linked Open Data (LOD) – bieten Bibliotheken völlig neue Chancen, ihre Bestände und besonderen Angebote international sichtbar zu machen. Die neue Technik des Cloud Computings, also der Bereitstellung und Nutzung von Software, Rechen- und Speicherkapazitäten als Onlinedienstleistung, ermöglicht es zudem, Routinearbeiten im eigenen Haus zu reduzieren und sie stattdessen durch Automatisierung, Vernetzung und Abgleich von z.B. Katalogen und Metadaten auf viele Schultern in der Bibliothekswelt zu verteilen. Auch die Informationsauslieferung auf mobile Computer und Smartphones lässt sich über Cloud Services relativ problemlos organisieren.

Soll man seine Daten wirklich für die ganze Welt öffnen? Wie sichert man in einer völlig offenen Welt die nationalen Interessen, wie die des eigenen Hauses? Wer sorgt für eine vernünftige, nachhaltige Ordnung in der Datenablage, der Datenverknüpfung und der Informationspräsentation? Wie kann man es als Bibliothek schaffen, mit dem atemberaubenden Forschungsfortschritt mitzuhalten? Viele grosse Fragen, auf die es heute noch keine abschliessenden Antworten gibt.

Globale Clouds: Problemlösung oder Herstellerabhängigkeit?

Als eine der ersten Bibliotheken in Europa hat die Universität Tilburg Services der cloudbasierten Softwareinfrastruktur WorldShare™ Management Services (WMS) von OCLC in Betrieb genommen: «Wir betreiben, historisch gewachsen, seit Jahren mehrere Bibliotheksinformationssysteme (LIS) paral-



Portrait

Vera Münch (55), Hildesheim (D) begleitet als freie Journalistin und PR-Beraterin für wissenschaftliche Einrichtungen die Entwicklungen in der Fachinformation und im Bibliothekswesen. Sie berichtet für die deutsche Bibliotheksfachzeitschrift b.i.t.online über Konferenzen, Messen und Trends.

lel. Mit der Cloud ist jetzt die Technologie da, hier etwas zu ändern», so ein Mitarbeiter.

Durch die Nutzung einer Cloud-Plattform bindet man sich technisch eng an einen Anbieter. Entsprechende Befürchtungen (in Fachkreisen unter dem Begriff «Vendor Lock-In») diskutiert, begleiten die neue Technologie. Anja Smit, Direktorin der niederländischen Universitätsbibliothek Utrecht und amtierende Vorsitzende des OCLC EMEA Regional Council hielt den Bedenken in Strassburg das Verbundargument entgegen: «Alle Macht geht von den Mitgliedern aus».

Die Schweizerische Nationalbibliothek hat eine Strategie

Die Direktorin der Schweizerischen Nationalbibliothek (NB), Marie-Christine Doffey¹, hat mit ihrem Team für ihr Haus eine klare Entscheidung getroffen. In ihrem Konferenzvortrag² brachte sie diese auf den Punkt: «Daten öffnen, weltweit kooperieren». Das bedeutet, die NB kooperiert nicht nur mit den Schweizer Bibliotheken wie zum Beispiel im Rahmen des Metakatalogs Swissbib und «den anderen etwa 900 Bibliotheksdiensten, die im Metakatalog sind», sondern auch mit der Europeana, mit dem WorldCat «und allen anderen Möglichkeiten», so Doffey. Sie begründete die Entscheidung vor allem mit der Überzeugung, dass alle Metadaten und Informationsbestände der Schweizerischen Nationalbibliothek

mit öffentlichen Geldern aufgebaut sind und deshalb auch der Öffentlichkeit zur Verfügung stehen müssten; sei es vor Ort in der Bibliothek oder virtuell. «Durch die Öffnung können andere Bibliotheken und Interessenten die Metadaten für eigene Zwecke und Innovationen nachnutzen.»

Die NB hat für die Europeana den Schweizer Bibliothekskatalog Helveticat vollumfänglich geöffnet. Alle Metadaten sind unter der Corporate-Commons-Lizenz CC-O für alle Interessenten verfügbar und können angebunden werden. Doffey erklärte: «Wir müssen dort sein, wo andere Partner, andere Bibliotheken auch dabei sind, damit wir an der Diskussion teilnehmen können. Die Schweiz ist so klein, dass wir diese Öffnung auch weltweit nutzen wollen und können.» Genevieve Clavel-Merrin, bei der NB verantwortlich für den Bereich «Nationale und Internationale Kooperation» ergänzte, durch den Big-Data-Austausch auf Basis vernetzter Metadaten würde ein Bestandsabgleich mit grossen anderen Metadatenansammlungen möglich. Zur Veranstaltung sagte Direktorin Doffey, dass auch «die ganze Diskussion über die Cloud» sehr wertvoll sei. «Als einzelne Institution haben wir zwar nicht die gleichen Probleme wie Konsortien, aber ich denke, es ist auch einmal ganz interessant, Leute zu hören wie Raymond Bérard, zu erfahren, was es da für Probleme gibt und welche Fragen gestellt werden.» Bérard, Direktor des französischen staatlichen bibliografischen Instituts ABES berichtete u.a., dass ABES für die französischen akademischen Bibliotheken eine Studie zur Ablösung der bisherigen LIS durch ein neues, geteiltes «shared LIS» durchgeführt hat. Auf Basis der gewonnenen Erkenntnisse hat ABES sein Bibliothekssystem neu ausgeschrieben.

Frankreich setzt stark auf Wikipedia für Bibliotheken

Wie Wikipedia in Bibliotheken eingebunden und von ihnen genutzt werden kann, ist in Bezug auf Big Data und Linked Data ebenfalls ein grosses Thema. Die Wissenschaftlerin Titia van der Werf, die vom niederländischen Leiden aus für OCLC die F&E-Aktivitäten mit europäischen Bibliotheken koordiniert,

griff die Diskussion mit Rémi Mathis als geladenem Gast und Max Klein von OCLC Research³ in einem Workshop auf. Mathis ist Bibliothekskurator und Präsident der Wikimedia France. Er berichtete, dass das französische Kultur- und Kommunikationsministerium mit Wikimedia France kooperiert, um die französischsprachigen Ressourcen in Wikipedia semantisch zu erschliessen. Klein beschäftigt sich als «Wikipedian in Residence» bei OCLC Research mit allen Fragen rund um die Weiterentwicklung von Wikipedia als Quelle für verlinkbare Daten und Informationen. Er versuchte, in 45 Minuten die Idee der freien, von Maschinen und Menschen gleichermaßen lesbaren Datenbank Wikidata zu erklären – und wie die Forscher dieser Datenbank Multilingualität beibringen wollen.

Culturomics: Kulturgeschichte aus Digitalisaten herausrechnen

Auf diesem hohen Niveau internationaler Forschung ging es nahtlos weiter. Jean-Baptiste Michel erforscht in Harvard «die Nutzung von Millionen digitalisierter Bücher und Dokumenten, um durch quantitative Analysen daraus z.B. Rückschlüsse zur Geschichte und Kultur zu ziehen». «Culturomics»⁴ haben er und sein Kollege Erez Lieberman Aiden dieses neue, auf Digitalisaten aufsetzende Forschungsgebiet getauft. Ihre Arbeiten führten zu einer Software, die aus grossen Datenbeständen in den Daten verborgene Trends berechnet und grafisch als Kurven in Diagrammen sichtbar macht. Die Forschung brachte ihnen im letzten Jahr Titelseiten auf «Nature and Science» und auf der «New York Times». Google hat das Programm gekauft.

In Deutschland entsteht ein «Culturegraph»

Die Deutsche Nationalbibliothek (DNB) baut an einer Plattform, über die Dienste und Projekte zu Datenvernetzung, Persistent Identifier und Linked Open Data für kulturelle Entitäten zusammengeführt werden sollen. Markus Geipel, Softwareingenieur bei der DNB, stellte die «Culturegraph» getaufte Plattform vor. Sie soll zu einem globalen Kulturschatz-Netzwerk weit über die Grenzen einzelner Unternehmen und Einrichtungen hinaus werden.

1 www.nb.admin.ch/org/amtsvorsteherin/index.html?lang=de

2 Aufzeichnungen der Konferenzbeiträge sind, soweit sie von den Referenten freigegeben wurde, auf der Website von OCLC bereitgestellt. www.oclc.org/en-europe/events/2013/emearc-annual-meeting-feb-26.html. Von einigen Präsentationen gibt es zudem Videoaufzeichnungen auf YouTube, u.a. vom Vortrag von M.-C. Doffey <https://www.youtube.com/watch?v=aONFHqMbdig> <http://www.oclc.org/en-europe/events/2013/emearc-annual-meeting-feb-26.html>

3 OCLC-Research betreibt eine eigene Website, auf der sehr interessante Forschungsprojekte und -ergebnisse stehen www.oclc.org/research.html

4 www.culturomics.org/

Neue Informationsdienste mit offenen Daten

Wie auf Basis offener Daten mit neuen IKT entwickelte Dienste aussehen können, zeigten nach Jean-Baptiste Michel zwei weitere Referenten auf sehr beeindruckende Weise: Silver Oliver, Informationsarchitekt mit Spezialgebiet strategische Ausrichtung der Datennutzung für Onlinepublishing bei der britischen Unternehmensberatung Ontoba⁵, und Dr. Klaus Ceynowa, stellvertretender Direktor der Bayerischen Staatsbibliothek (BSB). Oliver referierte über das «Zeitalter des Hinweisens auf Dinge, die in Beziehung zum persönlichen Interesse stehen» («The age of pointing at things»). Gemeint ist das automatische Anbieten von weiterführenden Informationen direkt in den Informationsraum hinein, in dem sich der Nutzer gerade bewegt.

Dr. Klaus Ceynowa zeigte innovative Nutzungsszenarien für digitalisierte Inhalte aus Bibliotheksbeständen, die von der BSB mithilfe von Technologien wie GPS-gestützte Lokalisierung (Location based Services), durch virtuelle Darstellung angereicherte Informationsbereitstellung (Augmented Reality) sowie gestengesteuerte Präsentation und Informationssuche (Gesture based Computing and Search) entwickelt wurden.

OCLC Research informiert umfassend über Datennutzung

Nicht nur die eingeladenen Redner, auch OCLC Research hatte Spannendes aus den Forschungslaboren zu berichten. Roy Tennant, Senior Programm Officer, führte in die «Entfesselung des Katalogisierens»⁶ durch Datamining in LOD ein. Die über 290 Millionen WorldCat-Einträge, die auf rund 1,95

Milliarden Holdings hinweisen, liefern den Rohstoff für schier unendliche Nutzungsmöglichkeiten. Van der Werf zeigte in ihrem Vortragsthema «Metadaten ausser Kontrolle: Metadatenaggregation auf Netzwerkebene», welche Aufgabe gute Metadaten erfüllen und wo die Grenzen der Technik liegen.

W3C Group für LOD-Erweiterung von Schema.org eingerichtet

Last, but not least sprach OCLC Technology Evangelist Richard Wallis im letzten Vortrag über «Linked Data und OCLC»⁷. Dabei berichtete er unter anderem, dass das World Wide Web Consortium (W3C) an der Weiterentwicklung des Metadatenbeschreibungsstandards Schema.org in Richtung LOD-Nutzung arbeitet. Dazu wurde eine «Schema Bib Extend Community Group»⁸ eingerichtet. Wallis leitet sie als Vorsitzender.

Kontakt: vera-muench@kabelmail.de

ABSTRACT

Ouvrir les données et coopérer au niveau mondial

Les nouvelles possibilités que présentent la mise en réseau pratiquement illimitée des données, l'agrégation et l'analyse d'énormes quantités de données (Big Data, Linked Data et Linked Open Data) offrent aux bibliothèques de nouvelles chances de valoriser leurs fonds à l'échelle internationale et de proposer des offres spécifiques. La nouvelle technique du Cloud Computing, autrement dit la mise à disposition et l'utilisation de logiciels, de capacités de calcul et de stockage comme prestation en ligne, permet en outre de réduire les travaux de routine et de répartir ces derniers via l'automatisation et la mise en réseau de catalogues et de métadonnées. La fourniture d'informations sur des ordinateurs portables et des téléphones intelligents peut également être organisée sans trop de problèmes via les Cloud-Services. (traduction: sg)

OCLC hat sein 4. EMEA Regional Council Meeting 2013 unter das Thema «Dynamic Data: a World of possibilities» gestellt.

5 www.ontoba.com/blog/bbc-news-labs

6 www.oclc.org/research/news/2013/02-27.html

7 Ein Interview mit Richard Wallis zu LOD erschien in B.I.T.-Online 15 (2012), Ausgabe 6, S. 566

8 www.w3.org/community/schemabibex/

[a [r [b | i] d] o]

ANNONCES/INSERATE

– arbido newsletter

– arbido print

– arbido website

inserate@staempfli.com

031 300 63 89

Les institutions patrimoniales suisses sont favorables à l'open data

Beat Estermann,
collaborateur scientifique,
Institut pour le gouvernement
électronique de la Haute école
spécialisée bernoise

Depuis l'avènement de l'Internet, les institutions patrimoniales connaissent des changements successifs nombreux. On par exemple la numérisation systématique des objets patrimoniaux, la mise en réseau accrue des organisations, la coproduction de contenus par les utilisatrices et les utilisateurs ainsi que la mise à disposition de données et de contenus sur l'Internet sous des licences «libres». Une étude de la Haute école spécialisée bernoise¹ montre où les institutions patrimoniales suisses se situent dans ce contexte.

Les institutions patrimoniales en pleine mutation

Le Musée de Derby mise sur des codes QR pour enrichir via des articles de Wikipédia les explications concernant les objets d'exposition, et ce en plusieurs langues bien entendu. La Bibliothèque nationale australienne propose

à des bénévoles des tâches de correction et de transcription sur l'Internet, offre qui rencontre un véritable succès. Le Tropenmuseum à Amsterdam met des centaines de milliers de photos haute résolution sur Wikimedia Commons, se félicite que ses fonds soient désormais plus faciles d'accès et en espère en tirer profit par l'augmentation du nombre de visiteurs. Le Brooklyn Museum s'en remet au jugement des visiteurs de son site web pour sélectionner des photographies destinées à une exposition. Un consortium de bibliothèques européennes utilise le crowdfunding pour financer la numérisation des œuvres tombant dans le domaine public. A Vienne, des chercheurs misent sur les données libres d'accès pour visualiser d'une nouvelle façon d'appréhender l'histoire de l'art.

Ces exemples montrent les profonds changements qui traversent le champ des institutions patrimoniales depuis l'avènement de l'Internet dans les années 1990 et les innovations qu'il a induites.

On peut, concrètement, identifier plusieurs tendances successives:

Depuis le début des années 2000, les institutions patrimoniales en Europe conjuguent leurs efforts afin de numériser de manière coordonnée et la plus complète possible les objets patrimoniaux et leurs métadonnées. Outre une mise en réseau renforcée qui permet aux institutions d'échanger leur savoir-faire et d'éviter les doublons, sont également proposées dans la foulée des offres de «single-point-of-access» sous la forme de systèmes de catalogues inter-organisations et de bibliothèques virtuelles.

Dès 2005, on assiste à l'avènement du Web social qui offre encore plus de possibilités d'interaction. Cette tendance se caractérise par de nouvelles formes de personnalisation et par une implication plus forte des utilisatrices et des utilisateurs, ou des visiteuses et des visiteurs respectivement, qui sont appelés à coproduire.

Le développement spectaculaire de Wikipédia, Flickr et d'offres analogues basées sur le principe de la communauté virtuelle focalise de plus en plus l'attention sur la constitution et la collecte collaborative de contenus sur l'Internet. De premières collaborations s'installent entre des institutions patrimoniales et des communautés en ligne, tandis que certaines institutions lancent des projets spécifiques en faisant appel au crowdsourcing².

Les communautés Wikipédia et Flickr découvrent de leur côté que les institutions patrimoniales sont des partenaires importants et cherchent depuis quelques années à collaborer activement avec elles. A partir de 2009, c'est le mouvement pour l'ouverture des données publiques qui, depuis les

Le développement spectaculaire de Wikipédia, Flickr et d'offres analogues basées sur le principe de la communauté virtuelle focalise de plus en plus l'attention sur la constitution et la collecte collaborative de contenus sur l'Internet.

Etats-Unis et la Grande-Bretagne, se propage rapidement pour toucher quelque 50 Etats de par le monde. Son exigence: les données des autorités publiques doivent être rendues accessibles sous une forme lisible par ordinateur afin qu'elles puissent être utilisées par des tiers. Si le mouvement pour l'ouverture des données publiques trouve son origine dans une vision de la politique étatique que l'on veut plus transparente et plus participative, il est également motivé par une vision technico-économique: la mise en réseau du plus grand nombre possible de données «ouvertes» doit permettre l'émergence de ce qu'il est convenu d'appeler le «Web sémantique». Les données des autorités publiques ainsi que, le cas échéant, celles d'entreprises privées, deviennent ainsi des ressources infrastructurelles libres d'accès que des tiers

¹ En allemand: Estermann Beat (2013): Schweizer Gedächtnisinstitutionen im Internet-Zeitalter. Ergebnisse einer Pilotbefragung zu den Themenbereichen Open Data und Crowdsourcing, Berner Fachhochschule, E-Government-Institut. / En anglais: Estermann Beat (2013): Swiss Heritage Institutions in the Internet Era. Results of a pilot survey on open data and crowdsourcing, Bern University of Applied Sciences, E-Government Institute.

² Le concept de «crowdsourcing» désigne des activités participatives en ligne dans le cadre desquelles une personne ou une organisation demande publiquement à un groupe (en général non défini de manière précise) de personnes d'horizons différents d'accomplir une tâche à titre bénévole. On parle également de «coproduction» ou, en particulier dans le contexte de Wikipédia, de création collaborative de contenus.

peuvent utiliser pour offrir des services à valeur ajoutée: pour la valorisation des données, la création et la visualisation de combinaisons de données d'un nouveau genre ou pour d'autres prestations basées sur des données. L'utilisation de licences standardisées «libres», qui permettent également de modifier et de diffuser plus loin les œuvres à des fins commerciales, est primordiale dans ce contexte, étant donné qu'elles réduisent sensiblement les coûts de transaction pour l'utilisation d'œuvres par des tiers³.

Où se situent aujourd'hui les institutions patrimoniales suisses par rapport à l'open data/open content et au crowdsourcing?

Afin de montrer où se situent actuellement les institutions patrimoniales suisses par rapport à l'open data / open content et au crowdsourcing, comment elles évaluent les chances et les risques de ces tendances et où elles voient leurs bénéfices potentiels, la Haute école spécialisée bernoise a réalisé une enquête-pilote à la fin 2012. Quelque 200 institutions de Suisse alémanique d'importance nationale ont été sollicitées dans ce contexte⁴, dont 72 ont répondu au questionnaire.

Environ 60% des institutions interrogées mettent, à tout le moins en partie, à disposition sur l'Internet des objets patrimoniaux et des métadonnées. Quelque 40% échangent des métadonnées dans le cadre d'alliances. Ces deux pratiques, qui représentent des conditions importantes du point de vue de l'open data / open content et du crowdsourcing, sont donc déjà très répandues. Les institutions qui ne sont pas encore parvenues à numériser leurs fonds les plus importants, pour des raisons de manque de moyens ou de savoir-faire, peuvent toutefois causer quelque souci.

On constate deux dynamiques différentes en ce qui concerne l'open data / open content et le crowdsourcing. Pour ce dernier, on constate un pourcentage certes relativement élevé (10%) des institutions interrogées qui disposent déjà de collaborateurs qui apportent leur contribution à Wikipédia dans le cadre de leur travail, ainsi qu'un pourcentage

tout aussi élevé de participants à l'enquête qui estiment que le travail bénévole sur l'Internet joue un rôle plutôt important pour leur organisation. Cela étant, la plupart des institutions interrogées estiment que le crowdsourcing recèle davantage de risques que de chances. Il n'en va pas de même pour ce qui est de l'open data/open content: à la fin 2012, il n'y avait certes que très peu d'institutions qui étaient disposées à mettre des contenus disponibles en ligne libres de droit, mais les institutions interrogées estimaient néanmoins que les chances qu'offre une telle approche étaient bien réelles. Le feu, pour l'open data, est donc pour ainsi dire au vert! Comme l'on s'y attendait, les linked data constituaient également déjà un thème pour quelques-unes des institutions interrogées, mais aucune d'elles n'avait toutefois de projet dans ce contexte à la fin 2012.

Open data/open content: besoin d'information concernant les licences «libres»

Les résultats de l'enquête permettent de conclure qu'à ce jour moins de 10% des institutions interrogées mettent «librement» à disposition sur l'Internet des illustrations de leurs objets patrimoniaux, mise à disposition qui permet, autrement dit, à des tiers de les modifier et de les utiliser gratuitement. Plus de la moitié mettent certes des illustrations à disposition sur le Net, mais en limitent toutefois leur utilisation. Le plus souvent, les institutions sont prêtes à mettre gratuitement à disposition des contenus pour la formation et la recherche, pour une utilisation privée ou des projets d'intérêt public. Elles sont nettement plus réticentes lorsqu'il s'agit d'autoriser la mise en ligne de contenus qui peuvent être utilisés à des fins commerciales ou faire l'objet de modifications.

Par rapport aux licences «libres», il existe en outre un besoin d'information: près de la moitié des institutions qui rendent des objets patrimoniaux accessibles en ligne ne semblent pas être conscientes du fait qu'il n'est pas possible de libérer des œuvres pour leur utilisation dans Wikipédia et en même temps d'interdire à des tiers de modifier des images ou de les utiliser à des fins commerciales.

Plus de la moitié des institutions interrogées considèrent l'open data comme important. Les groupes cibles les plus importants dans ce contexte sont les institutions de recherche et de formation, les personnes privées et d'autres institutions culturelles. Les institutions interrogées estiment que les plus grandes chances de l'open data/open content résident dans la meilleure visibilité et accessibilité des fonds, dans la meilleure visibilité des institutions elles-mêmes ainsi que dans leur meilleure mise en réseau. Concernant les risques de l'open data, les institutions interrogées craignent surtout des dépenses supplémentaires en ressources et une certaine perte de contrôle sur les données, en particulier pour ce qui est des violations des droits d'auteur, de la protection des données et des infractions contre les règles de confidentialité.

Crowdsourcing: important mais porteur de risques

Près de 40% des institutions interrogées considèrent le crowdsourcing comme étant important. La mise en œuvre du crowdsourcing, c'est-à-dire la possibilité de créer ou de compléter des contenus sur l'Internet par des bénévoles qui collaborent entre eux, est considérée par ces mêmes institutions comme envisageable surtout dans les domaines de la classification et de la complétion des métadonnées ainsi que pour des tâches de correction et de transcription. Près de la moitié des institutions constatent par ailleurs des besoins d'amélioration dans le domaine des métadonnées, en particulier pour ce qui est de l'intégralité, de la disponibilité et de la numérisation de ces dernières.

En revanche, les institutions interrogées estiment que le crowdsourcing présente lui aussi une série de risques. Parmi ces derniers figurent la charge

³ P. ex. les licences de droits d'auteur de Creative Commons, telles que CC-by ou CC-by-sa (pour plus d'informations à ce sujet, voir: www.creativecommons.org)

⁴ L'inventaire suisse des biens culturels d'importance nationale a servi de base pour la sélection des institutions.

importante qui y est liée au niveau de la préparation et du suivi. Autres risques: la continuité incertaine de l'entretien des données, de la difficulté d'influer sur les résultats et d'estimer correctement la charge temporelle ainsi que la faible sécurité en matière de planification.

Conclusions

Les résultats de l'enquête montrent que miser sur des approches de type crowdsourcing implique une certaine charge et des impondérables. Cela s'explique également par le fait que les institutions doivent participer activement aux communautés de bénévoles et de partenaires impliquées, ce qui ne manque souvent pas de poser des problèmes compte tenu des différentes cultures organisationnelles. Les processus d'intégration qui y sont liés peuvent certes, en fin de compte, déboucher sur un bénéfice pour les deux parties, mais cela ne va en général pas sans un investissement temporel considérable.

En contrepartie, opter pour une stratégie open data/open content est nettement plus simple pour les institutions. Il est donc conseillé de miser, à court terme et à grande échelle, surtout sur des approches qui permettent aux institutions la « libre » mise à disposition des données et des contenus, sans qu'elles aient à assumer elles-mêmes les dépenses qu'impliquent de longs processus de coopération. Néanmoins, un certain niveau de standardisation des métadonnées, de la qualité des données et des interfaces techniques pourrait malgré tout être utile pour réduire au minimum les coûts de transaction. Il s'agit également de mettre l'offre en rapport avec la demande, car si tel n'est pas le cas, on court le risque que les données et les contenus mis à disposition restent inutilisés.

Compte tenu du fait que de nombreuses institutions considèrent la thématique

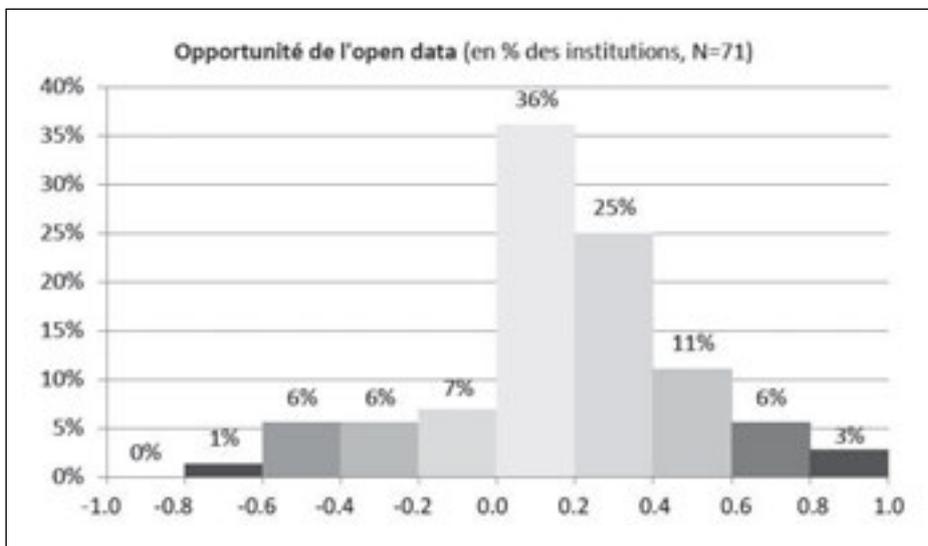


Illustration 1: Opportunité de l'open data (chances vs risques)

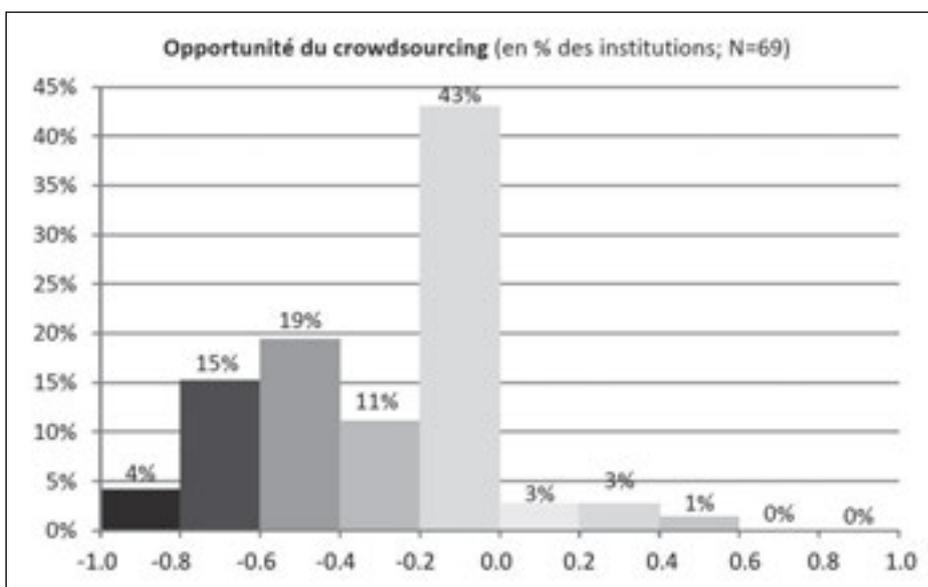


Illustration 2: Opportunité du crowdsourcing (chances vs risques)

du crowdsourcing et de la création collaborative de contenus comme étant aussi importante que l'open data/open content, il serait judicieux de mener quelques projets-pilotes dans ce domaine et d'entamer un processus d'apprentissage collectif parmi les institutions patrimoniales et les communautés en ligne concernées. Les conditions

pour ce faire sont une culture positive de l'erreur, la documentation et l'évaluation systématique des expériences faites ainsi que la création de plateformes d'échange appropriées, tels que conférences, ateliers ou offres de formation continue.

Contact: beat.ester mann@bfh.ch

Les applications basées sur les LOD en bibliothèque: un tour d'horizon

Nicolas Prongué et Jasmin Hügi,
spécialistes en information documentaire

Les Linked Open Data (LOD) sont devenus aujourd'hui un thème récurrent en bibliothèque, mais la plus-value qu'ils peuvent apporter est encore très difficile à percevoir. Quels résultats concrets ont déjà été atteints? Quelles sont, en bibliothèque, les applications LOD existantes à ce jour? Qu'en est-il de la situation en Suisse?

Tour d'horizon

De nombreux projets de mise à disposition de données bibliographiques en LOD sont en cours de réalisation ou ont abouti. En 2008 déjà, la Library of Congress (Etats-Unis) publiait le LCSH¹ en RDF. Mais le phénomène touche en particulier les bibliothèques européennes. Par exemple, le réseau suédois Libris² fut, la même année, le premier à publier l'intégralité de son catalogue en RDF.

D'autres institutions ont suivi. Plusieurs bibliothèques nationales (Allemagne, France, Grande-Bretagne, Hongrie, etc.) ainsi que des bibliothèques et réseaux universitaires, principalement allemands et français, ont publié certaines de leurs données.

Les données bibliographiques incluent également les fichiers d'autorités, qui sont particulièrement visés par les projets LOD. Ainsi, les fichiers d'autorités auteurs de plus de 30 bibliothèques ont été convertis en RDF et reliés (ils sont accessibles sur le site du VIAF³), tout comme les fichiers d'autorités matières RAMEAU en français, LCSH en anglais et GND en allemand⁴. Du fait qu'ils représentent des concepts du monde réel, comme des personnes ou des lieux, ces fichiers ont un potentiel de liaison bien plus grand que les données décrivant des documents: ils peuvent être liés à des données qui ne sont pas d'ordre bibliographique.

Néanmoins, la plupart de ces initiatives, après avoir converti et lié les don-

nées, se contentent de les mettre à disposition. Mais au-delà de la simple exposition des données, la possibilité de les exploiter, au moyen d'applications, est encore rarement considérée par les bibliothèques. Très peu d'applications LOD existent à ce jour. Et pourtant, c'est là que réside le réel potentiel de plus-value pour l'utilisateur.

Des données aux applications

Une application consiste en un logiciel traitant des données en vue de fournir un service à l'utilisateur. Dans cet article, nous avons donc considéré comme applications LOD en bibliothèque uniquement les services web destinés directement aux utilisateurs et apportant une plus-value basée sur les LOD. Ceci exclut notamment certaines interfaces dont l'utilisation nécessite des connaissances spécifiques, telles que le langage SPARQL par exemple.

Quelques bibliothèques disposent aujourd'hui de services web rendant accessibles les données bibliographiques en LOD. Néanmoins, la plus-value est bien souvent peu perceptible pour l'utilisateur direct, voire inexistante. De tels services sont entre autres proposés par le Sudoc⁵ (France), Libris (Suède), le réseau HBZ⁶ (Allemagne), la bibliothèque de l'Université de Winchester⁷ (Angleterre) ou encore les bibliothèques nationales allemande⁸, hongroise⁹ et britannique¹⁰.

Il existe néanmoins de réelles applications destinées directement à l'utilisateur lambda. En voici trois des plus abouties:

- Le Virtual International Authority File (VIAF). Maintenu par OCLC, ce portail fournit un accès centralisé aux notices d'autorités auteurs de différents réseaux et grandes bibliothèques, reliées entre elles.
- Le Centre Pompidou virtuel¹¹. Le nouveau site du Centre Pompidou à Paris, inauguré en octobre 2012, propose une interface moderne donnant accès à toutes les ressources de ses diverses institutions. On y trouve notamment

des notices de documents, des œuvres d'art numérisées, mais également toutes les ressources produites par le Centre lui-même: affiches, vidéos et enregistrements d'événements, dossiers pédagogiques, etc. Cependant, les données ne sont pas disponibles sous une licence ouverte, donc il ne s'agit pas d'une application strictement LOD.

- Data.bnf.fr. Cette application propose un accès unique à trois bases de données de la Bibliothèque nationale de France: le catalogue général, Gallica ainsi qu'Archives et manuscrits. Les données ont été FRBRisées pour être regroupées au sein de pages d'œuvres, d'auteurs et de sujets. De plus, des liens renvoient vers des notices d'autorités externes et des articles Wikipédia. Dans ses pages «ateliers», data.bnf.fr propose également de nouvelles manières d'exploiter les données, par exemple en les présentant sous forme de cartes et de frises chronologiques générées automatiquement.

La plus-value des applications

L'exploitation des LOD au moyen d'applications est encore très peu courante dans les bibliothèques. Pourtant, par rapport à un catalogue en ligne, cela

1 Library of Congress Subject Headings, fichier d'autorités matières en anglais, disponible sur <http://id.loc.gov/authorities/subjects.html>

2 <http://libris.kb.se/>

3 Virtual International Authority File: <http://viaf.org/>

4 ISAAC, Antoine. RAMEAU subject headings as SKOS linked data. Department of Computer Sciences, University of Amsterdam (en ligne). 2012. [Consulté le 14 octobre 2013]. Disponible à l'adresse: www.cs.vu.nl/STITCH/rameau/

5 www.sudoc.abes.fr/

6 Hochschulbibliothekszentrum (Rhénanie-du-Nord-Westphalie): www.lobid.org/

7 <http://prism.talis.com/winchester-ac/>

8 <https://portal.dnb.de/>

9 <http://nektar2.oszk.hu/>

10 <http://bnb.data.bl.uk/>

11 www.centrepompidou.fr/

peut apporter plusieurs avantages à l'utilisateur final, parmi lesquels:

La recherche fédérée

Dans data.bnf.fr par exemple, les diverses données, initialement encodées en MARC, Dublin Core et EAD, ont été converties dans un même format du modèle RDF¹². Grâce à cela, une seule requête permet d'interroger trois réservoirs de manière optimisée. La page d'un auteur fournit ainsi un accès unique à toutes les ressources liées à cet auteur, quels qu'en soient leurs réservoirs de base. Cela permet également, de pair avec la FRBRisation, d'éviter les doublons.

L'enrichissement des catalogues

Le principe de base des LOD réside dans la liaison des données. Les applications peuvent donc proposer des liens vers d'autres ressources, internes ou externes, pour favoriser la découverte et la sérendipité lors de la recherche. Le Centre Pompidou propose par exemple d'augmenter l'expérience de l'utilisateur par de nombreux liens internes, notamment vers des vidéos et des dossiers pédagogiques créés par ses institutions.

Les recherches multilingues

Les LOD sont prometteurs dans ce domaine. Par exemple, lors d'une requête dans The European Library¹³, le système propose des sujets connexes au terme recherché, en français, allemand et anglais, grâce à l'alignement des fichiers d'autorités matières RAMEAU, LCSH et GND en format RDF.

La visibilité des catalogues sur le Web

Un format compatible avec les standards du Web permet aux données d'apparaître dans les résultats des moteurs de recherche.

Et en Suisse?

Pour connaître l'état de la situation en Suisse, nous avons contacté les principaux acteurs susceptibles d'être inté-



Figure 1: Situation des LOD en Suisse

ressés par le développement des LOD, soit les grandes bibliothèques et les réseaux. Selon les résultats obtenus (figure 1), un certain retard peut être constaté. En effet, seules quatre institutions sur les onze ayant répondu ont un projet LOD en cours de préparation: RERO, la Bibliothèque nationale, Swissbib et la bibliothèque du CERN. Cependant, rien n'a encore abouti à un résultat concret. RERO, par exemple, a déjà préparé le terrain pour le passage aux LOD, notamment en participant au VIAF, en intégrant RAMEAU, en encourageant la création de notices d'autorités auteurs au sein du réseau ainsi qu'en attribuant des URI à ses données.

Ce retard peut être dû à différents facteurs. D'une part, les institutions suisses sont très petites en comparaison à des géants tels que la Bibliothèque nationale de France, et cela représente un frein. D'autre part, les institutions suisses n'ont pas encore toutes déterminé à qui revient la responsabilité du développement des LOD. Est-ce la tâche des réseaux ou de chaque bibliothèque?

La Bibliothèque nationale doit-elle montrer l'exemple? Faut-il organiser un projet au niveau national? Cette problématique est essentiellement d'ordre stratégique.

Conclusion

Les LOD reposent sur une technologie très jeune et encore en pleine phase de recherche. Très peu d'applications ont déjà été créées par les bibliothèques directement pour leurs utilisateurs. La plus-value des LOD est encore difficile à percevoir: aucune «killer application», révolutionnant la recherche et l'exploitation des données, n'a pour l'instant été inventée. Les bibliothèques suisses, en retard dans ce domaine, devraient se fixer des objectifs et prendre position par rapport aux LOD. Si elles ne souhaitent ou ne peuvent pas encore investir dans un tel projet, elles peuvent du moins déjà soigner et enrichir leurs fichiers d'autorités, qui sont la clé de voûte des LOD en bibliothèque.

Contact: nicolas.prongue@etu.hesge.ch
jasmin.hugi@hesge.ch

ABSTRACT

Applikationen, die mit LOD arbeiten, in Bibliotheken: eine Übersicht

LOD beruht auf einer sehr jungen Technologie, die sich noch im Forschungsstadium befindet. Es gibt erst sehr wenige Anwendungssoftware, die von Bibliotheken für ihre Nutzer entwickelt wurde. Der Wert der LOD kann noch nicht beziffert werden: Es wurde bis jetzt keine eigentliche «Killer-Applikation» erfunden, die die Suche und die Verwaltung von Daten revolutioniert hätte. Die Schweizer Bibliotheken, die diesbezüglich der Entwicklung etwas hinterher hinken, müssen in Bezug auf LOD Stellung beziehen und sich Ziele setzen. Wenn die Bibliotheken sich noch nicht in solche Projekte einbringen können oder wollen, so können sie zumindest ihre Autoritätsdaten, die den Zugangsschlüssel zu LOD in Bibliotheken darstellen, anreichern und pflegen. (Uebersetzung: sh)

12 Bibliothèque Nationale de France. Web sémantique et modèle de données. data.bnf.fr (en ligne). [Consulté le 14 octobre 2013]. Disponible à l'adresse: <http://data.bnf.fr/semanticweb>

13 www.theuropeanlibrary.org/tel4/

Alle Bibliotheken können die Zukunft der Katalogisierung gestalten

Eliane Blumer, Karsten Schuldt*

Linked Open Data ist Bestandteil einer grösseren Bewegung hin zur datengetriebenen Bibliotheksarbeit. Alle Bibliotheken haben die Chance, bei der Gestaltung der zukünftigen Katalogisierungsregeln und -formate aktiv mitzuwirken. Dazu ist es notwendig, sich mit den Möglichkeiten und aktuellen Diskussionen in diesem Feld auseinanderzusetzen und sich in diese einzubringen.

Einige wichtige Stichworte sind dabei FRBR, RDA, Bibframe, Kataloganreicherung und Metadatenmanagement. Die Functional Requirements for Bibliographic Records (FRBR) stellen ein Framework für die bibliografische Beschreibung von Bibliotheksmedien aus Sicht der Nutzerinnen und Nutzer dar, welches als Grundlage für die Formulierung zukünftiger Regelwerke dienen soll. (IFLA Study Group on the Functional Requirements for Bibliographic Records, 2009). FRBR unterteilt die bibliografischen Daten eines Mediums in vier Entitäten: (Work, Expression, Manifestation und Item). Die Entität Item stellt dabei das konkrete, in einer Bibliothek vorhandene Medium dar. Von dieser Ebene wird der Inhalt aufsteigend abstrahiert. Wichtig sind die Beziehungen zwischen den Ebenen. Resource Description and Access (RDA) stellt das erste auf FRBR aufbauende Regelwerk dar. RDA betont die Verfügbarmachung eines Mediums als Aufgabe des Katalogs. Es wird seit 2013 in der Library of Congress und ab 2014 in der Deutschen Nationalbibliothek genutzt. Der Entscheid der Schweizerischen Nationalbibliothek, RDA einzuführen, ist ebenfalls prinzipiell gefallen (Schweizerische Nationalbibliothek, 2013). Bibframe stellt den Versuch eines Austauschformats dar, welches die Bibliothekskataloge an das Semantic Web anschliesst. Es soll MARC und seine Dialekte ablösen. (Library of Congress, 2012).

Gleichzeitig gibt es in Bibliotheken die Tendenz, Bibliotheksdaten möglichst automatisch mit anderen Daten anzureichern. Die Anreicherung mit

Coverabbildungen ist heute Standard, genannt werden in der internationalen Diskussion weitergehende Anreicherungen, zum Beispiel Reviews, Inter-

Gleichzeitig gibt es in Bibliotheken die Tendenz, Bibliotheksdaten möglichst automatisch mit anderen Daten anzureichern.

pretationen, oder Querverbindungen zu anderen Medien. Begründet werden die Anreicherungen regelmässig damit, dass die Nutzerinnen und Nutzer diese wünschen, obgleich das nicht immer nachgewiesen wird.

Grundsätzlich kann gesagt werden, dass Fragestellungen im Bezug auf bibliografische Arbeit und Bibliothekskataloge international einer intensiven Diskussion unterliegen. Sicher ist, dass die Katalogarbeit in zehn Jahren radikal anders aussehen wird. Dabei beziehen einige Einrichtungen oder Einzelpersonen klare Positionen und bestimmen so die zukünftige Gestaltung der Katalogarbeit mit.

Trends und Szenarien

Wer sich in Diskussionen um die Zukunft der Katalogarbeit einbringen will, sollte den heutigen Zustand kennen und gleichzeitig Vorstellungen von möglichen und anstrengenswerten Zukünften haben. Dies ermöglicht, Positionen dazu zu beziehen, was möglich sein soll, was notwendig ist und was verhindert werden muss. In diesem Abschnitt wollen wir als Diskussionsgrundlage drei erkennbare Trends als Zukunftsszenarien darstellen.

Kollaborative Katalogisierung

Die aktuell diskutierten Regelwerke und Austauschformate zeigen eine klare Tendenz: Es werden Netzwerke von Wissen generiert, in denen einzelne Medien Endpunkte darstellen. Diese Netzwerke entstehen nicht, indem bibliografische Daten jedes einzelnen Bibliotheksmediums aufgenommen und anschliessend untereinander verbun-

den werden. Vielmehr implizieren die Diskussionen, dass bei Integration in einen Bibliotheksbestand die einzelnen Medien als Item in ein Wissensnetzwerk eingebunden werden.

Die Pflege dieses Wissensnetzwerks ist als Ergebnis kollaborativer Arbeit möglich: Alle Bibliotheken übernehmen gemeinsam die Weiterentwicklung des Netzwerks, ergänzen und updaten die notwendigen Entitäten und vertrauen gleichzeitig darauf, dass die Kolleginnen und Kollegen in den

Die aktuell diskutierten Regelwerke und Austauschformate zeigen eine klare Tendenz: Es werden Netzwerke von Wissen generiert, in denen einzelne Medien Endpunkte darstellen.

anderen Bibliotheken, im Idealfall auf internationaler Ebene, das Gleiche mit dem gleichen qualitätsbewussten Impetus tun.

Szenario

2021, der Bibliothekar L. der juristischen Bibliothek im Centovalli hat soeben die 5. Auflage eines Buches erhalten und stellt während seiner Suche innerhalb des Wissensnetzwerks fest, dass er der erste Besitzer der besagten Auflage ist. Mit einem Klick fügt er die Auflage unter die Entität «Manifestation» zur betreffenden Expression hinzu. In einem weiteren Schritt fügt er das Buch – als Entität «Item» – hinzu. Früher hätte er zeitraubend den bereits katalogisierten Datensatz des Buches als Fremdtiteldaten übernommen, überprüft und seinen Hausregeln angepasst.

* Eliane Blumer ist Lehr- und Forschungsassistentin im Bereich «Information documentaire» an der Haute Ecole de Gestion, Genève.
Dr. Karsten Schuldt (Chur/Berlin) ist Wissenschaftlicher Mitarbeiter am Schweizerischen Institut für Informationswissenschaft, HTW Chur.

Zentrale Katalogisierung

Die aktuellen Diskussionen beinhalten auch die Möglichkeit, dass ein Grossteil der Katalogarbeit nicht mehr in den einzelnen Bibliotheken, sondern zentral stattfindet, bspw. bei den Verbundzentralen, Nationalbibliotheken oder zentralen Katalogisierungsstellen. Wenn konkrete Bibliotheksmedien als Item an ein bestehendes Wissensnetzwerk angehängt werden, kann es sich als sinnvoll herausstellen, wenn dieses Wissensnetzwerk von einigen Expertinnen und Experten gepflegt wird. Die Bibliotheken hätten weniger Personalzeit in die Katalogisierung zu investieren und könnten sich anderen Aufgaben widmen. Gleichzeitig würden sich die Expertinnen und Experten, welche

Die Bibliotheken hätten weniger Personalzeit in die Katalogisierung zu investieren und könnten sich anderen Aufgaben widmen.

das bibliografische Wissensnetzwerk pflegen, intensiver mit den Konsequenzen von Entscheidungen in Regelwerken auseinandersetzen und sich aktiv in Diskussionen über diese einbringen. Für solche Systeme gibt es heute schon in öffentlichen Bibliothekssystemen Vorbilder.

Szenario

2018, die Mitarbeiterin M. der Verbundzentrale Gesamtschweiz verbringt ihren Morgen mit der Metadatenerstellung auf allen Entitätsebenen eines Pakets von Büchern zum rätoromanischen Schulunterricht und schickt die Bücher anschliessend an die Bibliotheken, welche die Bestellung in Auftrag gegeben hatten. Diese können die Medien direkt zur Ausleihe verwenden.

Da sie von einer Kollegin des Verbunds Ostdeutschland vernommen hat, dass das Sorbische sich ähnlich entwickelt, es aber im Wissensnetzwerk noch keine Verbindung der zwei Kleinsprachen gibt, beschliesst sie, dies auf der nächsten bibliografischen Konferenz zu thematisieren.

Metadata Librarian

Die erkennbare Tendenz, Katalogdaten innerhalb der Bibliothekssysteme mit

weiteren Daten anzureichern, wird sich verstärken. Die Frage ist nur, welche Anreicherungen sinnvoll sind. Dies hängt von den verfügbaren Daten, den Möglichkeiten der zukünftigen Bibliothekssoftware und den Einschätzungen der Bibliotheken, was für die Sammlungen sowie die Nutzerinnen und Nutzer sinnvoll ist, ab. Grundsätzlich wird mit einer steigenden Anzahl der Anreicherungen die Aufgabe an die Bibliotheken übergehen, aus diesen Möglichkeiten zu wählen. Dies könnte die heutige Katalogisierung ersetzen.

Szenario

2035, der Musikbibliothekar D. des Instituts für Populärmusik in Versoix erhält eine Sammlung von Stücken des Schweizerischen Kammermusikorchesters. Für die gespielten Stücke wählt er innerhalb internationaler und regionaler Anreicherungen, z.B. Lexikonartikel zu Musikstilen, Webportale von Musikfans oder musikwissenschaftliche Artikel, aus. Dadurch werden die Stücke in einen erweiterten Kontext gesetzt, und es wird so ermöglicht, weit über das einzelne Musikstück hinweg Informationen direkt im Katalog zu finden. Der alte Katalog hingegen beantwortete vor allem die Frage, ob ein Medium vorhanden war.

Fazit

Alle Bibliotheken können sich aktiv in die Gestaltung der Zukunft des Katalogs und der bibliothekarischen Regelwerke einbringen. Bislang tun dies aber nur wenige. So brachte sich die Deutsche Nationalbibliothek engagiert in die Diskussion um FRBR und RDA ein und hinterliess in diesen Spuren europäischer Katalogtheorie und -praxis. Die Schweizerische und Österrei-

che Nationalbibliothek arbeiten nun mit der Deutschen Nationalbibliothek bei der Umsetzung von RDA zusammen.

Die Mailinglisten für die Diskussionen um Bibframe (bibframe@listserv.loc.gov), FRBR (frbr@infoserv.ist.fr) oder die RAK-Liste (rak-list@lists.dnb.de) zeigen, dass sich Bibliothekarinnen und Bibliothekare aus unterschiedlichsten Einrichtungen tatkräftig daran beteiligen können, die bibliothekarischen Regelwerke, Frameworks und Werkzeuge zu gestalten. Dazu bedarf es eines Wissens darüber, was möglich wäre und was sein sollte. Wir haben drei Szenarien als Input für Diskussionen entworfen. Andere Meinungen sind möglich und erwünscht.

Kontakt: eliane.blumer@hesge.ch
karsten.schuldt@htwchur.ch

Literatur

- IFLA Study Group on the Functional Requirements for Bibliographic Records/Functional Requirements for Bibliographic Records: Final Report. [ohne Ort] : International Federation of Library Associations and Institutions, 2009
- Library of Congress/Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services. Washington, D.C.: Library of Congress, 2012
- Schweizerische Nationalbibliothek / Schweizerische Nationalbibliothek 99. Jahresbericht 2012. Bern: Schweizerische Nationalbibliothek, 2013

ABSTRACT

Bibliothèques et catalogage ou comment maîtriser le futur

Dans un proche avenir, le travail de catalogage changera dans les bibliothèques, notamment en raison de la place qu'occuperont les Linked Open Data. Toutes les bibliothèques peuvent s'insérer dans ce processus de changement et y participer activement. Les auteurs décrivent dans cet article trois tendances qui leur semblent particulièrement pertinentes à ce stade, à savoir le catalogage collaboratif, le catalogage central et le *metadata librarian*, et élaborent pour chacune d'elle un scénario envisageable pour le futur travail des bibliothèques. Ces scénarios devraient surtout permettre de susciter le débat sur les futurs possibles et souhaitables pour les bibliothèques. (*traduction: sg*)

II. Projekte und Initiativen aus Bibliotheken und Archiven / *Projets et initiatives dans le domaine archive et bibliothèque*

Die Bibliographic Framework Initiative (BIBFRAME) der Library of Congress – Offene, verknüpfte Datensätze

Bruno Wenk, HTW Chur

Die Bibliographic Framework Transition Initiative stellt ein zukunftsorientiertes Modell für den Übergang von MARC zu verknüpften offenen bibliografischen Daten vor. Das Modell orientiert sich sowohl an der Praxis als auch an bereits vorliegenden Modellen (FRBR) und neuen Katalogisierungsregeln (RDA).

Grenzen neuer Dienstleistungen für Bibliotheksnutzer/innen

Bibliotheken bemühen sich bei der Entwicklung neuer Angebote vermehrt um Zusammenarbeit, um ein grösseres Publikum zu erreichen und die Kosten für die Dienstleistungen teilen zu können. Ein Beispiel dafür ist der Velobuchkurier der fünf am «Bibliotheksplatz Chur»¹ beteiligten Bibliotheken: Die Kunden können sich Medien der teilnehmenden Bibliotheken kostenlos in die Lieferbibliothek ihrer Wahl brin-

gen lassen und Medien an der nächstgelegenen Bibliothek zurückgeben. Die innovative Dienstleistung hat allerdings einen Haken: Interessierte Nutzer/innen müssen in zwei verschiedenen Onlinekatalogen suchen, um den Gesamtbestand der beteiligten Bibliotheken abzufragen – wer ist heute noch bereit, diese Mühe auf sich zu nehmen? Der Aufwand für die Realisierung einer Webanwendung zur gleichzeitigen Suche in beiden Katalogen ist für die beteiligten Bibliotheken aber aus technischen und finanziellen Gründen nicht vertretbar.

Im Rahmen des Innovations- und Kooperationsprojekts der Schweizer Hochschulen «E-lib.ch: Elektronische Bibliothek Schweiz» wurde der Aufwand geleistet. Mit der Eingabe von Stichwörtern im Suchfeld von Swissbib² werden die Bestände von 900 Bibliotheken, Mediotheken und Archiven aus allen Landesteilen der Schweiz ab-

gefragt, darunter auch diejenigen der Churer Bibliotheken. Aber Swissbib wäre auch keine Lösung. Weil die Suche nur auf Bibliotheksverbünde und nicht auf einzelne Bibliotheken eingeschränkt werden kann, müssen zur Abfrage des Gesamtbestands des «Bibliotheksplatzes Chur» trotzdem zwei getrennte Suchen durchgeführt werden.

Verteilte geschlossene Datenbanken sind keine zukunftsorientierte Form der Datenhaltung für Gedächtnisinstitutionen. Um Bibliotheksnutzerinnen und -nutzern innovative Dienstleistungen anbieten zu können, ist Open Data eine (notwendige) Grundlage.

Open Data

Open Data ist eine Bewegung mit dem Ziel, Daten aus geschlossenen Datenbanken zu «befreien» und im Web öffentlich zugänglich und nutzbar zu machen. Wer eine auf offenen Daten beruhende Dienstleistung anbieten möchte, kopiert die Daten und programmiert eine passende Webapplikation. Ein Beispiel für offene geografische Daten (Orte, Strassen, Gebäude usw.) bietet OpenStreetMap³. Basierend auf den Daten von OpenStreetMap³ stellt eine an der HTW Chur realisierte Webanwendung die (auf OpenStreetMap erfassten) Schweizer Bibliotheken (und Museen) auf einer Karte dar (Bild 1)⁴.

Die Veröffentlichung von Daten bewegt momentan vor allem die Behörden

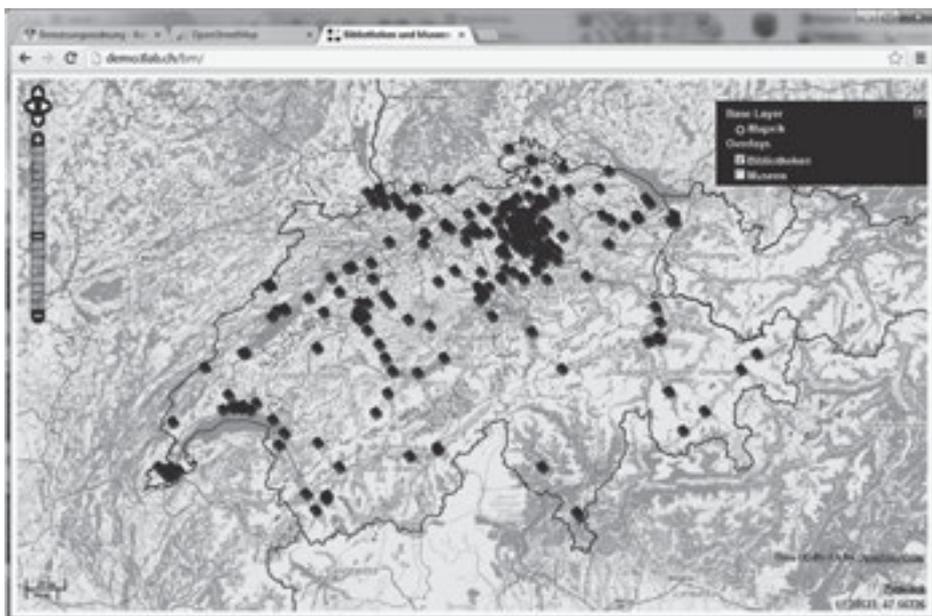


Bild 1: Karte mit den auf OpenStreetMap erfassten Schweizer Bibliotheken

- 1 www.bibliotheksplatz-chur.ch
- 2 www.swissbib.ch/
- 3 www.openstreetmap.org
- 4 <http://demo.tlab.ch/bm/>

1. The curious case of Benjamin Button das Hörbuch zum Sprachen lernen Verfasser / Beitragende: F. Scott Fitzgerald	2009	Ungelirnte Originalfassung	Medienkategorie: Englisch	5 Bibliotheken
2. The Curious Case of Benjamin Button das Hörbuch zum Sprachen lernen Verfasser / Beitragende: F. Scott Fitzgerald; David Hillerman (Sprecher)	2009	JUF Ton	Deutsch	2 Bibliotheken

Bild 2: Wegen der unterschiedlichen Belegung der MARC-Felder ist für ein Programm nicht erkennbar, dass sich die beiden dargestellten Datensätze auf dasselbe Hörbuch beziehen.

(Open Government Data). Beispielsweise hat am 16. September 2013 die Bundesverwaltung ihr Pilotportal für offene Behördendaten der Schweiz⁵ eröffnet. Ein erster Schritt von Bibliotheken in Richtung Open Data wäre die Öffnung der OPAC und die Publikation ihrer bibliografischen Daten im Web. Vorreiterin diesbezüglich war die Bibliothek des Kernforschungszentrums CERN in Genf, welche im Dezember 2009 die Datensätze des CERN Document Server im Format MARCXML im Web zum Download zur Verfügung stellte⁶.

Die Open-Data-Bewegung legt das Format der offenen Daten nicht fest. Das Format MARCXML wäre demnach für die Veröffentlichung bibliografischer Daten durchaus geeignet. Aber das kryptische Feldformat ist nicht mehr zeitgemäss und die Belegung der Felder von Bibliothek zu Bibliothek zu unterschiedlich (vgl. Library of Congress, 2011). Die Programmierung von Metasuchfunktionen ist deshalb sehr aufwendig (Bild 2).

Linked Open Data

Das Konzept von Linked Open Data ist strikter als dasjenige von Open Data. Es legt nicht nur fest, dass die Daten im Web veröffentlicht und durch einen Unified Resource Locator⁷ (URL) eindeutig adressierbar sein müssen; das Konzept von Open Data legt auch fest,

dass die Datensätze in einer standardisierten Form (RDF⁸) maschinenlesbar formuliert sein und Verknüpfungen (Links) mit weiteren, verwandten Datensätzen enthalten müssen (Berners-Lee, 2006).

Auf bibliografische Daten übertragen ermöglichte das Konzept von Linked Open Data den Nutzerinnen und Nutzern, Werke, Autorinnen und Autoren, Institutionen und viele weitere Fakten zu entdecken. Im Linked-Open-Datensatz zu «The Great Gatsby» wäre anstelle des Namens des Autors (die Buchstabenfolge «Fitzgerald, F. Scott») eine Verknüpfung zum Datensatz mit Angaben zum Autor Francis Scott Fitzgerald zu finden und würden beispielsweise auch Verknüpfungen zu Datensätzen mit Angaben zu den Übersetzungen und Verfilmungen des Romans enthalten sein. Und weil diese Datensätze maschinenlesbar wären, würde eine entsprechende Webanwendung einer Person, die sich für den grossen Gatsby interessiert, durch automatisches Verfolgen der verknüpften Datensätze auch anzeigen können, dass Francis Scott Fitzgerald Ernest Hemingway kannte.

Einen Weg von offenen bibliografischen Daten im Format MARCXML zu verknüpften offenen bibliografischen Daten bietet die Bibliographic Framework Transition Initiative (BIBFRAME) der Library of Congress.

BIBFRAME

BIBFRAME⁹ ist eine Initiative der Library of Congress. In Zusammenarbeit mit der Firma Zepheira¹⁰, Library and Archives Canada, der British Library, der Deutschen Nationalbibliothek und anderen Nationalbibliotheken, zahlreichen MARC-Nutzergruppen und weiteren Communities wird ein Modell ent-

wickelt, das den Grundstein für die Zukunft verknüpfter offener bibliografischer Daten legt. Das praxisorientierte Modell berücksichtigt die Anforderungen an detaillierte bibliografische Beschreibungen, an Beschreibungen anderer kultureller Objekte und an weniger detaillierte Beschreibungen von Inhalten gleichermaßen. Die Initiative befasst sich aber nicht nur mit einem zukunftsorientierten Ersatz für das Machine Readable Cataloging (MARC) Format, sondern untersucht auch Aspekte der Erfassung und des Austauschs bibliografischer Daten (z.B. Katalogisierungsregeln und Datenaustauschprotokolle).

Das Modell von BIBFRAME lehnt sich an dasjenige der Functional Requirements for Bibliographic Records (FRBR) an. Es unterscheidet klar zwischen einem Werk (work) und dessen physischen/digitalen Erscheinungsformen (instances), umfasst aber nur diese beiden Ebenen (anstelle der vier Ebenen Werk/Expression/Manifestation/Exemplar von FRBR). Werke und Instanzen können Beziehungen zueinander und untereinander sowie Beziehungen zu Autoritäten (authorities) und Notizen (annotations) aufweisen (Bild 3).

Zur Beschreibung der Beziehungen stellt BIBFRAME ein Vokabular zur Verfügung, das erweiterbar ist. Beispielsweise beschreiben die Ausdrücke «hasInstance» bzw. «instanceOf» die Beziehung zwischen einem Werk (z.B. ein Roman) und dessen Erscheinungsformen (z.B. Taschenbuch). Der Ausdruck «creator» beschreibt die Beziehung zwischen dem Roman und dessen Autor/in. Und mit dem Ausdruck «providerName» wird die Beziehung zwischen einer Instanz (z.B. das Taschenbuch) und dem publizierenden Verlag beschrieben.

Auf der Website von BIBFRAME¹¹ steht ein Onlinedienst zur Verfügung, um MARCXML-Datensätze automatisch in das Format BIBFRAME zu konvertieren. Bei der Konversion werden die Verknüpfungen mit Autoritätsdatensätzen ebenfalls automatisch eingefügt. Damit können auch Verbünde kleinerer (öffentlicher) Bibliotheken

5 <http://opendata.admin.ch>

6 <http://oldlibrary.web.cern.ch/oldlibrary/Library/bookdata.html>

7 Webadresse, z.B. <http://www.worldcat.org/oclc/727645936>

8 Resource Description Framework

9 www.loc.gov/bibframe/

10 <http://zepheira.com>

11 <http://bibframe.org>

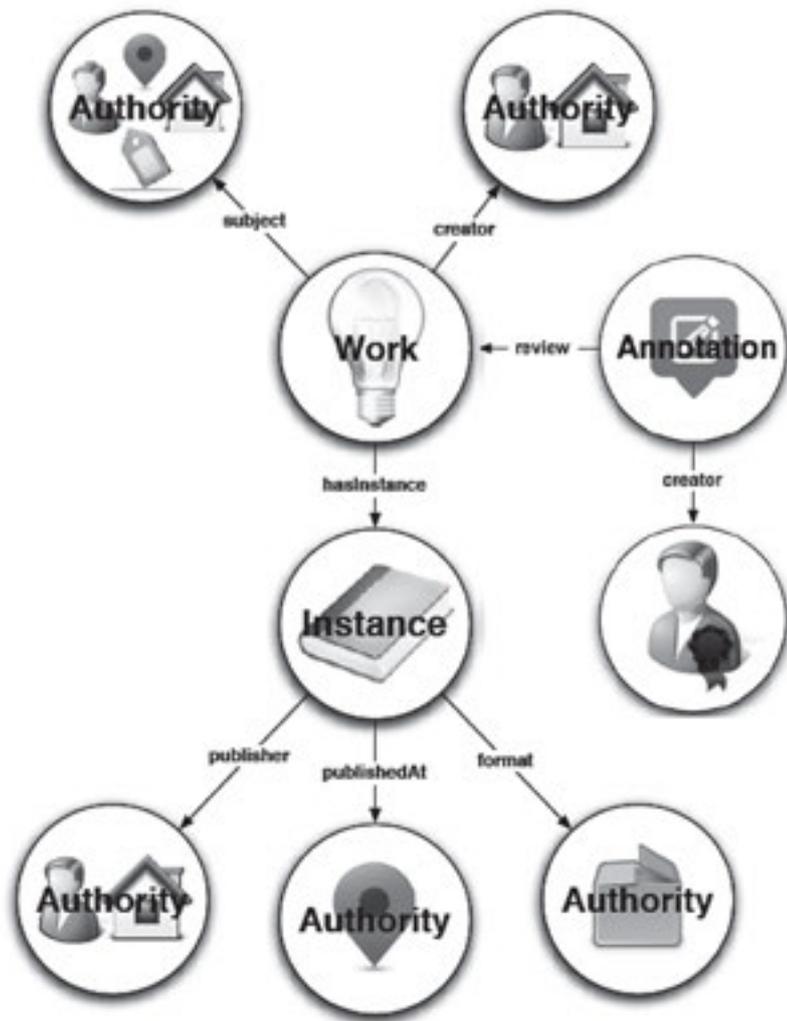


Bild 3: BIBFRAME-Modell (Quelle: Library of Congress, 2012; modifiziert)

Experimente mit Linked Open Library Data durchführen.

Die automatische Konversion löst allerdings das in Bild 2 illustrierte Problem der unterschiedlichen Belegung von MARC-Datenfeldern nicht. Die Inhalte der generierten BIBFRAME-Datensätze müssen auf die Einhaltung allgemein gültiger Katalogisierungsregeln

geprüft und nötigenfalls überarbeitet werden. Obwohl BIBFRAME keine bestimmten Katalogisierungsregeln voraussetzt, orientiert sich dessen Vokabular an RDA¹² (Resource Description and Access). Bibliotheken, die sich bereits mit der Einführung von RDA auseinandergesetzt haben, können beim Übergang von MARC zu BIBFRAME darauf aufbauen.

ABSTRACT

Des données ouvertes et reliées au lieu de banques de données fermées

L'initiative Bibliographic Framework Transition présente un modèle d'avenir pour la transition de MARC à des données biographiques ouvertes et connectées. Le modèle s'oriente aussi bien vers la pratique que vers des modèles existants (FRBR) et de nouvelles règles de catalogage (RDA). Le travail sur BIBFRAME est encore en cours, raison pour laquelle seuls des projets-pilotes réalisés dans un cadre restreint sont judicieux pour l'instant. Cela étant, se pencher sur BIBFRAME et participer activement à la discussion en cours concernant le modèle BIBFRAME est utile et nécessaire pour le développement des futures offres d'informations des bibliothèques. *(traduction: sg)*

Um schliesslich den Nutzerinnen und Nutzern den im Zusammenhang mit Linked Open Library Data versprochenen Mehrwert zu bieten, müssen die konvertierten BIBFRAME-Datensätze noch mit anderen offenen Datensätzen verknüpft werden, beispielsweise mit Datensätzen der Linked Movie¹³ Database (LinkedMDB) und Datensätzen von Wikidata¹⁴, der freien Wissensbasis mit aktuell mehr als 13 Millionen Datensätzen.

Fazit

Die Bibliographic Framework Transition Initiative stellt ein zukunftsorientiertes Modell für den Übergang von MARC zu verknüpften offenen bibliografischen Daten vor. Das Modell orientiert sich sowohl an der Praxis als auch an bereits vorliegenden Modellen (FRBR) und neuen Katalogisierungsregeln (RDA). Die Arbeit an BIBFRAME ist noch im Gang, weshalb vorläufig nur Pilotprojekte in kleinerem Rahmen sinnvoll sind. Aber die Beschäftigung mit BIBFRAME und die aktive Beteiligung an der laufenden Diskussion über das BIBFRAME-Modell sind für die Entwicklung zukünftiger Informationsangebote von Bibliotheken auf jeden Fall nötig.

Kontakt: bruno.wenk@htwchur.ch

Quellen

- Berners-Lee, T., «Linked Data», 2006. (www.w3.org/DesignIssues/LinkedData.html)
- Library of Congress, «A Bibliographic Framework for the Digital Age», 2011. (www.loc.gov/bibframe/news/framework-103111.html)
- Library of Congress, «Web of Data: Linked Data Model and Supporting Services», 2012. (www.loc.gov/bibframe/pdf/marclid-report-11-21-2012.pdf)

¹² www.rdatoolkit.org

¹³ <http://datahub.io/de/dataset/linkedmdb>

¹⁴ www.wikidata.org

Vom Projekt zum Produkt: Der Linked Data Service der Deutschen Nationalbibliothek

Julia Hauser,
Deutsche Nationalbibliothek

Das Umfeld, in dem Informationsspezialisten wie Bibliothekare, Archivare und Dokumentare arbeiten, hat sich mit der Verbreitung des World Wide Webs und den zugrunde liegenden Informationstechnologien grundlegend verändert. Damit ergaben sich neue Möglichkeiten der Datenbereitstellung, der Datennutzung und interdisziplinärer institutionenübergreifender Kooperationsmöglichkeiten.

Motivation und Hintergrund

Das Umfeld, in dem Informationsspezialisten wie Bibliothekare, Archivare und Dokumentare arbeiten, hat sich mit der Verbreitung des World Wide Web und den zugrunde liegenden Informationstechnologien grundlegend verändert. Damit ergaben sich neue Möglichkeiten der Datenbereitstellung, der Datennutzung und interdisziplinärer institutionenübergreifender Kooperationsmöglichkeiten. Die vollständige Überführung existierender Arbeitsprozesse und lang etablierter Strukturen hin zu einer modernen webbasierten Informationsinfrastruktur ist ein langer Weg – erste Schritte dahin sind bereits getan; weiteres Potenzial ist vorhanden und entsteht durch den fortschreitenden technologischen Entwicklungsprozess immer wieder neu.

Während Bibliotheksdaten viele Jahre lang in geschlossenen Katalogsystemen isoliert waren, sind sie heute dabei, zu einem stabilisierenden Element des World Wide Web zu werden. Kontrollierte Vokabulare, standardisierte Metadatenformate und ein persistenter Nachweis der Ressourcen sind Komponenten, die die Arbeit von Informationsspezialisten seit je begleiten. Bibliotheksdaten können nun ei-

nen Beitrag zur Verlässlichkeit, Beständigkeit und Datenqualitätssicherung des Semantic Web leisten. Zur Bedeutung etwa der Nationalbibliografie im Zeitalter des World Wide Web, aber auch des WWW für die Nationalbibliografie, schreiben Kett et. al.: «Es sind Datennetze wie die Linked Open Data Cloud, in die Nationalbibliografien integriert werden müssen, um den Erwartungen des World Wide Web gerecht und auch in Zukunft noch als relevant wahrgenommen zu werden.»¹

Zukünftig sind weiterführende Schritte notwendig, um das volle Potenzial des Linked-Data-Modells auszunutzen. Damit es für externe Dienstleister und Portale möglich wird, die bibliografischen Daten in ihre Angebote zu integrieren, sollen beispielsweise vorliegende textuelle Informationen (Literele) in Referenzen auf Entitäten umgesetzt werden, also mit Personen, Körperschaften, Ereignissen und Orten sowie anderen Objekten verknüpft werden. Dies ist ein weiterer Schritt in die Richtung, die Nationalbibliografie als einen Teilgraphen in der Linked Data Cloud zu verankern².

Viele Bibliotheken haben die darin liegende Chance erkannt, sich bibliotheksfremden Sparten und Kooperationspartnern aus der Forschung zu öffnen und die bibliografischen Daten für weitere Anwendungsfälle attraktiver zu vermarkten. Dies schafft eine bessere Nachnutzbarkeit in anderen Kultur- und Gedächtniseinrichtungen wie Museen und Archiven und ermöglicht eine bessere (technische) Zitierbarkeit im Bereich der Forschung und Wissenschaft – einem Bereich, der zunehmend webbasiert arbeitet und auf verlässliche und persistente Datenquellen im Web angewiesen ist.

Je stärker Bibliotheksdaten mit anderen Datenanbietern und Institutionen vernetzt sind, je mehr wird ihre Bedeutung in der heutigen Informationsgesellschaft gestärkt: der Wert eines Knotens im Geflecht des Semantic Web bestimmt sich danach, wie viele Fäden zu ihm hinführen.

Erreichte Meilensteine im Linked Data Service der Deutschen Nationalbibliothek

Der Linked Data Service begann als kleines internes Forschungs- und Entwicklungsprojekt im Jahr 2009 mit ersten Überlegungen dazu, wie die Daten der Deutschen Nationalbibliothek im Kontext des Semantic Web bereitgestellt und nutzbar gemacht werden könnten. Funktionalitäten und inhaltliche Modellierungsaspekte wurden in mehreren Iterationsschleifen in Folgeprojekten bearbeitet. Hierbei halfen sowohl DNB-interne Evaluationsrunden sowie Feedback aus der Linked Data Community. Auf diesem Wege wuchs der Linked Data Service Schritt für Schritt aus dem Projektstatus hinaus und wird nun zum Produkt mit verbindlicher Releaseplanung. Hiermit steht RDF (Resource Description Framework)³ fortan gleichberechtigt neben weiteren Exportformaten, die von der Deutschen Nationalbibliothek angeboten werden⁴.

Porträt

Julia Hauser ist Diplom-Bibliothekarin in der IT-Abteilung der Deutschen Nationalbibliothek (DNB). Hier arbeitet sie seit 2009 an Projekten zur Repräsentation der DNB-Daten im Kontext des Semantic Web mit. Julia Hauser verantwortet den Linked Data Service der DNB und koordiniert den DNB-Beitrag zur BIBFRAME-Initiative der Library of Congress.



1 Kett, Jürgen; Manecke, Mathias; Beyer, Sarah: Die Nationalbibliografie im Zeitalter des Internets. In: ZfBB 59(2012), 2, S. 70

2 vgl. ebd., S. 71 ff.

3 www.w3.org/TR/rdf-primer/

4 www.dnb.de/datendienst

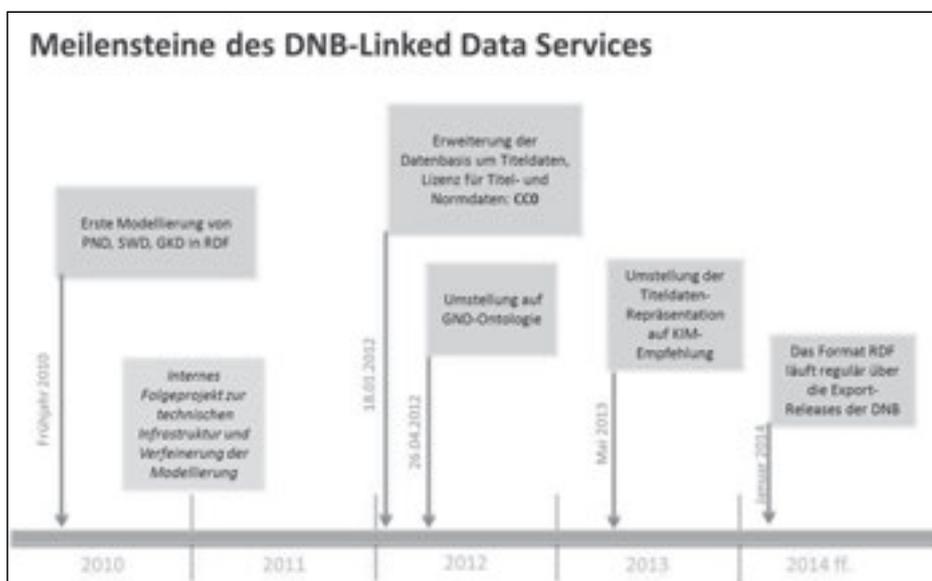


Abbildung: Meilensteine des DNB Linked Data Services

Im Folgenden werden die einzelnen Evolutionsschritte des Linked Data Services erläutert.

Normdaten als Linked (Open) Data

Im Jahr 2010 hat die DNB sich der Semantic Web Community geöffnet und erstmals ihre Normdaten als Linked Data veröffentlicht. In der Gemeinsamen Normdatei (GND)⁵ sind Personen, Familien, Körperschaften, Kongresse und Veranstaltungen, Geografika, Schlagwörter und Werke (z.B. die Einheitssachtitel des Deutschen Musikarchivs) enthalten. Die heutige GND bestand damals aus drei getrennten Normdateien: der Personennamendatei, der Schlagwortnormdatei und der Gemeinsamen Körperschaftsdatei. Die Modellierung erfolgte mittels bestehender Ontologien wie dem Friend-of-a-friend-Vocabulary (FOAF)⁶ und dem Elementset der Resource Description and Access (RDA)⁷. Diese Umsetzung konnte den Datenstrukturen nur zum Teil gerecht werden. Daher wurde im Zuge der Zusammenführung der drei früheren Normdateien zur GND auch die RDF-Repräsentation angepasst: Die DNB spezifizierte eine eigene GND-Ontologie⁸, die im April 2012 produktiv ging. Die GND-Ontologie wird sukzessive mit Elementen aus anderen, bereits existierenden Vokabularen verknüpft, im ersten Schritt mit dem RDA Element Set und FOAF⁹. Sie wird seit ihrer Erstveröffentlichung laufend aktualisiert und auch zukünftig an kommende Formatänderungen angepasst.

Umstellung auf die offene Lizenz CCO

Ein wichtiger Aspekt bei der Datenveröffentlichung im Kontext von Linked Data ist die Frage nach der Nutzungslizenz: kommerzielle externe Anwendungen und Services können nur dann auf einem Dataset aufbauen, wenn durch dessen Nachnutzung keine rechtlichen Konsequenzen zu befürchten sind, sie also unter einer offenen Lizenz stehen. Daher entschloss sich auch die DNB dazu, ab Januar 2012 sämtliche im Format RDF verfügbaren DNB-Daten unter die Nutzungslizenz Creative Commons Zero (CCo)¹⁰ zu stellen¹¹. Seit dem 1. Juli 2012 stehen darüber hinaus alle Normdaten und ein Grossteil der Titeldaten der DNB inklusive Kataloganreicherungsdaten auch in anderen Formaten kostenfrei zur Verfügung und werden unter CCo zur freien Nachnutzung angeboten¹².

Die Entscheidung, die DNB-Daten schrittweise freizugeben und unter die Lizenz CCo zu stellen, ist eine logische Konsequenz aus den Entwicklungen in der zunehmend globalisierten und vernetzten Informationswelt. Eine freie Lizenz ist die erste Voraussetzung für die Nachnutzung der Daten im Semantic Web.

Bereitstellung von Titeldaten

Anfang 2012 wurde der Linked Data Service um Titeldaten erweitert – die Voraussetzung, um auch im Kontext von Linked Data auf die Titel der Deutschen Nationalbibliografie zu referenzieren.

Der Umfang der konvertierten Titeldaten umfasst derzeit die Deutsche Nationalbibliografie inklusive der darin enthaltenen fortlaufenden Sammelwerke der Zeitschriftendatenbank (ZDB)¹³. Musikalien, Tonträger und die Bestände des Deutschen Exilarchivs und des Deutschen Buch- und Schriftmuseums sind aktuell noch nicht umgesetzt. Während für bibliografische Textdaten bereits ein Application Profile durch die Titeldaten der DINI-AG-KIM-Gruppe¹⁴ spezifiziert wurde, gibt es für Archivalien und Sonderbestände derzeit noch keinen Standard, der implementiert werden kann. Eine enge Zusammenarbeit mit Fachleuten aus dem Bereich Archiv und Dokumentation ist erforderlich, um eine RDF-Repräsentation für Sonderbestände zu erarbeiten. Aktuell ist eine Kooperation der DINI-AG-KIM-Gruppe-Titeldaten mit Musikbibliothekaren in Planung, um ein Anwendungsprofil für Musikalien und Tonträger zu erarbeiten. Langfristiges Ziel ist es, sämtliche Daten der Deutschen Nationalbibliothek inklusive kontrollierter Vokabulare als Linked Open Data bereitzustellen. Weitere Schritte sind geplant¹⁵.

5 www.dnb.de/gnd

6 <http://xmlns.com/foaf/spec>

7 <http://rdvocab.info>

8 <http://d-nb.info/standards/elementset/gnd#>

9 vgl. Haffner, Alexander: Internationalisierung der GND durch das Semantic Web, 2012, http://www.kim-forum.org/Subsites/kim/DE/Materialien/Dokumente/dokumente_node.html

10 <http://creativecommons.org/publicdomain/zero/1.0>

11 Svensson, Lars G.: Licensing Library and Authority Data Under CCo: The DNB Experience, 2013. http://www.w3.org/2013/04/odw/odw13_submission_57.pdf

12 www.dnb.de/geschaeftsmodell

13 www.zeitschriftendatenbank.de

14 <https://wiki.dnb.de/display/DINIAGKIM/Titeldaten+Gruppe>

15 Vgl. S. Hartmann, J. Hauser: Die Deutsche Nationalbibliografie in der Linked Data Cloud. In: Bibliotheken: Tore zur Welt des Wissens / 101. Bibliothekartag in Hamburg 2012. Hrsg. von Klaus-Rainer Brintzinger u.a. – Hildesheim; Zürich; New York, NY: Olms, 2013 (S. 57-71)

Geplante Schritte und Ausblick

Neben bereits erwähnten inhaltlichen Erweiterungen ist vor allem der Ausbau externer Verlinkungen für die Zukunft vorgesehen: Die Titeldaten sollen in kommenden Releases mit Titeln der British Library (BL)¹⁶ verknüpft werden. Auch die Normdaten sollen weiter mit externen Verknüpfungen angereichert werden, so etwa zu Geonames¹⁷, den LCSH¹⁸ und RAMEAU¹⁹.

Die Deutsche Nationalbibliothek arbeitet ausserdem in unterschiedlichen Projekten und Initiativen mit, die sich mit der Bereitstellung und Nutzung von Daten im Semantic Weg beschäftigen, wie der Deutschen Digitalen Bibliothek²⁰ oder der europeana²¹. Darüber hinaus ist die DNB Mitglied der Early Experimenters Group der von der Library of Congress initiierten Bibliographic Framework Transition Initiative (BIBFRAME)²². Ziel dieser Initiative ist es, das MARC-21-Format durch ein nachhaltiges, RDF-basiertes Rahmenkonzept abzulösen. Die vorhandene DNB-Expertise hinsichtlich Konkordanzen, Formatumstiegen und

der Konvertierung sowie Bereitstellung von RDF-Daten sind hierbei wichtige Grundlagen.

Verstetigung des Services: RDF als gleichberechtigtes Exportformat

Ab Januar 2014 werden die Norm- und Titeldaten in RDF, die über den Linked Data Service bereitgestellt werden, in regulären Export-Releases mit bearbeitet²³. Mit diesem Schritt hat der Linked Data Service endgültig seinen Projektstatus verloren und wird in den Regelbetrieb überführt: Von nun an gelten verbindliche Zeitrahmen für die Anforderungserhebung, Tests, Realisierung und die Bereitstellung. Diese Abläufe werden für die Kunden transparent gemacht, sodass ggf. ausreichend Vorlauf für Anpassungen in ihren Systemen zur Verfügung steht. Die DNB schafft damit eine zuverlässige

Struktur für die Nachnutzung der Daten und die Einbindung in Anwendungen durch externe Nutzer. Zukünftig werden verbindlich in den Monaten Januar, Mai und September neue Releases veröffentlicht und anschliessend zeitnah Gesamtabzüge (Dumps) in RDF bereitgestellt.

Für Rückmeldungen, Fragen und Änderungsankündigungen sowie als allgemeiner Informationskanal wurden Mailinglisten eingeführt für den Linked Data Service im Allgemeinen²⁴ und die GND-Ontologie im Speziellen²⁵. Anregungen und Kritik zur Verbesserung des Linked Data Services sind willkommen über die Mailinglisten – weitere Kontaktmöglichkeiten sind auf der Website genannt²⁶.

Kontakt: j.hauser@dnb.de

16 www.bl.uk/

17 www.geonames.org/

18 <http://id.loc.gov/authorities/subjects.html>

19 <http://rameau.bnf.fr/>

20 www.deutsche-digitale-bibliothek.de/

21 www.europeana.eu

22 <http://bibframe.org/>

23 <https://wiki.dnb.de/display/ILTIS/ILTIS-Change-Management>

24 <http://lists.dnb.de/mailman/listinfo/lds>

25 <http://lists.dnb.de/mailman/listinfo/gnd-ontology>

26 www.dnb.de/lds

ABSTRACT

Du projet au produit. Le service de Web de données de la Bibliothèque nationale allemande

Depuis 2010, la Bibliothèque nationale allemande (DNB) fournit ses données au format RDF, conforme aux besoins du Web sémantique, via un service de Web des données (linked data).

Par cette initiative, les utilisateurs et groupes d'utilisateurs ont eu, pour la première fois, la possibilité de réutiliser les données, sans avoir besoin d'une quelconque connaissance des formats propres aux bibliothèques.

Initialement, il s'agissait d'un prototype, lequel a été développé ensuite dans le cadre de projets. Après plusieurs versions, il a été optimisé sur les plans technique, fonctionnel et organisationnel. Parallèlement, RDF s'est imposé comme format d'exportation: le service de Web de données est ainsi passé du projet au produit.

La contribution donne un aperçu touchant les motifs de la démarche, évoque les problèmes de licence, les étapes de développement, rappelle l'évolution des contenus et fournit un aperçu des prochaines étapes.

(traduction: fs)

Andreas Kellerhals, Direktor des Schweizerischen Bundesarchivs

Nach nur zweijähriger Entwicklungszeit haben Mitte September 2013 fünf Bundesämter das Open-Government-Data- (OGD-)Pilotportal opendata.admin.ch eröffnet und den rund sechsmonatigen Pilotbetrieb gestartet, während dem noch Daten des Kantons Zürich und weiterer Bundesstellen aufgeschaltet werden. In dieser Zeit wird auch politisch und praktisch über die weitere OGD-Politik zu entscheiden sein.

Am 16. September 2013 hat der Bund anlässlich der Open Knowledge Conference in Genf sein Open-Government-Data-(OGD-)Pilotportal mit 1617 Datensätzen und vier Anwendungen der Öffentlichkeit vorgestellt¹. Die zentral verfügbar gemachten Daten stammen vom Schweizerischen Bundesarchiv BAR, dem Bundesamt für Statistik BFS, dem Bundesamt für Landestopografie

swisstopo, dem Bundesamt für Meteorologie und Klimatologie MeteoSchweiz und der Schweizerischen Nationalbibliothek NB. Während der Testphase von rund sechs Monaten werden noch Daten des Kantons Zürich und anderer Bundesämter dazukommen. In dieser Zeit wird gleichzeitig eine schweizerische Open-Government-Data-Strategie formuliert, die Überführung des Pilotportals in ein definitives schweizerisches Portal vorbereitet und die Massnahmen für dessen dauerhaften Betrieb beschlossen².

Die Schweiz ist kein OGD-Trendsetter, aber der Weg zum Pilotportal war doch ein erstaunlich kurzer: Von der ersten

Offene Behördendaten sind frei zugängliche und wiederverwertbare Datenbestände öffentlicher Verwaltungen, welche diese in Erfüllung ihrer gesetzlichen Aufgaben sammeln.

OGD-Konferenz im Bundesarchiv 2011 bis zum Launch in Genf hat es nur zwei Jahre gedauert. Der Weg ist gesäumt von verschiedenen Hackdays und wichtigen Publikationen zu OGD in der Schweiz sowie einer Abklärung zum Potenzial volkswirtschaftlicher Wertschöpfung³. Aktive Akteure waren der Verein opendata.ch und die parlamentarische Gruppe Digitale Nachhaltigkeit, die das Thema auf die politische Agenda gebracht hat⁴. Parallel dazu wurde unter der Leitung des Bundesarchivs pragmatisch dieses Pilotportal realisiert, welches sich als Tatbeweis für eine wirtschaftliche Machbarkeit in das E-Government-Vorhaben integriert.

Was sind offene Behördendaten?

Offene Behördendaten sind frei zugängliche und wiederverwertbare Datenbestände öffentlicher Verwaltungen, welche diese in Erfüllung ihrer gesetzlichen Aufgaben sammeln. Diese Datenbestände sind umfangreich, in der Regel von hervorragender Qualität und enthalten wertvolle Informationen zu unterschiedlichsten Themen auch

jenseits des ursprünglichen Verwendungszwecks⁵.

Damit interessierte Dritte diese Daten direkt und frei verarbeiten können, dürfen keine datenschutz-, informationsschutz- oder urheberrechtlichen Einschränkungen bestehen; am besten sollten sie unter einer einheitlichen Lizenz nutzbar sein⁶. Sie müssen zudem klar beschrieben sein und technisch in einer Form vorliegen, die einfache Weiterverarbeitung überhaupt ermöglicht. Gemäss einem Fünf-Sterne-Anforderungskatalog⁷ müssen sie

- 1 opendata.admin.ch. Kontrolle aller Links: 23.–27.9.2013.
- 2 Vgl. dazu den Bericht des Bundesrates zu Open Government Data vom 13. September 2013 (www.news.admin.ch/message/index.html?lang=de&msg-id=50250).
- 3 André Golliez et. al. (2012): Open Government Data Studie Schweiz, Bern (www.itopia.ch/repository/Publikationen/OGD_Studie_Schweiz_Juni_2012.pdf); Beat Estermann (2013): Schweizer Gedächtnisinstitutionen im Internet-Zeitalter. Ergebnisse einer Pilotbefragung zu den Themenbereichen Open Data und Crowdsourcing, Bern (www.wirtschaft.bfh.ch/de/forschung/institute/e_government_institut/tabs/publikationen.html). Adelheid Bürgi-Schmelz (2013): Wirtschaftliche Auswirkungen von Open Government Data. Verfasst im Auftrag des Bundesarchivs (www.bar.admin.ch/aktuell/index.html?lang=de).
- 4 OGD als priorisiertes E-Government-Vorhaben B2.12 (www.egovernment.ch/de/umsetzung/katalog_vorhaben.php) resp. in der Diskussion um das neue Meteorologiegesez (vgl. www.meteoschweiz.admin.ch/web/de/meteoschweiz/portrait/datenliberalisierung.html).
- 5 Vgl. dazu Golliez, wie Fussnote 3, S. 5. Die Studie weist auf die grosse Bedeutung unterschiedlichster Datenbestände etwa im Umwelt-, Gesundheits- oder Verkehrs- und Kommunikationsbereich hin.
- 6 Z.B. Creative Commons, die Praxis ist heterogener (vgl. andere Datenportale, wie data.gov.uk/ oder www.govdata.de/). Die konkrete rechtliche Ausgestaltung in der Schweiz ist noch in Arbeit.
- 7 Vgl. <http://5stardata.info/>.

Porträt

Andreas Kellerhals ist seit dem 1.11.2004 Direktor des Schweizerischen Bundesarchivs. Er studierte an der Universität Bern Geschichte, Architekturgeschichte und Staatsrecht. 1985 schloss er mit dem Lizentiat ab und war 6 Jahre als Wissenschaftlicher Assistent an der Universität Bern tätig. 1991 trat er als Direktionsadjunkt in den Dienst des Bundesarchivs, wo er ab 1995 als Vizedirektor amtierte.



demnach im Web auffindbar und strukturiert sein, in nicht proprietären Formaten vorliegen, eine URI haben und verlinkbar sein, d.h. beispielsweise im Resource Description Framework (RDF)⁸ vorliegen.

8 RDF ermöglicht «logische Aussagen über beliebige Dinge» und ist «ein grundlegender Baustein des Semantischen Webs» (vgl. dazu http://de.wikipedia.org/wiki/Resource_Description_Framework).

9 Das BAR ist beispielsweise daran, die bereits digitalisierten Staatsrechnungen seit 1848 zu datafizieren, damit interessierte Forschende diese nicht nur konsultieren, sondern auch direkt auswerten und bearbeiten können. Das Risiko von «Fehlinterpretationen» besteht, aber Fehlinterpretationen gab und gibt es auch bei nicht offenen und nichtverlinkten Daten.

10 EU-Schätzungen nennen rund 40 Milliarden, das Spektrum ist nach oben offen. Allerdings ist nicht einfach vergleichbar, was genau unter Public Sector Information verstanden wird und auf welchen Annahmen die Berechnungen beruhen. Vgl. dazu mit zurückhaltend-vorsichtigem Ansatz und trotzdem positivem Schlussergebnis Bürgi-Schmelz, wie Fussnote 3.

11 Beispielsweise: Artikel zum Beschaffungsfilz (Ex-Beamte profitieren vom Filz, *Sonntagszeitung* 25.11.2012, S. 7), Karte zum Hochwasser-Risiko in der Schweiz (www.sonntagszeitung.ch/fileadmin/dj/vis/flut_soz/flutsoz.html) in der *Sonntagszeitung*, die *NZZ-Sommerserie* mit verschiedenen Schweizerkarten (www.nzz.ch/aktuell/schweiz/ende-der-sommerserie-schweizer-karten-1.18130619) bzw. die Datenseite von *Le Temps* (www.letemps.ch/datas/) alle mit häufig starkem Echo in den Social Media.

12 Leitbilder in der Diskussion um offene Daten sind immer *The Manchester Guardian*, 5. Mai 1821, Publikation einer Liste von Schulen in Manchester und Salford mit Angaben zu Schüleranzahlen sowie den Jahresdurchschnitt an Ausgaben (vgl. z.B. www.nzz.ch/aktuell/digital/open-data-republica-13-1.18078390#); anderes klassisches Beispiel sind die Mortalitätsstatistiken mit Diagrammen von Florence Nightingale aus dem Krimkrieg von 1858 (www.theguardian.com/news/datablog/2010/aug/13/florence-nightingale-graphics#).

Was für eine Bedeutung haben offene Daten – speziell auch Linked Open Data – für Archive? Grundsätzlich unterliegen alle offenen Behördendaten einer Archivierungspflicht. Wenn sie öffentlich zugänglich bleiben, muss eine angemessene Archivierungslösung auf dem Konzept einer ununterbrochenen Nutzungsmöglichkeit aufbauen. Fragen des Zugangs und auch der Erschliessung stellen sich damit neu. Gleichzeitig eröffnet diese Entwicklung für Archive neben einem enormen Modernisierungspotenzial die Möglichkeit, ihr spezielles Fachwissen in den Dienst zeitgemäss standardisierter Katalogisierung und Sammlung von Daten über Daten einzubringen, um Orientierung zu schaffen und allen Interessierten Zugang zu Informationen zu eröffnen. Gleichzeitig kann Archivgut produktiv in einen offenen Datenraum integriert und an andere Informationen im oder ausserhalb des Archivs anschlussfähig gemacht werden. Archivgut zu digitalisieren und zu datafizieren⁹ braucht zwar einen nicht zu unterschätzenden Aufwand, verspricht aber erfreuliche Erfolgsaussichten.

Nicht zuletzt eröffnen Linked Open Data ebenfalls wesentlich einfachere Wege für archivische Kooperation in freier Form, ohne dass für jedes Projekt neuer Aufwand betrieben werden muss und jedes Update einer Datenbank zu einem ressourcenmässigen Flaschenhals wird. Archive gewinnen damit an Flexibilität und Archivgut kann wesentlich besser valorisiert werden – möglicherweise auch auf neue, aber kreative und spannende Art und Weise. Archivarinnen und Archivare müssen dabei allerdings auch einen Kontrollverlust in Kauf nehmen: Daten offen anbieten heisst, dass wir nicht mehr bestimmen können, was damit geschieht – aber das ist ohnehin nicht unsere Aufgabe. Umso wichtiger wird damit aber unser Beitrag zur Authentizitätsgarantie, welche nach wie vor nur die Archive leisten können.

Offene Behördendaten: ein Angebot und ...

Was ist der Nutzen frei verwendbarer Daten? Das legitimatorische Mantra ist kurz: Offene Behördendaten stärken die Transparenz sowie die demokratische Partizipation. Auch haben sie ein

grosses Potenzial für (volks-)wirtschaftliche Wertschöpfung, selbst wenn die Schätzungen hier weit auseinandergehen¹⁰. Neue Business-Modelle werden sich entwickeln und bereits jetzt profitieren beispielsweise die Medien vom sogenannten Datenjournalismus¹¹.

Ebenso gewinnen öffentliche Verwaltungen intern durch offene Behördendaten an Effizienz: Teilen der vorhandenen Daten dank Standardisierungen, bessere Qualität durch mehr Nutzung und intensivere Nutzung durch den Wegfall verwaltungsinterner Verrechnungen. Einnahmeausfälle für einzelne Stellen können nicht al-

Die Forderung nach offenen Behördendaten schliesst sich dem Diskurs um das Öffentlichkeitsprinzip an, grenzt sich aber gleichzeitig klar von Wikileaks ab.

lein durch volkswirtschaftliche Wertschöpfung kompensiert werden, sondern müssen im Budget der Verwaltungsstellen ausgeglichen werden.

Ein staatspolitisch-demokratischer Nutzen ergibt sich aus der zusätzlichen Transparenz. Die Forderung nach offenen Behördendaten schliesst sich dem Diskurs um das Öffentlichkeitsprinzip an, grenzt sich aber gleichzeitig klar von Wikileaks ab. Auch steht und stand die Forderung nach mehr Transparenz immer in einem spannungsvollen Verhältnis zur Arkanpolitik. Seinen Anfang nahm dieses Spannungsfeld aber nicht mit der vielzitierten Publikation der Schulstatistik des *Manchester Guardian* von 1821¹², sondern lässt sich anhand des Beispiels der (amtlichen) Statistik über die ganze (europäische) Geschichte verfolgen. Informationsmonopolisierung von Daten und anderen Dokumenten ist stets ein Herrschaftsmittel, und so galt auch die Veröffentlichung von Statistiken als Verrat, der mit dem Tod bestraft werden konnte. Nichtsdestotrotz wurden im Laufe der Zeit immer mehr statistische Werke publiziert, nicht zuletzt in propagandistischer Absicht, und statistische Bevölkerungsdaten zirkulierten im 18. Jahrhundert ziemlich frei unter den Gelehrten und Pfarrherren über die Staatsgrenzen hinaus.

Die konkrete Wirkung offener Daten auf die politische Diskussion ist allerdings nicht einfach zu ermitteln; auch mag man bezweifeln, ob die Publikation von Wetterdaten oder der Fliessgewässertemperaturen – wie in der bei Berner Badenden beliebten Aare-App – politische Partizipation fördert. Wenn auch positive nichtökonomische Wirkungen und lebenspraktische Erleichterungen mehr dem Wohlergehen als dem Gemeinwohl (welfare effects) dienen, tragen sie doch zum «grössten Glücks der grössten Zahl» (Jeremy Bentham) bei.

...eine Herausforderung

Was sind mögliche Risiken und Nebenwirkungen? Bürgerinnen und Bürger mit qualitativ hochstehenden Daten und Informationen in die Lage zu versetzen, die Begründungen politischer Entscheidungen besser zu verstehen, sich selber eine eigene Meinung zu bilden oder überhaupt wohlbegründete Entscheidungen zu treffen – mit anderen Worten: durch intellektuelle Anstrengung aus möglicher Unmündigkeit herauszutreten –, ist ein hehres Anliegen. Offene Daten ersetzen allerdings das Denken nicht. Technisch frei kombinierbare Daten ergeben nicht automatisch inhaltlich zulässige Aussagen. Es muss immer ernsthaft geprüft werden, ob eine Kombination dekontextualisierter Daten zu logisch über-

zeugenden Ergebnissen führt. Jede Diskussion wird offen, der Umgang mit Daten überprüfbar und Resultate können fundiert bestritten, korrigiert

Offene Daten ersetzen allerdings das Denken nicht. Technisch frei kombinierbare Daten ergeben nicht automatisch inhaltlich zulässige Aussagen.

oder gar widerlegt werden. Damit geht die Pflicht einher, die neuen Möglichkeiten mit Sorgfalt und Respekt zu nutzen. Diskussionen werden sich von Fragen nach blossen Facts & Figures auf Methodenfragen verschieben: Was ist eine zulässige Argumentation? Was kann gesagt werden?¹³

Es kann als Aufgabe des Staats verstanden werden, Voraussetzungen für eine aktive Partizipation von Bürgerinnen und Bürgern zu schaffen: Daten sind nicht einfach Belege einer Meinung, sondern Rohstoff zur Meinungsbildung und müssen sich im Licht anderer Daten behaupten. OGD ist somit Teil einer Wissensinfrastruktur, die präventiv gegen Uninformiertheit und strukturell, nicht individuell begründetes Unwissen wirken soll¹⁴.

Unsere Gesellschaft braucht aber nicht allein offene Behörden-, Wissenschafts-

und anderen Daten, sondern auch intelligente Information. Das umfasst sowohl Inhalte aus verlässlichen Quellen als auch Instrumente mit transparenter Funktionsweise. Wir brauchen nicht bloss informationelle, sondern auch algorithmische Autonomie¹⁵. Ausserdem braucht es über die offenen (Behörden-)Daten hinaus auch einen offenen Geist. Im herrschenden Zwang zu argumentativer Auseinandersetzung kann der Staat dazu beitragen, dass alle gleich gute Voraussetzungen zur Partizipation haben. Dazu braucht es aber auch eine Haltung, wie sie Voltaire zugeschrieben wird: «Ich verachte Ihre Meinung, aber ich gäbe mein Leben dafür, dass Sie sie sagen dürfen.»¹⁶ Wenn nicht nur die Daten, sondern auch der Geist offen ist, dann haben wir etwas Positives erreicht.

Kontakt: andreas.kellerhals@bar.admin.ch

13 Berühmtes Beispiel ist die Reinhart-Rogoff-Debatte um die Studie *Growth in a Time of Debt*, in der sich weltwirtschaftliche Ratschläge als Ergebnis von falschen Formeln und unklar begründeter Datenauswahl bei der Modellbildung entpuppte – wie es die *New York Times* zugespitzt formulierte: in der Wirtschaftsexpertise von Excel-Expertise abhing. Vgl. dazu u.a. Paul Krugman, *The Excel Depression*, *New York Times*, 18. April 2013 (www.nytimes.com/2013/04/19/opinion/krugman-the-excel-depression.html?_r=0).

14 Willke, Helmut (1996): *Die Steuerungsfunktion des Staates aus systemtheoretischer Sicht. Schritte zur Legitimierung einer wissensbasierten Infrastruktur*. In *Staatsaufgaben* (Hg. von Dieter Grimm), S. 685–711. Frankfurt/M.

15 Rainer Kuhlen (1999): *Die Konsequenzen von Informationsassistenten. Was bedeutet informationelle Autonomie oder wie kann Vertrauen in elektronische Dienste in offenen Informationsmärkten gesichert werden?* Frankfurt/M (www.kuhlen.name/MATERIALIEN/Publikationen1995-2000/informationsassistenten.pdf).

16 Evelyn Beatrice Hall (Pseudonym Stephen G. Tallentyre) (1906): *The Friends of Voltaire* (vgl. <http://de.wikiquote.org/wiki/Voltaire>).

ABSTRACT

opendata.admin.ch

Après seulement deux années de développement, cinq offices fédéraux ont publié à la mi-septembre le portail-pilote Open Government Data (OGD) *opendata.admin.ch* et ont démarré ainsi un projet d'environ six mois durant lesquels seront ajoutées les données du canton de Zurich et d'autres offices fédéraux. Cette même période devrait permettre de décider de l'orientation politique et pratique de la stratégie de l'OGD.

L'OGD possède le potentiel – surtout sous sa forme en réseau, Linked Open Data – de transformer profondément l'archivage, aussi bien du point de vue de la permanence de la disponibilité des données que de celui de leur description et de leur organisation. Le patrimoine archivistique s'intègre dans un espace de données ouvertes. Ce potentiel de modernisation ne doit pas nous faire perdre de vue les risques et effets secondaires. Les données ouvertes représentent aussi un défi et une obligation pour leurs utilisateurs: il est également nécessaire de recourir, parallèlement aux données ouvertes, à des outils transparents, et de faire preuve d'un esprit ouvert. (traduction: fs)

Inscrire les données du Service de la culture dans le Web pour mieux valoriser le patrimoine

Alain Dubois, archiviste paléographe,
Archives de l'Etat du Valais

Les questions relatives à la mise en réseau et à l'interopérabilité entre les systèmes d'information et de description, ainsi qu'à l'échange et/ou à la réutilisation des données constituent aujourd'hui un enjeu professionnel majeur pour les institutions de conservation du patrimoine.

Elles sont actuellement au cœur des réflexions menées par le Service de la culture du canton du Valais, dans le sillage de la mise en place d'une plate-forme de pérennisation et d'archivage pour les collections numériques et numérisées des Archives de l'Etat du Valais, de la Médiathèque Valais et des Musées cantonaux conforme aux normes et standards internationaux en la matière. Elles aboutiront concrètement à la fin du premier semestre 2014 par la mise en ligne d'un outil de recherche fédérée dédié au patrimoine valaisan, l'ouverture d'un entrepôt OAI-PMH (open data culturel) et l'alignement des notices descriptives selon les principes du Web de données (linked open data).

Le présent article retrace en trois parties les étapes successives d'un projet qui a débuté en 2011 et qui a pour ambition d'offrir à toute personne intéressée un accès aux ressources patrimoniales conservées par les Archives de l'Etat du Valais, la Médiathèque Valais et les Musées cantonaux à partir d'un portail unique.

1. L'outil de recherche fédérée du Service de la culture

L'origine du projet de recherche fédérée se situe en fait dans le sillage de l'ouverture de la plate-forme de pérennisation et d'archivage en mai 2011. Il s'est agi, à cet instant, de mettre en réseau et de donner accès aux ressources numériques et numérisées disponibles sur cette plate-forme, indépendamment des outils métier institutionnels usuels – scopeArchiv pour les Archives de l'Etat du Valais et le catalogue collectif RERO pour la Médiathèque Valais. Rapidement, toutefois, s'est imposée l'idée de réaliser un outil qui, au-delà des seules ressources numériques et numérisées, donne accès à toutes les ressources en lien avec le patrimoine conservées par le Service de la culture. C'est ainsi qu'est né le projet d'outil de recherche fédérée. Actuellement en voie de finalisation, ce projet s'est déroulé en deux phases:

- une phase-pilote (avril-décembre 2011), qui a permis de démontrer la faisabilité d'une mise en réseau de ressources disparates;
- une phase de déploiement (janvier 2012 juin 2014), qui permet de consolider les acquis de la phase-pilote.

Sans revenir sur tous les éléments du projet, j'en mentionnerai ici les principaux.

Particularité de la structure de projet

Le projet est piloté par la cellule e-culture du Service de la culture. Cette structure originale, qui regroupe des représentants des Archives de l'Etat du Valais, de la Médiathèque Valais, des Musées cantonaux et de l'Encouragement des activités culturelles, a notamment pour objectif de mener des projets transversaux en lien avec les technologies numériques et relatives à la collecte, à la conservation, à la mise en valeur et à la promotion du patrimoine culturel valaisan et des ressources qui lui sont associées.

Particularité du projet

Dans sa philosophie, mais à une échelle nettement moindre, le projet de re-

cherche fédérée se rapproche par ailleurs de la bibliothèque numérique européenne lancée en novembre 2008 (Europeana). Il s'agit, en effet, de mettre à la disposition du public intéressé par le patrimoine valaisan un outil qui fédère des ressources descriptives qui ne le sont pas encore (fonds d'archives, collections audiovisuelles, bibliothèque numérique et collections muséales). Sur ce point, le défi principal porte sur la mise en commun et l'alignement de ressources décrites selon des normes et standards qui diffèrent fortement.

Choix de l'outil informatique

La particularité du projet a eu une incidence significative sur le choix de la solution informatique. Au terme d'une procédure sur invitation, c'est l'entreprise française Antidot SAS qui a été retenue pour réaliser l'outil de recherche fédérée¹. Ce choix s'est notamment justifié en raison des solides références dont disposait l'entreprise dans le domaine de la gestion de l'information, son expérience avérée dans la mise en œuvre d'outils de recherche fédérée et d'accès à l'information à travers de multiples portails et l'intégration des outils du web de données qu'elle proposait².

Choix du corpus

L'outil de recherche intègre actuellement 362 704 notices descriptives qui se répartissent comme suit:

- 68 947 notices de fonds d'archives conservés aux Archives de l'Etat du Valais (Etat du Valais, communes, bourgesses et paroisses, ainsi que fonds privés) et progressivement mis à la disposition du public sur le portail scopeQuery³;
- 279 182 notices de la bibliographie valaisanne disponibles sur le portail valaisan de RERO;
- 14 225 notices des ressources audiovisuelles conservées à la Médiathèque Valais – Martigny (photographies, enregistrements sonores et films) et mises à la disposition du public sur le portail «mémoire audiovisuelle du Valais»⁴;

1 www.antidot.net/ [consulté le 15 octobre 2013].

2 L'entreprise Antidot a notamment réalisé le portail du Centre national de la recherche scientifique pour les sciences humaines et sociales (www.rechercheisidore.fr/ [consulté le 15 octobre 2013]).

3 <http://scopequery.vs.ch/suchinfo.aspx> [consulté le 15 octobre 2013].

4 <http://archives.memovs.ch/> [consulté le 15 octobre 2013].

– 350 notices issues des deux catalogues «Collectionner au cœur des Alpes» publiés par le Musée d'art et le Musée d'histoire du Valais⁵.

Livraison et traitement des données

L'un des principaux avantages du module Antidot Finder Suite, utilisé pour le traitement des données, réside dans le fait qu'il accepte en entrée des formats très variés, indépendamment de leur source. C'est ainsi que chaque office a pu déterminer son mode de livraison: les Archives de l'Etat du Valais exportent leurs unités de description en EAD (Encoded Archival Description), tandis que les notices de la bibliographie valaisanne sont livrées en Dublin Core et celles des collections audiovisuelles sont extraites de l'entrepôt OAI-PMH de la Médiathèque Valais – Martigny. Quant aux Musées cantonaux, ils fournissent leurs notices descriptives dans un fichier Excel.

Le module Antidot Finder Suite aligne ensuite les différents flux et les transforme en un format pivot dérivé du Dublin Core, avant d'appliquer sur le corpus ainsi constitué différents filtres d'indexation.

Editorialisation de l'outil de recherche fédérée

Le site dédié à la recherche fédérée est actuellement en cours de création⁶. A terme, il offrira au public un outil de recherche dont les résultats pourront être affinés à l'aide de facettes (provenance institutionnelle, lieux, périodes historiques et personnes). Il mettra également

en exergue un «document du mois» qui se fera l'écho de l'actualité d'un office (exposition, conférence, journée thématique ou encore nouvelle publication).

2. Vers un open data culturel

Dès les origines du projet, soucieux d'en assurer une diffusion aussi large que possible et dans le sillage des initiatives actuelles en la matière, le Service de la culture a décidé de rendre ses données interopérables et réutilisables. C'est ainsi que s'est très rapidement imposée l'idée de créer un portail qui rendent interopérables et moissonnables les données, dans la perspective de créer une plate-forme d'open data culturel au niveau valaisan. Sur ce point, le Service de la culture a pu bénéficier de l'expérience de la Médiathèque Valais, qui rend disponibles depuis 2009 déjà ses collections audiovisuelles à travers un entrepôt OAI-PMH⁷. Ce modèle a donc été repris et adapté en fonction des besoins. C'est ainsi que l'entrepôt OAI-PMH du Service de la culture, actuellement en cours de validation auprès de l'*Open Archive Initiative*⁸ et mis à disposition au début de l'année 2014 sur le site Internet de l'Etat du Valais⁹, proposera ses données dans deux formats: le format oai_dc, constitutif de tout entrepôt OAI-PMH, et le modèle de données d'Europeana (EDM).

Les sets seront ensuite rendus accessibles sous deux formes:

– par provenance institutionnelle (sets «Archives de l'Etat du Valais», «Bibliothèque numérique», «Mémoire audio-

visuelle du Valais» et «Musées cantonaux»);

– par type de ressources (sets «Photographies», «Enregistrements sonores», «Films», «Monographies», «Articles» et «Fonds d'archives»).

3. L'intégration du linked open data

Rendre interopérables ses données à travers un entrepôt OAI-PMH constitue certes une première étape importante, mais elle reste néanmoins insuffisante au regard des possibilités qu'offre désormais le Web de données en matière d'échange et de réutilisation des données. C'est ainsi que l'alignement des notices descriptives selon le modèle de

5 Ruedin, P., *Collectionner au cœur des Alpes: le Musée d'art du Valais, Sion, Sion – Paris 2007*; Elsig, P., Morand, M. C., *Collectionner au cœur des Alpes: le Musée d'histoire du Valais, Sion, Sion – [Paris] 2013*.

6 C'est l'entreprise Conchita-Plus qui a été choisie pour créer l'interface de consultation (www.conchita-plus.ch/ [consulté le 15 octobre 2013]).

7 Consultable à l'adresse <http://xml.memovs.ch/oai/>, l'entrepôt OAI-PMH de la Médiathèque Valais rend disponibles sous forme de sets ses collections de photographies, d'enregistrements sonores et de films. Les données sont disponibles dans les formats suivants: oai_dc, qdc (Dublin Core qualifié), ead, marcxml et ese (Europeana) [consulté le 15 octobre 2013].

8 www.openarchives.org/ [consulté le 15 octobre 2013].

9 www.vs.ch/oai.

Impressum

arbido print

arbido newsletter

arbido website

print: © arbido ISSN 1420-102X
website + newsletter: © arbido newsl. ISSN 1661-786X
www.arbido.ch

Offizielle Publikationsorgane – Organes officiels de publication

Bibliothek Information Schweiz BIS (www.bis.info)
Verein Schweizerischer Archivarinnen und Archivare VSA-AAS (www.vsa-aaa.org)
Bibliothèque Information Suisse BIS
Association des archivistes suisses AAS-VSA

Chefredaktor – Rédacteur en chef

Stéphane Gillioz (sg), E-Mail: stephane.gillioz@gmail.com

Redaktion BIS – Rédaction BIS

Stephan Holländer (sth), Katja Böspflug (kb), Daniela Rüegg (dr)

Redaktion VSA-AAS – Rédaction AAS-VSA

Frédéric Sardet (fs), Annkristin Schlichte (as)

Layout, Druck – Graphisme, Impression

Stämpfli Publikationen AG, Wölflistrasse 1, Postfach 8326, 3001 Bern, www.staempfli.com

Inserate – Annonces

Tel. 031 300 63 89, Fax 031 300 63 90
E-Mail: inserate@staempfli.com

Abonnement – Abonnements

Tel. 031 300 63 41, Fax 031 300 63 90
E-Mail: abonnement@staempfli.com

4 Hefte/Jahr: Fr. 115.– (im Mitgliederbeitrag inbegriffen).
Auslandsabonnemente: Fr. 135.– (inkl. Porto).
Einzelnummer: Fr. 30.– (plus Porto und Verpackung).
4 cahiers/année: Fr. 115.– (compris pour les adhérents).
Abonnements de l'étranger: Fr. 135.– (affranchissement compris).
Vente au numéro: Fr. 30.– (plus affranchissement et emballage).

Termine – Dates

Nächste Ausgabe erscheint am 20.2.2014
Le prochain numéro paraît le 20.2.2014

Thema – Thème

Alpenraum / Espace alpin / Spazio alpino

Redaktionsschluss – Délai de rédaction: 18.1.2014

Inserateschluss – Délai d'insertion: 14.1.2014

Newsletter

Pro Jahr erscheinen 8 bis 12 elektronische Newsletters.
Siehe www.arbido.ch
Parution de 8 à 12 newsletters électroniques par an.
Voir www.arbido.ch

Publiziert mit Unterstützung der Schweizerischen Akademie der Geistes- und Sozialwissenschaften SAGW
Publié avec le soutien de l'Académie suisse des sciences humaines et sociales ASSH

données d'Europeana permettra aux données du Service de la culture d'être présentées selon la structure du RDF (Ressource Description Framework).

Le référentiel des lieux, par exemple, sera aligné sur GeoNames, base de données géographique de plus de 6,5 millions de lieux existants¹⁰. Des discussions sont par ailleurs en cours à l'interne pour aligner le référentiel des personnes avec le fichier d'autorités VIAF¹¹. Par contre, l'abandon – temporaire toutefois – du référentiel des matières à ce stade du projet reporte l'alignement de ce dernier avec le thésaurus-matières pour l'indexation des archives locales proposé par le Service interministériel des Archives de France¹² ou la CDU abrégée¹³.

10 www.geonames.org/ [consulté le 15 octobre 2013].

11 <http://viaf.org/> [consulté le 15 octobre 2013].

12 <http://data.culture.fr/thesaurus/page/ark:/67717/Matiere> [consulté le 15 octobre 2013].

13 <http://udcdata.info/> [consulté le 15 octobre 2013].

Au final, les réflexions autour de la création d'un outil de recherche fédérée permettront non seulement d'offrir à toute personne intéressée un portail d'accès unique au patrimoine valaisan, qui pourrait du reste tout à fait être enrichi par des ressources en lien avec le Valais provenant d'autres institutions, mais surtout de l'inscrire dans le sillage des

projets d'open data et de linked open data actuellement en cours en matière de valorisation du patrimoine. En somme d'inscrire pleinement les données du Service de la culture dans le Web pour les valoriser au mieux.

Contact: alain.dubois@admin.vs.ch

ABSTRACT

Die Valorisierung des Kulturguts durch Onlinestellung von Daten der Dienststelle für Kultur des Kantons Wallis

Die Überlegungen zur Onlinestellung und Interoperabilität zwischen verschiedenen Informationssystemen, ebenso wie der Austausch und/oder die Wiederverwendung von Daten stehen im Zentrum der Überlegungen der Dienststelle für Kultur des Kantons Wallis, die im Gefolge der Inbetriebnahme der Speicher- und Archivierungsplattform für die digitalen und digitalisierten Bestände des Staatsarchivs des Kantons Wallis, der Mediathek und der kantonalen Museen angestellt wurden. Am Ende des ersten Semesters 2014 sollen diese Überlegungen ganz konkret zur Onlinestellung einer gemeinsamen Suche für das Walliser Kulturgut, zur Öffnung eines Datenlagers Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) und zum Abgleich der beschreibenden Daten nach den Grundsätzen des semantischen Webs führen.

Der vorliegende Artikel zeigt die wichtigsten Etappen des Projekts auf, das 2011 begonnen wurde und das zum Ziel hat, allen Interessierten den Zugang zu den kulturellen Ressourcen der Dienststelle für Kultur des Kantons Wallis über ein Guichet virtuel («One-Stop-Portal») zu ermöglichen.

(Übersetzung: as)

Linked Open Data in archivischen Fachinformationssystemen

Peter Sandner, Archivar,
Leiter der Abteilung Archivfachliche
IT/Digitales Archiv im Hessischen
Hauptstaatsarchiv in Wiesbaden

Die grosse Idee der vernetzten offenen Daten erreicht über den Umweg der übergreifenden Portalangebote auch die Archive und ihre Fachinformationssysteme. Bei der Erneuerung des hessischen Archivinformationssystems¹, das gemeinsam für das Hessische Landesarchiv und das Niedersächsische Landesarchiv entwickelt wird, ist von Anfang an erörtert worden, wie die Idee der Linked Open Data (LOD) in der Praxis realisiert werden kann.

URI für archivische Erschliessungsobjekte

Um Daten miteinander verknüpfen zu können, ist ein eindeutiger Identifikator unverzichtbar. Seit Langem werden Archivalien und Archivbestände, wie es ISAD (G)² beschreibt, mit Signaturen versehen, um sie eindeutig identifizieren zu können. Auch die Archive selbst haben in der Regel eindeutige Kürzel. In den meisten Fällen genügt also die Archivsignatur den Ansprüchen an einen einheitlichen Ressourcenidentifikator (URI – Uniform Resource Identifier), aber eben doch nicht immer. Probleme können sich ergeben bei Korrekturen, Umsignierungen oder bei

Mehrfachverzeichnungen eines Objekts. Daher ist die Verwendung maschinell generierter Identifikatoren der sicherere Weg zur eindeutigen Bezeichnung archivischer Erschliessungsobjekte.

- 1 Peter Sandner: Von der Findmitteldatenbank zum virtuellen Lesesaal im Netz. Rechtliche Fragen bei der Umgestaltung des Archivinformationssystems HADIS, in: Alles was Recht ist. Archivische Fragen – juristische Antworten. 81. Deutscher Archivtag in Bremen, Fulda 2012, S. 135–145
- 2 www.ica.org/10207/standards/isadg-general-international-standard-archival-description-second-edition.html

jekte. Dabei können beispielsweise Datenbank-ID verwendet werden. Sicherlich ist die längerfristige Stabilität solcher ID wünschenswert, noch wichtiger als ewige Geltung aber ist die Eindeutigkeit.

Verwendung von HTTP-Links für archivische Erschließungsobjekte

Eine Vorschrift zur Verwendung von HTTP-Links für archivische Erschließungsobjekte findet sich in den Archivgesetzen zwar nicht. Eine solche Forderung steht aber im Einklang mit dem Auftrag des neu gefassten Hessischen Archivgesetzes, Archivgut «unter Anwendung moderner Technologien für die öffentliche Nutzung zugänglich zu machen»³. Schon vergleichsweise früh haben die hessischen Staatsarchive ihre Erschließungsdaten im Archivinformationssystem HADIS unter www.hadis.hessen.de recherchierbar gemacht⁴. Allerdings waren dort die standardmäßig angebotenen HTTP-Adressen keine festen, sondern nur temporäre Web-Links, deren Gültigkeit mit Ablauf der Session verfiel. Diese Einschränkung führte zu Unmut bei jenen Anwendern, die Links zu HADIS-Objekten setzen wollten. Seit 2006 werden von HADIS auch permanente Links für diverse Portale bereitgestellt. Infolgedessen sind hessische Erschließungsdaten u.a. aus spartenübergreifenden Angeboten wie dem BAM-Portal oder der Deutschen Digitalen Bibliothek (DDB) heraus verlinkt worden⁵. Von der HADIS-Nachfolgeanwendung werden durchweg feste HTTP-Adressen so angeboten, dass eine Verlinkung zu den Erschließungsobjekten für jeden leicht möglich ist. Dabei kommt das archivische Credo über die Bedeutung des Kontextes zur Geltung, indem jeder Erschließungsdatensatz im Zusammenhang mit der Gliederung (Klassifikation, Tektonik) angezeigt wird; ein Navigationsbaum ermöglicht das Überwechseln zu benachbarten Objekten.

Eignung archivischer Erschließungsdaten als Open Data

Manche hinterfragen, ob archivische Erschließungsdaten überhaupt als Open Data geeignet seien. Diese Zweifel betreffen zum einen die – verglichen mit Bibliotheksdaten – heterogene Struktur der Archivdaten. Mehr noch

aber führen die hohen Anforderungen des Persönlichkeitsschutzes zu einer Zurückhaltung mancher Archive bei Open Data. Die öffentlichen Archive verstehen sich zu Recht auch als Datenschutzbehörden. Sie haben die gesetzliche Aufgabe, das Freiheitsrecht auf informationelle Selbstbestimmung⁶ auch nach Abgabe der Unterlagen ans Archiv zu gewährleisten. Dazu dienen vor allem die archivgesetzlichen Schutz- oder Sperrfristen. Aber deshalb

Die öffentlichen Archive verstehen sich zu Recht auch als Datenschutzbehörden. Sie haben die gesetzliche Aufgabe, das Freiheitsrecht auf informationelle Selbstbestimmung auch nach Abgabe der Unterlagen ans Archiv zu gewährleisten.

müssen die betreffenden archivischen Erschließungsdaten keineswegs vollständig unter Verschluss gehalten werden. Vielmehr ist es Aufgabe der Archive, das Recht auf Persönlichkeitsschutz mit dem Grundrecht der Wissenschafts- und Forschungsfreiheit⁷ in Einklang zu bringen. Deshalb hat die bundesdeutsche Archivreferentenkonferenz 2007 eine Empfehlung zum Umgang mit sensiblen Erschließungsdaten verabschiedet⁸. Dieses Papier arbeitet heraus, dass Erschließungsdaten durchaus auch dann schon in öffentlich zugänglichen Netzen publiziert werden dürfen, wenn die Akteninhalte selbst noch Schutzfristen unterliegen. Entscheidend ist, dass der Erschließungstext die schutzwürdige Information (etwa einen Personennamen) nicht preisgibt. Archivinformationssysteme sollten diese Unterscheidung – so wie in Hessen seit vielen Jahren der Fall – abbilden und die beiden Merkmale «Zeitpunkt der Onlinepublikation der Erschließung» und «Ende der Schutzfrist» getrennt speichern. Dies sollte auf der Ebene der einzelnen Verzeichnungseinheit geschehen, und nicht pauschal für den ganzen Archivbestand, denn unter Onlinebedingungen entfällt die Notwendigkeit, die Fertigstellung eines kompletten Onlinefindbuchs abzuwarten. Dieses Verfahren erfordert unter den Vorzeichen von Linked Open Data eine erhöhte Wachsamkeit bei der Bearbeitung, denn ein versehentlich veröffentlichter Datensatz lässt sich kaum zurückholen.

Bereitstellung standardisierter Informationen zu den Objekten

Inzwischen hat sich EAD⁹ zum Standardaustauschformat für archivische Informationen entwickelt. In den letzten Jahren ist es im Förderprogramm der Deutschen Forschungsgemeinschaft (DFG) zur Retrokonversion von Findmitteln verwendet worden. Auch die Deutsche Digitale Bibliothek (DDB) nutzt EAD, um Daten von den Archiven in Empfang zu nehmen. Inzwischen haben die deutschen Archive sich über ein einheitliches EAD-Profil verständigt¹⁰. Die DDB reicht die Daten weiter an die Europeana, die diese in das klassische LOD-Format RDF verwandelt¹¹ und so öffentlich bereitstellt.

3 § 1 (1) Hessisches Archivgesetz (HArchivG). Vom 26. November 2012, Gesetz- und Verordnungsblatt für das Land Hessen Teil I, S. 458

4 Bernward Helfer: Das hessische Erschließungssystem HADIS 2000, in: Angelika Menne-Haritz (Hg.): Online-Findbücher, Suchmaschinen und Portale, Marburg 2002, S. 153–171, auch unter http://archivschule.de/uploads/Publikation/VOE35/Voe35_14.HELPER.pdf; Peter Haberkorn: Das Fachinformationssystem HADIS der hessischen Staatsarchive, in: Gerald Maier, Thomas Fritz (Hg.): Archivische Informationssysteme in der digitalen Welt. Aktuelle Perspektiven, Stuttgart 2010, S. 181–195

5 www.bam-portal.de; www.deutsche-digitale-bibliothek.de

6 In Deutschland abgeleitet aus der Menschenwürde und dem Recht auf freie Entfaltung der Persönlichkeit: Artikel 1 (1) bzw. Artikel 2 (1) Grundgesetz

7 Artikel 5 (3) Grundgesetz

8 www.bundesarchiv.de/imperia/md/content/bundesarchiv_de/fachinformation/ark/20070320_veroeffentlichungsgrundsaeetze_ark.pdf

9 EAD = Encoded Archival Description; www.loc.gov/ead

10 www.landesarchiv-bw.de/ead; Ulrich Fischer, Sigrid Schieber, Wolfgang Krauth, Christina Wolf: Ein EAD-Profil für Deutschland.

EAD(DDB) als Vorschlag für ein gemeinsames Austauschformat deutscher Archive, in: *Archivar. Zeitschrift für Archivwesen* 2/2012, S. 160–162, auch unter www.archive.nrw.de/archivar/hefte/2012/ausgabe2/ARCHIVAR_02-12_internet.pdf

11 <http://pro.europeana.eu/web/guest/in-brief>

Offenheit der Daten darf nicht auf Kosten der Datenqualität gehen. Daher widmete der LOD-Summit 2013 im kanadischen Montreal der Datenpflege («Curation») bei LOD eine eigene Sektion¹². Unter diesem Gesichtspunkt sind auch die Modalitäten der Datenaktualisierung zu klären. Die hessischen Staatsarchive hatten 2006 mit dem BAM-Portal eine tägliche Aktualisierung vereinbart, um mögliche Fehler – insbesondere solche, die den Persönlichkeitsschutz betrafen – ohne Zeitverzug bereinigen zu können. Derzeit werden die Daten bei der DDB noch nach Absprache im Einzelfall aktuali-

Offenheit der Daten darf nicht auf Kosten der Datenqualität gehen. Daher widmete der LOD-Summit 2013 im kanadischen Montreal der Datenpflege («Curation») bei LOD eine eigene Sektion.

siert, aber das ist nur eine Übergangslösung in der Startphase. Nach gegenwärtiger Planung werden die hessischen Archive ihre EAD-Dateien ohne Zugangsbeschränkung im Internet verfügbar machen. Über einen Zeitstempel lässt sich erkennen, wann eine Aktualisierung erfolgt ist. Ob darüber hinaus auch unmittelbar RDF-Daten publiziert werden, ist noch zu klären.

12 <http://summit2013.lodlam.net/2013/08/18/curation-of-lod/>

13 <http://pro.europeana.eu/web/guest/linked-open-data>; <http://creativecommons.org/publicdomain/zero/1.0/deed.de>

14 <http://summit2013.lodlam.net/2013/08/18/curation-of-lod/>; <http://summit2013.lodlam.net/2013/07/08/notes-from-normalizing-licensing-and-data-models/>

15 <http://de.wikipedia.org/wiki/Wikipedia:Lizenzbestimmungen>; <http://creativecommons.org/licenses/by-sa/3.0/>

16 Richtlinie 2013/37/EU des Europäischen Parlaments und des Rates vom 26. Juni 2013, Amtsblatt Nr. L 175 vom 27/06/2013 S. 0001–0008, auch unter <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:175:0001:01:DE:HTML>

17 www.lagis-hessen.de/de/subjects/index/sn/hadis

18 www.regesta-imperii.de; <http://monasterium.net>

Offenheit

Es wird diskutiert, wann Daten wirklich als «offen» zu bezeichnen sind. Die Europeana publiziert Erschließungsdaten nur, wenn sie unter der (Nicht-)Lizenz CCo verwendet werden können¹³. Dagegen wurden beim LOD Summit sachliche und rechtliche Zweifel an dieser umfassenden Lizenzierung artikuliert¹⁴. Ein Verzicht auf die Nennung des Datenurhebers widerspricht dem Ziel der Archive, ihre Archivalien besser auffindbar zu machen. Und bei Weiterverwendung der Daten zu rechtswidrigen Zwecken könnte die Reputation des Archivs leiden. Die deutschsprachige Wikipedia erlaubt eine Nachnutzung ihrer Artikel unter der Creative Commons-Lizenz BY-SA, die die Nennung des Urhebers fordert und die Weitergabe von Bearbeitungen unter den gleichen Bedingungen zulässt¹⁵. Die möglichen Nachteile einer CCo-Lizenzierung sind abzuwägen gegenüber den Vorteilen, die im positiven Fall einen gesellschaftlichen Mehrwert für alle Beteiligten erzeugen. Sicherlich wird die 2013 neu gefasste EU-Richtlinie über die Weiterverwendung von Informationen des öffentlichen Sektors, die innerhalb von zwei Jahren in nationales Recht umzusetzen ist, die Entwicklung beflügeln: Erstmals werden nun nämlich auch Bibliotheken, Museen und Archive verpflichtet, die freie (wenn auch nicht kostenfreie) Weiterverwendung ihrer bereitgestellten öffentlichen Daten zu ermöglichen¹⁶.

Adressaten der verlinkten Daten

Wer mit einem gewissen Aufwand Links und weitergehende Informationen zu seinen Erschließungsobjekten im Internet bereitstellt, möchte auch, dass die Angebote genutzt werden. Bisher zeigen sich – neben den nationalen und europäischen Archivportalen – vor allem andere Archiv- oder Kulturinstitutionen daran interessiert, Links zu Verzeichnungsobjekten der regionalen Archivportale wie dem in Hessen zu setzen. Dadurch werden Quellen im Zusammenhang präsentiert oder Recherchemöglichkeiten durch Zusatzinformationen optimiert. Die landesgeschichtliche Anwendung LAGIS beispielsweise setzt Links, um die von den hessischen Staatsarchiven online gestellten digitalisierten Perso-

nenstandsregister gezielter recherchierbar zu machen¹⁷. Mehrfach haben Projektförderungen dazu geführt, dass mittelalterliche Urkunden der hessischen Staatsarchive im Verbund online präsentiert worden sind¹⁸.

In der nächsten Ausbaustufe der Datenverknüpfung könnten Archivdaten selbst als historische Forschungsdaten genutzt werden – für die Geschichtsforschung und für andere Disziplinen. Dazu dürften Erschließungsdaten nicht mehr nur als Verweise auf einzelne Akten und Urkunden verstanden werden, sondern als ein eigener Datenkorpus, der weiter gehende Auswertungen zulässt. Beispielsweise könnte eine systematische Auswertung eines grösseren Korpus von Erschließungsdaten Erkenntnisse darüber erbringen, welche Krankheiten im Zeitverlauf diagnostiziert wurden, welche Bodenschätze wann an welchen Orten gewonnen wurden; welche Personennamen über die Jahrhunderte hinweg verwendet wurden usw. Im Moment ist bei derartigen Forschungsthemen ein direkter Kontakt zum Archiv unumgänglich, und er wird auch in Zukunft sicherlich weiterhin angeraten sein, denn die archivischen Erschließungsdaten sind teilweise erklärungsbedürftig, und aufgrund archivischer Bewertung und Kassation haben statistische Auswertungen ihre Grenzen. Dennoch werden die Möglichkeiten der Forschung deutlich erweitert, wenn Daten in einer strukturierten Form frei zugänglich sind.

Open Government

Über die Onlinebereitstellung von Erschließungsdaten hinaus können staatliche und kommunale Archive – wie alle anderen öffentlichen Verwaltungen auch – ihren Beitrag zu einer offenen und transparenten Verwaltungskultur leisten. Im Sinne des Open Government lassen sich statistische Daten zur archivischen Tätigkeit in einer weiterverarbeitbaren Form publizieren. Bei diesem Aspekt der Datenbereitstellung stehen die meisten deutschen Archive noch am Beginn der Überlegungen. Zu denken wäre etwa an die jährliche Anzahl der Zugänge, der neu verzeichneten Archivalieneinheiten oder der Nutzerinnen und Nutzer im Archiv usw. – kurzum: an alle

Daten, die auch heute schon in Jahres- und Tätigkeitsberichten publiziert werden¹⁹, aber bislang, da sie nur in Papier- oder PDF-Dokumenten enthalten sind, nur schwer automatisiert ausgewertet werden können.

Links zu Ressourcen anderer Anbieter

Ein Kernpostulat der Linked-Open-Data-Idee ist die Verlinkung zu Ressourcen anderer Informationsanbieter, um die Daten zu Elementen des Semantic Web zu machen. Hierbei erweist es sich als Problem, dass die archivischen Erschliessungsdaten bei den meisten Archiven bislang nur bedingt strukturiert oder standardisiert sind. Zwar bieten EAD-Elemente eine Zuordnung, jedoch findet dabei oft kei-

ne Entitätenprüfung statt: Wenn im Titelfeld das Wort «Müller» vorkommt, geht häufig nur aus dem Kontext hervor, ob es sich um den Beruf oder um einen

Eine Anreicherung der Erschliessungsdaten um eindeutige Informationen ist dazu notwendig und auch erstrebenswert, jedoch mit den personellen Kapazitäten der Archive allein kaum zu bewältigen.

Personennamen handelt – und im zweiten Fall weiss man nicht, ob es sich um dieselbe Person namens Müller handelt wie in einer anderen Verzeichnung. Daher sind solche Datei-

en nur eingeschränkt für das Semantic Web geeignet. Eine Anreicherung der Erschliessungsdaten um eindeutige Informationen ist dazu notwendig und auch erstrebenswert, jedoch mit den personellen Kapazitäten der Archive allein kaum zu bewältigen. Ziel muss daher eine automatisierte Anreicherung um Normdaten sein, sodass der Aufwand der Archive sich auf eine Restmenge und auf die Qualitätskontrolle beschränkt. Hier sind Verbünde und Kooperationen unerlässlich. Vordringlich erscheint vor allem die Erstellung oder Optimierung von Personen- und Ortsregistern. Erschliessungsdatensätze sollten dazu mit Datensätzen grösserer Anbieter verlinkt werden. Dabei kommen etablierte Dienste wie GeoNames für ein Ortsregister und DBpedia für ein Personenregister infrage, die auch bei der Europeana hierfür verwendet werden²⁰.

Letztlich gilt es gerade für mittlere und kleinere Archivverwaltungen, sich verbreiteten Lösungen anzuschliessen, die im Verbund genutzt und weiter entwickelt werden. Nur dann kann die grosse Vision der Linked Open Data in der Praxis auch mit Leben gefüllt werden.

Kontakt: Poststelle@hhstaw.hessen.de

ABSTRACT

Linked Open Data dans les systèmes d'information professionnels d'archives

La grande idée des données ouvertes et reliées entre elles (Linked Open Data) intéresse également les archives et leurs systèmes d'information professionnels par le biais d'offres de portails transversaux. Dans le cadre de la rénovation du système d'information archivistique du Land de Hesse, développé conjointement pour les archives de Hesse et de Basse-Saxe, la question de savoir comment concrétiser l'idée des Linked Open Data (LOD) s'est posée dès le début. Les Archives du Land de Hesse permettent depuis des années de relier leurs données via des liens HTTP. L'EAD (encoded archival description; description archivistique encodée) a fait ses preuves comme standard d'échange pour les archives. La bibliothèque numérique européenne Europeana publie également des données archivistiques au format RDF. L'entretien («Curation») des open data représente un véritable défi pour le futur. La question porte également sur le degré d'ouverture des données. Jusqu'ici, ce sont surtout des institutions patrimoniales qui utilisent les liens des systèmes d'information archivistiques. Les données archivistiques devraient également être utilisées à l'avenir pour la recherche, et les archives devraient elles-mêmes se considérer comme des administrations transparentes au sens de l'open government. *(traduction: sg)*

¹⁹ www.hauptstaatsarchiv.hessen.de/irj/HHStAW_Internet?cid=b95c44612cd6101b5b641fbf2f40116b
²⁰ <http://pro.europeana.eu/web/guest/in-brief>

III. Projektberichte aus der Hochschulforschung / *Projets de recherches universitaires*

swissmetrix.ch – neutrale, transparente und vergleichende Webanalyse nicht nur für Bibliotheken

Wolfgang Semar, SII Schweiz. Institut für Informationswissenschaft

Spammails nach dem Motto – «Bringt Ihnen Ihre Internetseite eine befriedigende Anzahl an Besuchern? Wir können Ihnen helfen...» kennen Sie sicherlich. Dabei sieht die Hilfe so aus, dass Metatags in Ihre Website so eingefügt werden, dass Google Ihre Website besser positionieren soll. Aber wer sind eigentlich Ihre Besucher, nach welchen Begriffen haben sie gesucht, woher kommen sie und warum oder wo verlassen sie Ihre Website wieder? Das sind zunächst die entscheidenden Fragen, die geklärt werden müssen, bevor man die eigene Website anpasst.

Um diese und weitere Fragen zu beantworten, bedarf es einer ausführlichen Webanalyse (engl. Web Analytics), mit deren Hilfe das Verhalten der Websitebesucher ausgewertet wird. Web Analytics ist eine kontrollierte, statistische Analyse des Nutzerverhaltens auf Websites, und idealerweise wird mit dieser Methode aufgezeigt, woher die Besucher kommen, welche Bereiche auf der Website aufgesucht wurden und wie oft und wie lange welche Unterseiten und Kategorien angesehen wurden. Jede Website verfolgt ein bestimmtes Ziel, z.B. den Verkauf von Produkten, den Aufbau eines Markenimages oder die Erreichung neuer Zielgruppen. Mithilfe der Webanalyse können Sie überprüfen, ob und wie diese Ziele erreicht wurden. Die analysierten Daten helfen den Webverantwortlichen, die Effektivität des Webauftritts zu verstehen und die Website im Hinblick auf die Zielerreichung zu optimieren. Mit Web Analytics werden nur generelle und anonyme Nutzerdaten untersucht und keine persönlichen oder benutzerindividuellen Aussagen gemacht.

Auswahl eines passenden Instruments

Für die Etablierung eines kontinuierlichen Verbesserungsprozesses ist die Auswahl eines passenden Instruments nur der erste Schritt. Es muss auch festgelegt werden, wer in welchen Zeitabständen welche Kennzahlen analysiert, welche Abweichungen toleriert und welche Schritte bei grösseren Abweichungen eingeleitet werden müssen. Will man eine Newsletter- oder eine Offlinekampagne überwachen, empfiehlt sich die Kommunikation eines individuellen Links, um die Besucher eindeutig dieser Kampagne zuordnen zu können. Häufig werden für solche Zwecke auch spezielle Websites (sogenannte «Landing-Pages») eingerichtet, die zudem den Vorteil haben, dass die Besucher nicht auf der Startseite lange nach den beworbenen Inhalten suchen müssen. Kennzahlen, die in der Anfangsphase von Bedeutung sind, sind beispielsweise die Besucheranzahl, die Konversionsrate, die Abbruchquoten bei bestimmten Seiten und die durchschnittliche Verweildauer auf der Website. Auf Basis dieser Kennzahlen sollte zunächst das Hauptaugenmerk auf die Verbesserung der Nutzerführung gelegt werden. Erst wenn der Web-Controlling-Regelkreis mit diesen Kennzahlen einige Male durchlaufen wurde, sollte der Fokus auf weitere Kennzahlen ausgedehnt werden.

Analyse der Nutzung der Website

Nicht nur Bibliotheken kommen mehr und mehr in einen Nachweiszwang für ihr umfangreiches Dienstleistungs- und Produktangebot. Hier leistet eine ausführliche Analyse der Nutzung der eigenen Website und der Onlineaktivitäten bei der Argumentation eine sehr grosse Hilfe. Allerdings haben nicht alle Bibliotheken das informationstechnische Know-how oder die notwen-

digen Ressourcen, um solche Auswertungssysteme selbst zu betreiben. In solchen Fällen ist es sinnvoll, sich hierfür externe Hilfe zu holen. Bleibt die Frage, von wem und wie eine solche fundierte Analyse vorgenommen wird. Prinzipiell gibt es zwei Möglichkeiten der Datenerfassung und Auswertung.

Da wäre zum einen die Logfile-Analyse. Bei ihr werden die Daten aufgrund der vom Server permanent protokollierten Zugriffsstatistik, meist in einem Monatsrhythmus, analysiert. Zur Analyse stehen den Serverbetreibern umfangreiche Tools zur Verfü-

Die Auswertung der angefallenen Logfiles kann Anhaltspunkte dafür liefern, wie häufig eine Website aufgerufen oder ein PDF-Dokument heruntergeladen wurde.

gung. Ein Vorteil der Logfile-Analyse besteht darin, dass keine Änderungen an der Website erforderlich sind, da die Logfiles vom Server automatisch erzeugt werden. Die Auswertung der angefallenen Logfiles kann Anhaltspunkte dafür liefern, wie häufig eine Website aufgerufen oder ein PDF-Dokument heruntergeladen wurde. Zudem stehen für die Logfile-Analyse kostenlose Instrumente zur Verfügung. Die Nachteile bestehen in einem hohen Aufwand und der Ungenauigkeit des Verfahrens bei der Ermittlung von Besucherzahlen. Auch ist die Logfile-Analyse nicht dazu geeignet, das Nutzerverhalten auf der Website zu rekonstruieren. Zudem ist der Websiteinhaber meist gar nicht im Besitz der Logfiles. Man muss sie vom Administrator anfordern und selbst auswerten. Dieses Verfahren ist daher nur für Unternehmen geeignet, die eine eigene IT haben.

Bei der zweiten Methode wird auf jeder HTML-Seite ein kleiner Code eingeführt, der die Aktivitäten der Besucher zu einem zentralen Server schickt. Dort kann man über ein Webinterface auf die Daten zugreifen und eine eigene Auswertung und Analyse kann vorgenommen werden. Das bekannteste Werkzeug hierfür ist Google-Analytics. Der Nachteil bei Google-Analytics ist, dass zum einen nicht bekannt ist, was mit den Daten passiert, und zum anderen keine eigenen Auswertungswünsche angebracht werden können.

Diesen Nachteil behebt nun das Schweizerische Institut für Informationswissenschaft SII an der HTW Chur. Das SII betreibt seit einiger Zeit eine eigene Anwendung zur Webanalyse mit dem Namen swissmetrix.ch. Der Vorteil gegenüber Google als Anbieter eines eigenen Webanalysetools liegt auf der Hand: Das SII ist ein neutraler, schweizerischer Anbieter. Hierdurch ist gewährleistet, dass die gesammelten Daten dem schweizerischen Datenschutzrecht unterstellt sind und somit auch nicht in fremde Hände fallen können. Gerade bei Google gibt es immer wieder Diskussionen darüber, welchen Umgang das Unternehmen mit gesammelten Daten pflegt.

Ein flexibles Webanalyzesystem

Swissmetrix.ch bietet ein umfangreiches und flexibles Webanalyzesystem. Der Websitebetreiber installiert auf seiner Website einen HTML-Code, der alle notwendigen Daten sammelt und an den Server des SII weiterleitet. Mithilfe des am SII ständig weiterentwickelten Open-Source-Webanalysetools erhalten Websitebetreiber eine eigene Analyse, die individuell und exakt auf ihre Ansprüche zugeschnitten ist. Dabei ist die Bedienung relativ einfach. Derzeit werden die Analysen monatlich vom SII zentral erstellt und den Verantwortlichen zugeschickt. In der Zukunft wird es möglich sein, dass jeder Websitebetreiber seine eigene Analyse selbst zusammenstellt und seine individuellen Ergebnisse selbstständig (falls gewünscht) interpretieren kann. Die Oberfläche von swissmetrix.ch ermöglicht durch ihre nahezu unbegrenzten Filter- und Auswertungsmöglichkeiten tiefe Einblicke und Erkenntnisse in die Webanalyse. Und sollte

eine Analysemethode vermisst werden, so wird diese ganz nach Bedarf entwickelt. Sämtliche Analysen erfolgen in Echtzeit für frei definierbare Zeiträume und über alle erhobenen Daten hinweg.

Wichtige Kennzahlen (Key Performance Indicators; KPI) gewinnen erst durch einen Vergleich mit vorangegangenen oder branchenspezifischen Werten an Aussagekraft. Sie sind innerhalb eines sinnvollen Zeitraumes zu betrachten, um eine Veränderung als Trend und nicht als gewöhnliche, alltägliche Schwankung ansehen zu können.

Die Oberfläche von swissmetrix.ch ermöglicht durch ihre nahezu unbegrenzten Filter- und Auswertungsmöglichkeiten tiefe Einblicke und Erkenntnisse in die Webanalyse.

Hier wird eine regelmässige Analyse empfohlen, um Massnahmen gegen Kennzahlen mit negativer Tendenz rechtzeitig einleiten zu können. Neben einfachen Besucheranalysen bietet swissmetrix.ch weitere statistische Auswertungen, z. B. Traffic-Analysen (PI, Visits, Visitors) im Zeitverlauf, Geoanalysen (Land, Region, Stadt, Organisation, Provider, Browsersprache), Technikanalysen (z.B. Browser, Betriebssysteme, Cookies, Java, Internetgeschwindigkeit), Mobile-Analysen (Endgeräte, Betriebssystem, Steuerungsmethoden, Tastatur, Verbindungstypen), Grafik-

analysen (Bildschirm- und Browserauflösung, Farbtiefe) oder die Überwachung bestimmter Sozialer Medien (Tweets, auf Facebook Shares, Likes oder Kommentare, die auf die eigene Website verweisen).

Onlinemarketing-Analyse

Für viele Unternehmen ist es schwierig, eine eigene Onlinemarketing-Analyse zu erstellen und diese passend auszuwerten. Hier bietet swissmetrix.ch ebenfalls Unterstützung an. So können z.B. eigene Onlinekampagnen geplant und deren Erfolg mithilfe des Kampagnencontrollings (Affiliate, Banner, Newsletter, Keyword) überprüft werden. In Zukunft wird das swissmetrix-Tool so weiterentwickelt, dass Google-Analytics sowie die Facebook-Insite-Statistik integriert werden können. Sie haben dann in einem Tool alle wichtigen Kennzahlen zusammengefasst. Somit erhalten Sie ein umfangreiches Analyseinstrument zur effizienten Überprüfung der Leistungsfähigkeit Ihrer eigenen Websites.

Was müssen Sie tun, wenn Sie an diesem Dienst teilnehmen wollen? Derzeit ist der Dienst für Bibliotheken noch kostenlos, da swissmetrix.ch von e-lib.ch unterstützt wird. Falls Sie Interesse an einer Teilnahme haben, nehmen Sie über swissmetrix.ch Kontakt mit den Verantwortlichen auf, und lassen Sie sich ganz individuell beraten.

Kontakt: wolfgang.semar@htwchur.ch

ABSTRACT

swissmetrix.ch – une analyse du Web neutre et transparente, et pas seulement pour les bibliothèques

De nombreux opérateurs de sites web essaient d'accroître le succès de leur plateforme par le biais de campagnes et d'autres mesures visant à encourager les usagers à les consulter. Une telle mesure ne doit toutefois pas nécessairement être considérée comme une priorité. Elle ne permet en effet de toucher qu'un petit pourcentage de visiteurs potentiels et l'objectif final, par exemple vendre un produit, ne saurait être atteint de manière totalement satisfaisante par ce biais. C'est ici qu'intervient l'analyse de l'audience d'un site web ou «Web Analytics». Mis en œuvre de manière cohérente et suivie, cette méthode de mesure est une garantie du succès des sites web. Le SII exploite depuis quelques mois l'outil d'analyse de l'audience «swissmetrix.ch».

(traduction: sg)

Wissensvorsprung statt Datenflut mittels Web Monitoring und Web Intelligence

Anton Weichselbraun, SII Schweiz.
Institut für Informationswissenschaft

Im Rahmen des WISDOM (Web Intelligence for Improved Decision Making) Projekts werden an der HTW Chur Methoden zur intelligenten und automatischen Extraktion von entscheidungsrelevanter Information aus Onlinequellen entwickelt. Der folgende Artikel gibt einen Überblick über die Entwicklung von Web Monitoring und Web Intelligence und die potenziellen Einsatzmöglichkeiten dieser Technologien bei der Optimierung von Entscheidungsprozessen und Ressourcenallokationen in Organisationen.

Unter Web Monitoring versteht man die gezielte Überwachung von Onlinequellen, um aus der Häufigkeit und Art der Berichterstattung Rückschlüsse auf die

Von Web Intelligence spricht man hingegen, wenn die primäre Zielsetzung auf der Optimierung von Entscheidungsprozessen anhand des online verfügbaren Datenmaterials liegt.

Performance von Unternehmen, Personen oder Produkten ziehen zu können. Von Web Intelligence spricht man hingegen, wenn die primäre Zielsetzung auf der Optimierung von Entschei-

dungsprozessen anhand des online verfügbaren Datenmaterials liegt.

Der Begriff der «Web Intelligence» wurde von Zhong u.a. im Jahre 2000 geprägt¹. Dieses neue Forschungsgebiet beschäftigt sich damit, wie man das im Web und in sozialen Medien verfügbare Datenmaterial auszuwerten und für die Entscheidungsfindung in Unternehmen und Organisationen nutzt.

Der Ausgangspunkt – Business Intelligence

Vorbild dafür war ein wesentlich älteres Teilgebiet der Informatik – nämlich das Gebiet der «Business Intelligence». Ursprünglich wurde der Begriff der «Business Intelligence» von Luhn² verwendet, 1990 jedoch von Howard Dressner wieder aufgegriffen und im Folgenden von Beratungsfirmen wie zum Beispiel Gartner verwendet, um Technologien zu beschreiben, welche auf strukturierte Unternehmensdaten zugreifen, um Prozesse im Unternehmen, und somit in weiterer Folge deren Profitabilität, zu optimieren.

Eines der bekanntesten Beispiele für den erfolgreichen Einsatz von Business Intelligence stammt aus dem im Jahr 2004 von Anderson und Lehman veröffentlichte Artikel (Management Information Science Quarterly), welcher beschreibt, wie Continental Airlines durch den Einsatz von Business Intelligence die Kundenzufriedenheit und Profitabilität ihres Angebotes durch organisatorische Verbesserungen, welche basierend auf internen Unternehmensdaten umgesetzt wurden, erheblich steigern konnten³.

– Continental nutzte historische Flugpass- und Umsatzdaten, um dynamische Auslastungsprognosen zu erstellen und somit besonders gefragte und profitable Strecken an den betroffenen Tagen verstärkt anzubieten.

– Anhand von automatischen Analysen erstellte die Fluglinie Kundenprofile, mit denen die wichtigsten Kunden identifiziert wurden. Im Falle von Verspätungen wurden diesen eine schriftliche Entschuldigung und in manchen Fällen auch eine kostenlose

Probemitgliedschaft im «President Club» von Continental angeboten. In Folge buchte diese besonders profitable Kundengruppe in den kommenden 12 Monaten um 8% mehr Flüge als in den Jahren zuvor.

Neben diesen beiden Massnahmen wurden noch weitere Optimierungen in den Kundenbeziehungen und im Angebot von Continental vorgenommen, welche in Summe zu einem Umsatzplus von mehr als 13 Millionen US-\$, zu einer besseren Abdeckung der Kundenbedürfnisse und somit zu zufriedeneren Kunden führten.

Der nächste Schritt – Web Intelligence

Web Intelligence, das von einigen Autoren auch gerne als Business Intelligence 2.0 bezeichnet wird⁴, erweitert die Methoden der Business Intelligence auf das Web und auf soziale Medien. Im Gegensatz zu Unternehmensdatenbanken, Bibliotheken oder den internen Archiven grosser Organisationen liegt Information hier fragmentiert – spricht über verschiedenste Quellen und Systeme verteilt – vor. Das erschwert die automatische Extraktion von Wissen. Noch komplizierter wird dies durch den Siegeszug von sozialen Medien. Der Kontext einer Nachricht ist dort ungleich schwerer erkennbar und die Verwendung von Slang, Dialekten oder Begriffen aus anderen Sprachen zwingen existierende Analysetools oft in die Knie.

Daher werden in der Praxis Technologien aus den Bereichen des Natural Language Processings und der künstlichen Intelligenz mit semantischen Technologien und sprachspezifischen Heuristiken kombiniert, um zu besseren Ergebnissen zu kommen.

Ein weiteres Problem sind die immensen Datenmengen, welche für solche Analysen relevant sein können. Diese erfordern den Einsatz von «Big Data»-Technologien, um die verfügbaren Datenvolumina in einer vernünftigen Zeit abarbeiten zu können. Auch sind die erhaltenen Analyseergebnisse oft

1 N. Zhong, J. Liu, Y. Y. Yao, and S. Ohsuga, «Web Intelligence (WI)», in Proceedings of the 24th Annual International Computer Software and Applications Conference (COMPSAC 2000), pp. 469–470, 2000.

2 H. P. Luhn, «A business intelligence system», IBM J. Res. Dev., vol. 2, no. 4, pp. 314–319, 1958.

3 R. Anderson-Lehman, H. J. Watson, B. H. Wixom, and J. A. Hoffer, «Continental Airlines Flies High with Real-time Business Intelligence» MIS Q. Exec., vol. 3, no. 4, pp. 163–176, 2004.

4 H. Chen, R. H. L. Chiang, and V. C. Storey, «Business Intelligence and Analytics: From Big Data to Big Impac.», MIS Q., vol. 36, no. 4, pp. 1165–1188, 2012.

mit einer gewissen Schwankungsbreite versehen, da diese auf aggregierten Meinungen, Einzelaussagen und der Medienberichterstattung beruhen. Diese Einschränkungen müssen von Analysten hinreichend berücksichtigt werden, um die mannigfaltigen Analyse-möglichkeiten seriös zu nutzen und um die Ergebnisse korrekt zu interpretieren. Hierfür ist ein umfangreiches Methodenwissen und eine fundierte informationswissenschaftliche Ausbildung unerlässlich. In der englischsprachigen Literatur wird das entsprechende Berufsbild gerne unter dem Begriff «Big Data Scientists» subsumiert.

Bereits jetzt setzen viele Organisationen Web Intelligence ein, um ihre öffentliche Reputation, Marketingkampagnen und Sponsoringaktivitäten zu überwachen⁵. Die folgenden Beispiele sollen einen Eindruck von weiteren innovativen Anwendungsmöglichkeiten dieser Technologien vermitteln.

Kundeninduzierte Produktverbesserungen

Viele Unternehmen setzen auf Expertenpanels, Fragebogenanalysen und die Innovationskraft ihrer Mitarbeiter, um ihre Dienstleistungen und Produkte weiterzuentwickeln.

Mittels Web Intelligence besteht die Möglichkeit, aktuelle Markttrends und Kundenbedürfnisse in diesen Prozess zu integrieren. Trendanalysen erlauben es, Rückschlüsse auf geschäftsrelevante Themen wie zum Beispiel Trendsportarten (Tourismus), neue Smartphones (Elektronik, Handel) oder Gesundheitstrends (pharmazeutische Industrie) zu ziehen und diese zu bewerten. Zudem lässt sich feststellen, ob Themen beim Kunden

positive oder negative Assoziationen wecken und welche Produkt-beziehungsweise Serviceeigenschaften für diese relevant sind. So kann man zum Beispiel automatisch erkennen, dass Kunden mit einem bestimmten Mobiltelefon Eigenschaften wie «hochwertig», «leicht bedienbar» und «schönes Design» verbinden, aber auch, dass die Konkurrenz als innovativer wahrgenommen wird und dass für die Benutzer eine ausgezeichnete Usability oft wichtiger ist als so manches technische Detail.

Noch effektiver sind Techniken, in denen man Web Intelligence einsetzt, um anhand von Kundenfeedback zu eigenen und konkurrierenden Produkten gezielt jene Produktverbesserungen zu identifizieren, welche die Attraktivität des eigenen Produktportfolios besonders steigern. Diese Technik minimiert teure Fehlinvestitionen in Produkteigenschaften, die für den Kunden gar nicht oder nur wenig relevant sind, und erlaubt es, den Prozess der Produktentwicklung optimal an die Kundenbedürfnisse anzupassen.

Web Intelligence für Wirtschaftsinformationen

Das von der HTW Chur gemeinsam mit der Orell Füssli Wirtschaftsinformationen AG entwickelte WISDOM-Projekt beschäftigt sich damit, wie man Web Intelligence für die Extraktion von entscheidungsrelevanten Wirtschaftsinformationen nutzen kann.

So werden im Projekt automatisch Beziehungsnetzwerke zwischen Unternehmen und Führungspersonen extrahiert, anhand welcher sich simulieren lässt, wie sich die wirtschaftlichen Stärken und Schwächen einzelner Partner auf andere Unternehmen auswirken.

Weiterführende Information

- WISDOM-Projekt: www.htwchur.ch/wisdom
- Schweizerisches Institut für Informationswissenschaft (SII): www.informationswissenschaft.ch

Durch die automatische Erkennung von semantischen Zusammenhängen und Assoziationen wird es möglich, Trends zu antizipieren und Wahrnehmungsprofile für Personen, Produkte und Unternehmen zu ermitteln.

Im Rahmen von WISDOM wird auch die Interaktion zwischen wirtschaftlich relevanten Kennzahlen und den extrahierten Informationen untersucht. Dadurch wird es möglich, den Wert dieser Informationen für die Entscheidungsfindung abzuschätzen und verlässliche Modelle zu entwickeln, welche eine optimale Zuweisung von Unternehmensressourcen und eine durch entsprechendes Datenmaterial unterstützte Entscheidungsfindung im Unternehmen ermöglichen.

Die Entwicklung der dem WISDOM-Projekt zugrunde liegenden Infrastruktur, mit Technologien aus den Bereichen Natural Language Processing, künstliche Intelligenz und semantische Technologien, erfolgt durch die HTW Chur in Kooperation mit einem internationalen Forschungsteam von der MODUL University Vienna und der Wirtschaftsuniversität Wien. Die entsprechenden Tätigkeiten werden u.a. von der Europäischen Union im Rahmen des 7. Rahmenprogrammes, der Kommission für Technologie und Innovation (KTI) und dem österreichischen FIT-IT-Programm gefördert. Bei der Weiterentwicklung und Anwendung dieses Technologieportfolios auf wirtschaftliche Fragestellungen spielt die HTW Chur innerhalb des Teams eine Vorreiterrolle.

Kontakt: anton.wechselbraun@htwchur.ch

ABSTRACT

Connaissances vs flux de données: les atouts du Web monitoring et du Web intelligence
Des méthodes d'extraction intelligente et automatique d'informations pertinentes pour la prise de décision à partir de sources en ligne sont développées actuellement à la HTW Coire dans le cadre des projets WISDOM (Web Intelligence for Improved Decision Making). Cet article donne un aperçu du développement du monitoring de sites web (Web monitoring) et du Web intelligence ainsi que les possibilités d'utilisation de ces technologies pour optimiser les processus de décision et les allocations de ressources dans les organisations. *(traduction: sg)*

5 A. Scharl, A. Hubmann-Haidvogel, M. Sabou, A. Wechselbraun, and H.-P. Lang, «From Web Intelligence to Knowledge Co-Creation: A Platform for Analyzing and Supporting Stakeholder Communication», IEEE Internet Computing, vol. 17, no. 5, pp. 21–29, 2013.

IV. Rezensionen / *Recensions*

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

Stephan Holländer

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

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

Gerade diesen beiden erfahrenen Autoren und Referenten geschieht im Eröffnungskapitel ein Fehler, der mit der Thematik sehr vertrauten Autoren öfters passiert. Sie fragen sich nicht, wo man die Einsteiger ins Thema abholen sollte und legen gleich voller Elan mit der Schilderung der Entwicklungsgeschichte von Linked Open Data (LOD) seit 2007 los. Und die gerät sehr ausführlich, ersetzt aber die Erklärung der Grundlagen nicht. Wo findet man eine Erklärung des LOD-Konzepts und wo zu Graphen und Triples in diesem Zusammenhang dienen? Suchen im Ab-

schnitt «Grundlagen»? Fehlanzeige! Die Autoren setzen das unbewusst als bekannt voraus. Hingegen wird die Entwicklungsgeschichte des LOD im angelsächsischen und deutschsprachigen Raum in allen seinen Verästelungen mit Schneid und Akribie aufgezeigt. Gerade für das Verständnis der Ausführungen im später folgenden Kapitel von Pascal Christoph sind diese einführenden Erklärungen unerlässlich. Irgendwann scheint den Autoren aber die Geduld ausgegangen zu sein, denn die Übersicht zu den Entwicklungen zwischen 2007 und 2012 kommen als schnöde Tabellenübersicht daher. So einfach kann man es sich machen. Hier hätte ein Fachlektorat eingreifen sollen, das für eine ausbalanciertere Darstellung der Gesamthematik hätte besorgt sein müssen.

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

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

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

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

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

dass die Lizenzierungsfrage bei der Entscheidung, welchen Ansatz die Bibliotheken wählen können, eine zentrale Rolle spielt, da sie bei der Nutzung von Fremddaten von den vertraglichen Nutzungsbedingungen abhängig sind. Ein guter Beitrag, der aber Vorkenntnisse in der Computerlinguistik voraussetzt.

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

Umwandlung von bibliografischen Daten ins RDF schwierig, da noch keine standardisierten Konkordanztabellen zwischen den Bibliotheksformaten und dem RDF geschaffen wurden.

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

Big Data für Anfänger – ein Reiseführer für «Nichtinformatiker»

Eliane Blumer

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

Das Buch «Big Data. Das Ende des Zufalls» ist der Blick eines «alten Hasen» aus der Kommunikations- und Medienwelt, der es sich zur Aufgabe gemacht hat, tiefer in die Big-Data-Welt einzutau-

chen und seine Entdeckungsreise mittels eines Buches festzuhalten. Dabei will er seiner Auffassung, dass Big Data eines der Jahrhundertthemen ist, jedoch noch nicht die nötige Aufmerksamkeit erhalten hat, Nachdruck verleihen.

Die Reise von Rudi Klausnitzer beginnt ganz vorne – nämlich bei der Fragestellung, warum denn Big Data? Hier seine Erklärung zusammengefasst: Weil der Mensch, neurologisch bewiesen, den

Zufall nicht mag und ihn deshalb so oft als möglich vermeiden will. Und Big Data kann uns dabei helfen, denn durch die maschinelle Kontextualisierung der richtigen Datensets können Vorhersagen getroffen werden, die denn Zufall quasi ausschalten. So kann beispielsweise die Verlinkung von Fahndungsdaten mit bereits existierenden kriminellen Mustern dazu dienen, mögliche Verbrechensorte vorauszusagen und den effizienten Einsatz von Polizisten

bedeuten. Selbstverständlich war dies schon vor der Erfindung von Semantic-Web-Technologien möglich, bedurfte jedoch mehr Zeit und nur menschlicher Denkarbeit. Selbstverständlich wird dies auch nach dem ersten Rush rund um Big und Linked Data möglich sein und bedarf weiterhin menschlicher Denkarbeit, welche einfach durch den Computer unterstützt wird.

Ob «Datenabfall», «Data Scientist», «Big Data», «Dark Data» oder gar «Smart Living», es ist sicher, dass sich Rudi Klausnitzer sehr gut ins Thema der «Grossen Daten» eingearbeitet hat und mit den Begriffen umzugehen weiss. Treffsicher führt er diese im Verlauf der Lektüre ein, erklärt beispielsweise, dass unter «Smart Living» unter anderem intelligente Wohnungen verstanden werden, die den Energieverbrauch messen, und verwendet sie weiter bis zur letzten Zeile. Sein Buch liest sich einfach, gibt einen schnellen Einblick ins Thema der grossen Daten und ist vor allem gespickt mit veranschaulichenden Beispielen, die sowohl die negativen wie auch positiven Auswirkungen oder einfach die Einsatzmöglichkeiten von Big Data aufzeigen. Dabei scheint Klausnitzer Unvoreingenommenheit wichtig zu sein, was durch die jeweils gleiche Anzahl positiver wie negativer Beispiele sichtbar wird. Die Beispiele holt er beispielsweise aus dem Gesundheitswesen mit der Predictive Medicine oder aber dem Marketing, wo Einkäuferdaten analysiert und dementsprechend auf den Einkäufer zugeschnittene Werbeprodukte angepriesen werden.

Der etwas ungeduldige Leser wartet aber bis in die Hälfte des Buches, um eine Definition von Big Data zu erhalten. Einmal bis dorthin gelangt, wird ihm dann aufgezeigt, wo heutzutage überall Daten generiert werden – durch Smartphone, Auto, Haushaltsgeräte oder Applikationen wie «GoogleGlass» – und wie diese durch Algorithmen in Kontext gebracht, zu Smart Data werden können. Schade ist dabei, dass dem Leser bis ans Ende der Algorithmus als nur von «Informatikern», «Mathematikern» und «Statistikern» zu verstehende Rechnungsformel dargestellt wird, die die Intelligenz in die Daten bringt, jedoch nie ein eigentlicher Algo-

rithmus im Buch dargestellt und ausführlicher erklärt wird. Am Ende hat es den Anschein, als ob die Algorithmen die Lösung für alles seien, jedoch nur oben genannte «Auserwählte» ihren wirklichen Hintergrund verstehen können.

Im populärwissenschaftlichen Stil verfasst, soll das Buch ein möglichst breites Publikum ansprechen, denn, wie Klausnitzer selbst im Buch erwähnt: «Bei Big Data geht es um das reale Leben. Da geht es um uns.» Die schnelle Entwicklung der Technologie führt dazu, dass so bald als möglich Fragen zu Datenschutz, Datenerbe oder auch zu Datenkriegen gestellt werden müssen. Klausnitzer betont, dass eine Diskussion bereits stattfindet, dass sich aber jeder darin einbringen sollte, auch um den eigenen Standpunkt zu klären. Dies entspricht der Positionierungsdiskussion der Bibliotheken gegenüber «der Datenfrage» und ist deshalb auch richtig zu erwähnen.

Klausnitzer kommt, wie andere auch, zum Schluss, dass die neuen Trendbegriffe «Data Scientist», «Data Designer», «Data Journalisten» oder gar «Data-tainer» sein werden. Kurz gesagt, Fachkräfte, die mit den Daten umgehen können, sie analysieren, visualisieren, beschreiben beziehungsweise spielerisch weiterverwenden. Dies muss aber nicht heissen, dass nur noch solche Leute ausgebildet werden sollen, sondern zeigt eine neue Entwicklung der Professionen auf.

In einer leicht an Science Fiction anmutenden Zukunftsvision zeigt Klausnitzer auf, dass wir uns auf eine hybride Zeit hinbewegen, in welcher alltägliche Geräte wie Smartphones aber auch der Mensch noch mehr miteinander vernetzt werden. Dabei gibt es drei Lebewesen, «Hybrids», «Algos» und «MTurcs». Erstere sind eine Mischung aus Mensch und Maschine, zweitere sind die Computer und letztere sind diejenigen, die den Maschinen zuarbeiten und von ihnen kontrolliert werden.

Es kann gesagt werden, dass das Buch von Klausnitzer als Reiseführer durch die Welt von Big Data taugt und die aktuellen Beispiele von bereits existierenden Big-Data-Anwendungen aus der

Praxis auch im etwas erfahreneren Leser Neugier für mehr wecken. Dennoch mangelt es ihm manchmal an weiterführenden Definitionen oder tiefergehenden, auch vom Autor selbst vertretenen, Ansichten und Auseinandersetzungen mit der kritischen Thematik. Man merkt bei der gesamten Lektüre den Medienprofi, der weiss, wie er Neuigkeiten an den Mann und die Frau bringen muss, sich selbst und seine Meinung dabei so weit als möglich zurückhält und welche Beispiele dabei möglichst viele Menschen ansprechen, wie das Predictive Policing oder vermehrt auf Social Media ausgetragene Wahlkämpfe. Vergessen geht bei der Aneinanderreihung dieser Beispiele leider etwas, die verschiedenen Standpunkte innerhalb einer solchen Bewegung anzusprechen, damit die Beteiligung der «Nichtinformatiker», «Nichtstatistiker» und «Nichtmathematiker», wie es ja vom Autor auch gewünscht ist, wirklich erreicht werden kann. Denn, wie es im Buch zu lesen ist, ist man sich in vielen Bereichen des Potenzials von Big Data bereits bewusst und agiert dementsprechend, es bedarf deshalb sicherlich detaillierterem Hintergrundwissen und der Bereitschaft, um einen Beitrag dazu machen zu können. Abschliessend kann gesagt werden, dass sich das Buch für alle empfiehlt, die schnell einen weiten, aber an der Oberfläche bleibenden Überblick zu Big Data haben möchten. Wer sich, gerade auch an den informationswissenschaftlichen Bereich denkend, vertieft mit den Technologien hinter Big Data oder den eigentlichen Anwendungssprachen auseinandersetzen möchte, für den hat das Buch möglicherweise etwas zu wenig vertiefenden Inhalt.



Eine interessante Einführung zu einer Vielzahl von Themen der Informationswissenschaft

Stephan Holländer

Wolfgang G. Stock, Mechthild Stock, Handbook on Information Science, De Gruyter, Berlin/Boston 2013. Preis € 149.95 bzw. \$ 210.–

Im deutschsprachigen Raum gehört es zum guten Ton, bei Vorlesungen zur Einführung in die Informationswissenschaft auf eigene Skripte zu verweisen oder aber fallweise auf einzelne Kapitel in verschiedenen Fachbüchern hinzuweisen. In der Vorlesung wird dann gerne unterstützend noch auf Fachartikel und auf Beiträge auf Fachtagungen hingewiesen, um der Aktualität der Themen Rechnung zu tragen. Abgesehen von Einführungskapiteln im alten LaiLuMU und neuerdings in Kuhlen/Semar/Strauch, in neuer Ausgabe dieses Jahr erschienen, sind umfassende Handbücher zu dieser Thematik rar gesät.

Aus dem Hause Mechthild und Wolfgang Stock kommt nun ein Handbuch der Informationswissenschaft in englischer Sprache, das auf dem Umschlag mit dem Stichwort «Reference» wirbt. Ein hoher Anspruch in einem kompetitiven Umfeld, da in den Vereinigten Staaten und in Großbritannien einige Publikationen zur gleichen Thematik bereits erschienen sind.

Die Informationswissenschaft als Querschnittswissenschaft lässt je nach Standpunkt und Ausrichtung der jeweiligen Autoren unterschiedliche Darstellungsweisen des Betrachtungsgegenstands zu.

Mechthild und Wolfgang Stock befassen sich im ersten Kapitel ihres Buches mit der Informationswissenschaft als Gegenstand und Wissenschaftsdisziplin. Es werden Definitionen der Informationswissenschaft und eine Übersicht über ihre Nebendisziplinen vorgelegt. Sehr verkürzt wird die historische Entwicklung der Information beleuchtet. Diese Darstellung ist zu kurz ausgefallen und einige wichtige Meilensteine und Persönlichkeiten fehlen. Da die Geschichte der Informa-

tionswissenschaft aber nicht Hauptgegenstand des Buches ist, mag dies angehen.

Das Einleitungskapitel wartet auch mit einer Auslegeordnung dessen auf, was Informationswissenschaft beinhaltet. In ihrer konzisen Darstellung weiss diese zu gefallen. Man hält sich an Bewährtes und Akzeptiertes, wie es einem Handbuch wohl ansteht. Die Abgrenzung der Informationswissenschaft zur Archivwissenschaft und ihrer Unterdisziplin des Records Managements wird von den Autoren jedoch nicht aufgezeigt.

Ein zweiter Teil des Buches widmet sich dem Information Retrieval. Ein erstes Kapitel widmet sich anekdotenreich der Geschichte des Information Retrievals. Der Abschnitt über die Veränderungen kommerzieller Datenbanken mit Aufkommen des World Wide Webs (WWW) ist etwas kurz ausgefallen und blendet die europäische Entwicklung auf diesem Markt (Datastar, GBI, INKA Star usw.) interessanterweise völlig aus. Gut gelungen ist die Einführung in die Grundlagen des Information Retrievals. Die Einfügung eines Kapitels über Webcrawler ist folgerichtig, da diese Programme im heutigen WWW beim Data Harvesting eine wichtige Rolle spielen. Den Abschluss der Kapitel bildet jeweils die «Conclusion», was dem Buch eine etwas lehrbuchartige Ausrichtung gibt.

Ein dritter Teil beschäftigt sich mit der Semantik und der Analyse der natürlichen Sprache. Die verschiedenen Verfahren und Methoden werden erklärt und mit Beispielen unterlegt. Man hält sich auch in diesem Kapitel an Bewährtes. Ein abschliessendes Kapitel zu Trends und Entwicklungen in diesem Bereich hätte das Buch sehr bereichert.

In einem vierten Teil befassen sich die Autoren mit den Boole'schen Operatoren, deren Vor- und Nachteile beleuchtet werden. Ausgehend von diesen Betrachtungen werden verschiedene Suchstrategien und Methoden der Boole'schen Suche vorgestellt.

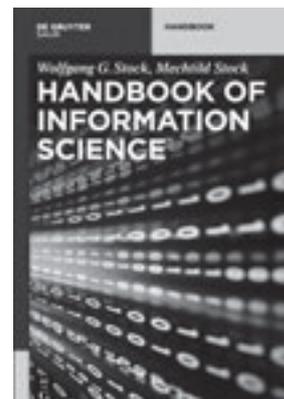
Im fünften Teil des Buches werden die klassischen Retrieval-Modelle sowie neuere Modelle wie beispielsweise das Vektorraum-Retrieval und probabilistische Modelle vorgestellt. Auch das Retrieval von nicht textbasierten Dokumenten wie Bildern, Videos oder Musik werden beleuchtet. Ein Ausblick auf Entwicklungen und Trends beim Retrieval wäre für den interessierten Leser zweifellos ein zusätzlicher Gewinn gewesen.

Im sechsten Teil des Buches beschreiben die Autoren die Suchtechnologien im Web. Die in Suchmaschinen verwendeten Technologien werden erläutert und analysiert. Dieser Teil ist solide geschrieben und stellt die gängigen Methoden vor, die diesen Suchtechnologien zugrunde liegen.

Der siebte Teil befasst sich mit den speziellen Fällen des Information Retrievals. Auch maschinenlesbare Wörterbücher, wie sie in webbasierten Übersetzungsdiensten Verwendung finden, sind dort beschrieben.

Modelle der Leistungsbewertung sind Gegenstand des achten Teils des Buches. Dieser Teil ist sehr knapp gefasst und beschränkt sich auf eine Einführung zu diesem Thema. Weitergehende Informationen sollen wohl der angefügten Bibliografie entnommen werden.

Die daran anschliessenden Teile des Buches sind der Wissensrepräsentation gewidmet. Es werden unter anderem Themen wie «Metadaten»,



«Folksonomies», «Wissensorganisations-systeme», «Indexation» und «Abstracting» angesprochen. Das Buch schließt mit einem Glossar und diversen thematischen Indices.

Das Buch bietet das, was man von einem Handbuch erwarten darf.

Formal mag das Buch in seiner Konzeption zu überzeugen. Dass Information als Gut aus wirtschaftlicher und juristischer Sicht nicht beleuchtet wird, erklärt sich mit der vorausgegangenen Publikation von Wolfgang Stock, die unter anderem diesem Thema gewidmet war.

Aus formaler Sicht überzeugen diesmal auch die dem Text beigefügten bibliografischen Angaben, die einen inhaltlichen Zusammenhang mit dem Text erkennen lassen. Hier kann man im Vergleich zu den letzten Büchern des Autors eine Aufwärtsentwicklung feststellen.

Inhaltlich bietet das Buch das, was man von einem Handbuch erwarten darf. Ein Handbuch soll eine strukturierte Zusammenstellung eines Wissensgebiets, im vorliegenden Fall der Informationswissenschaft, sein und auch als Nachschlagewerk dienen können. Eines zeigt das vorliegende Buch deutlich: Es ist nicht einfach, eine konzise und ausgewogene Übersicht zu allen verschiedenen Aspekten der Infor-

mationswissenschaft in einem vertretbaren Seitenumfang zu publizieren. Gewisse Themen sind ja bereits in anderen Büchern von Wolfgang Stock mit teilweise anderen Co-Autoren erschienen, worauf das Autorengespann im Vorwort hinweist.

Das Verfassen dieses Buches war mit einem grossem Ausarbeitungsaufwand verbunden. Dies wird bei der Lektüre der einzelnen Kapitel deutlich. Einzelne Kapitel stellen in ihrer konzeptionellen Gewichtung die ausdrückliche Meinung der Autoren dar. Dies gilt es zu respektieren. Den Autoren ist es gelungen, eine interessante Mischung von Beiträgen zu den jeweiligen Themenbereichen zu verfassen. Was den Literaturapparat betrifft, so vermisst man den einen oder anderen Hinweis auf englischsprachige Publikationen mit einer ähnlichen Themenstellung wie beispielsweise David G. Luenbergers «Information Science».

Wo bleibt die deutschsprachige Ausgabe?

Für den Diskurs in der Informationswissenschaft in den deutschsprachigen Ländern wäre es wünschenswert, dass das Werk auch in der Muttersprache der Autoren erscheinen würde. Dem Übersetzer, Paul Becker, ist eine gute Über-

tragung in die englische Sprache gelungen, die auch nicht vor Wortneuschöpfungen wie «Aboutness» und «Ofness» zurückschreckt. Ein Glossar wird damit zur unabdingbaren Voraussetzung, um den Kontext der jeweiligen Thematik zu verstehen.

Die Anschaffung des Werkes lohnt sich für alle Dozierenden, die das Fach «Einführung in die Informationswissenschaft» unterrichten, denn das Buch regt zur Auseinandersetzung mit der eigenen Vorlesungskonzeption an. Das Buch stellt eine vielfältige Zusammenstellung zu der grossen Themenbreite dar, die rund um den Begriff Information in der Informationswissenschaft erarbeitet wurden. Dieses Werk ist auch für die Kolleginnen und Kollegen in benachbarten Disziplinen wie der Telekommunikation, der Wirtschaftsinformatik, der Informatik und der Ökonomie von Interesse und kann den interdisziplinären Dialog fördern. Für Studenten wäre die Anschaffung des Buches als Nachschlagewerk von Nutzen, sofern eine preisgünstige Studienausgabe auf den Markt käme. Beim aktuellen Verkaufspreis bleibt den Studierenden lediglich die Hoffnung, dass die jeweiligen Fachbibliotheken sich für eine Anschaffung des Werkes entscheiden.

V. Aktualitäten / *Actualités*

Anna Pia Maissen zum Abschied

Gregor Egloff,
Staatsarchiv Luzern

Anna Pia Maissen, vielsprachige Bündnerin und Stadtarchivarin in Zürich, hat von 2007 bis 2013 den Verband der Schweizerischen Archivarinnen und Archivare VSA als Präsidentin geleitet. Mit Leidenschaft, Einsatz, Engagement. Mit Charme und Biss. Mit Erfolg.

«Nichts reflektiert die Vergangenheit besser als ein Blick in die Zukunft, denn dieser basiert in der Regel auf dem bereits Geleisteten.» Ein selbstcharakterisierender Satz im präsidentalen Editorial im Jahresbericht 2012 des VSA aus der Feder der scheidenden Präsidentin, deren Amtszeit statutengemäss nach sechs Jahren (bzw. nach acht Vorstandsjahren) im letzten September zu Ende gegangen ist.

Schon die Zahlen zeigen, dass der VSA unter Anna Pia Maissen quantitativ und qualitativ gewachsen ist. Mehr institutionelle Mitglieder, und vor allem mehr Einzelmitglieder, sind einer Vereinigung beigetreten, die organisatorisch und finanziell auf einer guten Basis in die Zukunft blicken darf.

Ja, Anna Pia, wir alle wissen, dass du dir diesen Erfolg nicht selber zuschreiben wür-

dest. Ein Team zu motivieren, entspricht deinem Führungsstil, und das machst du ausgezeichnet. Und doch hast du fast unheimlich viele Aufgaben der Verbandsleitung selber übernommen, und dies nicht nur, wenn du als Präsidentin den Verband offiziell vertreten durftest. Du warst dir auch nicht zu schade, viele kleinere Anfragen an den VSA selber zu beantworten, weil dir dieser Kontakt wichtig und deinem Amtsverständnis geschuldet war – von allen anderen Delegationen, Projektarbeiten und Sitzungen in den verschiedensten Gremien ganz zu schweigen.

Während ihrer gesamten Amtszeit hatte Anna Pia Maissen ihren Blick fest auf die Zukunft des VSA gerichtet. Dementsprechend prägte sie die Verbandschwerpunkte in Richtung Professionalisierung: Dazu gehören die Verbandsstrukturen, aber auch die berufliche Aus- und Weiterbildung, wo sich im Verbandsrahmen Angebote von «Profis für Profis» mit grossem Erfolg neu- und weiterentwickelt haben. Gleichzeitig legte sie grossen Wert darauf, dass auf der Plattform des Verbands ein engmaschiges, solidarisches Berufsnetzwerk gepflegt werden kann, das weitere Kooperationen ermöglicht, international, selbstverständlich, aber auch in unserer sehr kleinteiligen

Archivlandschaft. Gerade hier hat sie immer Wert darauf gelegt, dass die hohe Kultur der gegenseitigen kollegialen Unterstützung ebenso gepflegt wird wie das Angebot institutionell angebotener Bildungsangebote.

Die Professionalisierung in einem sich technologisch und gesellschaftlich wandelnden Umfeld führte Anna Pia Maissen zur «Archivierung als Qualitätsmerkmal». Klarer als in ihren eigenen Worten in ihrem letzten Jahresbericht kann man es nicht ausdrücken: «Es ist eine der wichtigsten Aufgaben des VSA sowie der Archivarinnen und Archivare, welche er repräsentiert, Archivierung als klar umrissenen und eindeutigen Arbeitsprozess zu vermitteln und dem beliebigen und willkürlichen Gebrauch dieser Qualitätsmarke entgegenzutreten.»

Der Einsatz für die professionellen Interessen der schweizerischen Archivarinnen und Archivare hatte sich Anna Pia Maissen schon ganz zu Beginn ihrer Amtszeit auf die Fahnen geschrieben. Nicht nur sie hat in den letzten Jahren feststellen müssen, dass «Archivieren» offenbar verbreiteter ist und positiver eingeschätzt wird als auch schon. Nur hat das, was mit diesem Begriff versehen wird, in den Augen pro-

fessioneller Archivarinnen und Archivare allzu oft rein gar nichts mit Archivierung zu tun. Was sie hierzu dem Verband ins Stammbuch schreibt, ist wiederum an Eindeutigkeit kaum zu übertreffen: «Überall, wo unser Kerngeschäft eingeschränkt oder behindert zu werden droht, müssen wir reagieren. Alle Eingriffe, welche unser Berufsethos kompromittieren könnten, müssen abgewehrt werden.»

Neben ihrem grossen persönlichen und professionellen Einsatz in allen Belangen der Verbandsführung ist es ihr grosser Verdienst, die Zeichen des Wandels erkannt und die Weichen in Richtung Zukunft gestellt zu haben. Sie übergibt ihren Nachfolgerinnen und Nachfolgern eine kerngesunde Organisation.

Es war ein grosses Privileg und eine dankbare Aufgabe, mit dir im Verband zu arbeiten (du hast es deinen Kolleginnen und Kollegen mit deiner Art auch leicht gemacht).

Als Mitglieder des VSA werden wir dir noch auf Jahre hinaus dankbar sein für alles, was du geleistet hast.

Für deine eigene Zukunft ist dir archivistisch wie privat nur das Beste zu wünschen!

Eloge de Anna Pia Maissen

Gregor Egloff, Staatsarchiv Luzern

Anna Pia Maissen, Grisonne polyglotte et archiviste de la ville de Zurich, a été présidente de l'Association des Archivistes Suisses de 2007 à 2013. Avec passion et engagement. Avec charme et mordant. Avec succès.

«Rien de reflète mieux le passé qu'un regard vers l'avenir, car celui-ci repose en général sur ce qui a déjà été accompli.» Cette phrase de l'éditorial que Anna Pia Maissen a signé pour le rapport d'activité 2012 de l'AAS traduit très bien son caractère. Conformément aux statuts, elle a quitté sa fonction en septembre dernier, après six ans d'activité (resp. huit ans au sein du Comité).



Les chiffres montrent à eux seuls que l'AAS s'est renforcée sous la présidence de Anna Pia Maissen, et ce, tant du point de vue quantitatif que qualitatif. Plus de

membres institutionnels et, surtout, plus de membres individuels ont rejoint une association dont les bases organisationnelles et financières sont saines.

Oui, Anna Pia, nous savons tous que tu n'attribuerais pas ce succès à toi-même. Motiver une équipe, tel est ton style de gestion, et tu le fais de manière remarquable. Et pourtant, tu as presque assumé toute seule de très nombreuses tâches de l'association, et ce, pas seulement lorsque tu pouvais agir officiellement en tant que présidente. Tu ne dédaignais pas non plus de répondre toi-même à de nombreuses petites demandes adressées à l'AAS, car tu trouvais ce contact important et que c'est ainsi que tu concevais ta fonction, sans parler de toutes les autres délégations, travaux de projet et séances dans des instances les plus diverses.

Pendant toute la période de son mandat, Anna Pia Maissen avait son esprit tourné vers l'avenir de l'AAS. Elle a donc tout particulièrement insisté sur le thème de la professionnalisation: structures associatives, mais aussi formation professionnelle initiale et continue, domaine où des offres «par des pros pour des pros» ont été proposées et développées avec succès dans le cadre de l'AAS. Elle estimait également qu'il était important d'entretenir un réseau professionnel dynamique et solidaire sur la plateforme de l'Association, afin de favoriser les coopérations, au niveau international bien évidemment, mais également au sein de notre petit monde des archives. C'est justement ici qu'elle estimait crucial d'entretenir une culture de soutien collégial réciproque et de proposer des offres de formation à un niveau institutionnel.

La professionnalisation dans un environnement changeant, tant sur le plan

technologique que social, a conduit Anna Pia Maissen à la notion d'«archivage comme critère de qualité». Elle a exprimé elle-même ce point de manière éloquent dans son dernier rapport d'activité: «L'une des principales tâches de l'AAS ainsi que des archivistes qu'elle représente est d'enseigner l'archivage en tant que processus de travail clairement délimité et défini, et de faire face à toute forme d'utilisation de ces marques de qualité.»

Dès le début de sa présidence, Anna Pia Maissen a œuvré en faveur des intérêts professionnels des archivistes suisses. Elle a certes dû constater ces dernières années que «archiver» était devenu plus répandu et perçu de manière plus positive que par le passé. Mais aussi que ce que l'on entend par là n'a souvent rien à voir avec l'archivage tel que le conçoivent les professionnels du domaine. Laissons-lui encore une fois la

parole à ce propos: «Nous devons réagir partout où notre activité clé risque d'être limitée ou empêchée. Toutes les attaques qui pourraient compromettre notre éthique professionnelle doivent être contrecarrées.»

Outre l'engagement personnel et professionnel dans toutes les affaires en rapport avec la direction de l'Association, elle a le mérite d'avoir reconnu les signes du changement et d'avoir posé les jalons pour assurer l'avenir. Elle transmet à ses successeurs une organisation saine.

Ce fut un grand privilège et une tâche gratifiante que de travailler avec toi au sein de l'Association (un travail que tu as également su alléger grâce à ton style). En tant que membres de l'AAS, nous te resterons encore longtemps reconnaissants pour tout ce que tu as accompli.

Nous te souhaitons un avenir fructueux tant sur le plan professionnel que privé!

BiblioFreak – Freaks und Leidenschaften für Bibliotheken

Gerhard W. Matter,
Kantonsbibliothek Baselland

Nach dem Vorbild der erfolgreichen US-Kampagne «geek the library» soll die Kampagne BiblioFreak auch in der Schweiz die Aufmerksamkeit auf die Bibliotheken lenken. Bibliotheken sollen in der Öffentlichkeit besser wahrgenommen werden und mit der Kampagne sich auch neue Nutzerkreise erschliessen können. Das gelingt umso besser, je mehr Bibliotheken mitmachen.

Ermuntert vom Erfolg der Kampagne in den USA lud OCLC Deutschland rund ein Dutzend Bibliotheksleiterinnen und -leiter aus Deutschland, Österreich und der Schweiz (DACH) ein, sich zu überlegen, ob die Kampagne «geek the library» (www.geekthelibrary.org) adaptiert und auch im deutschsprachigen Europa erfolgreich umgesetzt werden kann. Das Ergebnis der intensiven Überlegungen war ein überzeugtes Ja. Das neuartige

Konzept der Kampagne überzeugte. Und: Auch in Europa müssen die öffentliche Wahrnehmung und der Stellenwert von Bibliotheken gestärkt werden.

BiblioFreak kommt in die Schweiz

Das von einer internationalen Werbeagentur entwickelte Design mit Website, Facebookauftritt und Printmassnahmen konnte für Europa adaptiert werden. Momentan werden die Materialien der Kampagne von Pilotbibliotheken getestet. Die Schweizerische Konferenz der Kantonsbibliotheken SKKB will dieses Projekt in der Schweiz umsetzen. Für die Projektkoordination wurde das Aktionskomitee BiblioFreak ins Leben gerufen. Dieses Aktionskomitee wird die nationale Kampagne vorbereiten und koordinieren.

Start von BiblioFreak

Die Kampagne soll in der Schweiz im September 2014 am BIS-Kongress in Lugano starten. Je mehr Bibliotheken

sich mit kreativen Aktionen und Leidenschaft beteiligen, umso grösser wird das Echo in den Medien und der Öffentlichkeit sein. Erstmals werden Bibliotheken nicht nur eines Landes, sondern Bibliotheken von mehreren europäischen Ländern in einer gemeinsamen Kampagne wahrnehmbar. Dies ist eine neue Dimension und für das Ziel, die Wahrnehmung und den Stellenwert von Bibliotheken zu verbessern, sehr wichtig.

Französisch! Italienisch!

Die Kampagne BiblioFreak wurde für den deutschsprachigen Raum entwickelt und wird auch in allen deutschsprachigen Ländern umgesetzt. Mittlerweile hat Holland entschieden, die Kampagne zu übernehmen und auch Polen und Dänemark haben bereits grosses Interesse angemeldet. Abklärungen für eine französische und italienische Version laufen. Dies ist für eine nationale Kampagne in der Schweiz wichtig.

Wie kann sich meine Bibliothek beteiligen?

Für das Gelingen der Kampagne ist es sehr wichtig, dass sich möglichst viele Bibliotheken aus möglichst allen Landes-teilen beteiligen. Neben öffentlichen Bibliotheken jeglicher Grösse, sind auch Schulbibliotheken jeder Stufe – vom Kindergarten bis zur Universität – sowie Spezialbibliotheken angesprochen. Natürlich können die Bibliotheken einer Region oder eines Kantons sich auch zusammen der Kampagne anschliessen und gemeinsame Aktionen durchführen. Beginnen Sie sich frühzeitig mit BiblioFreak zu

beschäftigen. Integrieren Sie BiblioFreak in Ihre Bibliotheksplanung. Planen Sie die nötigen personellen Ressourcen für die Vorbereitung und Durchführung des Kampagnenstarts im September ein. Denken Sie daran, dass BiblioFreak eine mittelfristige Kampagne und nicht eine einmalige Aktion ist. In den USA läuft die Kampagne bereits seit 2004 sehr erfolgreich.

Kosten und nächste Schritte

Die Kampagne BiblioFreak ist für die einzelne Bibliothek kostengünstig. Die Website und Facebookseite BiblioFreak

werden Sie kostenlos nutzen können. Die gedruckten Materialien wie Plakate, Flyer, Buchzeichen, Postkarten, Buttons zum Anstecken, Roll-ups etc. sind preiswert, da sie in grossen Auflagen für die Bibliotheken in Deutschland, Österreich und in der Schweiz hergestellt werden. Zudem werden Druckvorlagen (Templates) all dieser Materialien zum kostenlosen Herunterladen zur Verfügung stehen. Das gesamte Kampagnenmaterial wird über die SBD.bibliotheksservice ag bezogen werden können.

Das Aktionskomitee BiblioFreak wird Sie auf dem Laufenden halten. Mit Ihren Fragen wenden Sie sich bitte direkt an das Mitglied des Aktionskomitees Ihrer Region oder dasjenige, das Ihnen am nächsten steht.

Und: Welcher Freak steckt in dir?

Mitglieder des Aktionskomitees BiblioFreak:

Vertreterinnen und Vertreter der Regionen:

- Sonia Abun-Nasr, KB St. Gallen, und Bernhard Bertelmann, KB Thurgau, für die Ostschweiz
- Hermann Romer, SB Winterthur, und Andrea Malits, ZB Zürich, für die Region Zürich
- Eliane Latzel, KB Uri für die Zentralschweiz
- Niklaus Landolt, ZB Bern, für die Region Bern
- Gerhard W. Matter, KB Baselland, für die Nordwestschweiz
- N.N. für die Romanische Schweiz

Vertreterinnen und Vertreter der nationalen Institutionen/Verbände:

- Hans-Dieter Amstutz, Schweizerische Nationalbibliothek
- Anita Büttiker, SBD.bibliotheksservice ag
- Edith Moser, SAB
- Herbert Staub, Bibliothek Information Schweiz BIS
- Peter Wille, Bibliomedia Schweiz

Weitere Informationen über Biblio-Freak finden Sie unter www.oclc.org/de-DE/about/Initiativen/BiblioFreak.html und www.bibliofreak.org.

BiblioFreak – Freaks et passions pour les bibliothèques

Gerhard W. Matter, Bibliothèque cantonale, Bâle-Campagne

Conçue d'après le modèle de l'action menée avec succès aux Etats-Unis, «geek the library», la campagne BiblioFreak a pour but de promouvoir les bibliothèques suisses auprès du public. Les bibliothèques doivent en effet être mieux perçues par le grand public et la campagne permettre d'élargir le cercle de leurs usagers. Le résultat dépend évidemment du nombre de bibliothèques qui y participent.

Encouragé par le succès rencontré par la campagne américaine, OCLC Deutschland a invité une douzaine de directrices et de directeurs de bibliothèques d'Alle-

magne, d'Autriche et de Suisse (DACH) à étudier la possibilité d'adapter à l'Europe germanophone la campagne «geek the library» (www.geekthelibrary.org). Leur réponse fut un «oui» clair et net. Le concept original de la campagne a convaincu et le besoin de reconnaissance et de valorisation des bibliothèques en Europe considéré comme évident.

BiblioFreak arrive en Suisse

La Conférence suisse des bibliothèques cantonales (CSBC) souhaite mettre en œuvre ce projet en Suisse. Un comité d'action spécifique a été créé afin de coordonner le projet. Ce comité d'action préparera et coordonnera la campagne au niveau national.

Lancement de BiblioFreak

La campagne doit être lancée en Suisse en septembre 2014, à l'occasion du Congrès BIS de Lugano. Plus il y aura de bibliothèques qui s'y engageront avec passion et des actions originales et plus l'écho sera grand dans les médias et le public.

En français! En italien!

La campagne BiblioFreak a été développée dans l'espace germanophone et sera également lancée dans tous les pays de langue allemande. Des clarifications sont en cours concernant une version française et italienne, un must si l'on entend réaliser une campagne nationale digne de ce nom.

Comment ma bibliothèque peut-elle participer?

Le succès de la campagne dépend essentiellement du nombre de bibliothèques qui y participent. Outre les bibliothèques publiques de toute taille, les bibliothèques scolaires de tous les niveaux – du jardin d'enfants à l'université – ainsi que les bibliothèques spécialisées sont sollicitées dans ce contexte. Il va de soi que les bibliothèques d'une région ou d'un canton peuvent également participer ensemble à la campagne et mener des actions communes.

Coûts et prochaines étapes

La campagne BiblioFreak est peu onéreuse pour les bibliothèques. Le site web et la page Facebook BiblioFreak peuvent être utilisés gratuitement. Les imprimés tels que affiches, prospectus, signets, cartes postales, etc. sont bon marché, étant donné qu'ils sont produits en grande quantité pour les bibliothèques, en Allemagne, en Autriche et en Suisse. De plus, les fichiers d'impression (templates) de tous ces matériels peuvent être téléchargés gratuitement.

Le comité d'action BiblioFreak vous informera régulièrement. Veuillez adresser vos questions directement au membre du comité de votre région ou à celui qui vous est le plus proche.

Vous trouverez d'autres informations sur BiblioFreak sur: www.oclc.org/de-DE/about/Initiativen/BiblioFreak.html et www.bibliofreak.org.

BIS-Kongress in Lugano: Attraktives Programm in der Sonnenstube

**Katia Röthlin, Projektmitarbeiterin
Geschäftsstelle BIS**

Unter dem Motto «Bibliotheken & Bildung» findet vom 3. bis 6. September 2014 der Kongress des BIS in Lugano statt. Reservieren Sie sich jetzt schon den Termin.

Nach dem erfolgreichen Kongress 2012 in Konstanz lädt der BIS im nächsten September ins Tessin ein. In der Sonnenstube der Schweiz treffen sich rund 400 Personen aus öffentlichen und wissenschaftlichen Bibliotheken, um sich weiterzubilden und auszutau-

schen. Sie sind herzlich eingeladen, dabei zu sein. Der Palazzo dei Congressi liegt direkt am See und bietet optimale Voraussetzungen für den Kongress.

Bereits haben zahlreiche Firmen Plätze für die Ausstellung reserviert und werden dort das Neueste aus der Bibliothekswelt präsentieren. Neben Besuchen in diversen Bibliotheken rund um Lugano – so zum Beispiel in der Bibliothek des neuen Bundesstrafgerichts in Bellinzona, stehen Fachvorträge auf dem Programm. Auch das gemütliche Zusammensein kommt nicht zu kurz, so am Eröffnungs- und am Festabend

vom Donnerstag: per Schiff fahren Sie in Grotti, die selbst für Tessiner ein Geheimtipp sind. Am Freitagnachmittag hält der BIS seine Generalversammlung ab. Wer ganz entspannt am Kongress ankommen will, der reist mit uns mit dem Sonderzug «Churchill» ins Tessin (siehe Kästchen).

Im «Churchill» an den Kongress

Der nächste BIS-Kongress beginnt schon unterwegs: Reisen Sie mit uns im Sonderzug «Churchill» am 3. September 2014 vormittags von Olten oder Zürich ins Tessin. In Bellinzona besichtigen wir die neue Bibliothek des Bundesstrafgerichtes und fahren von dort weiter nach Lugano. Der rote Doppelpfeil, der anlässlich der Landesausstellung 1939 gebaut wurde, hat 1946 dazu gedient, Winston Churchill als Staatsgast durch die Schweiz zu fahren. Seither ist er bekannt als «Churchill-Pfeil». Jetzt haben Sie die einmalige Gelegenheit, mit diesem besonderen Zug zu reisen und ganz entspannt am Kongress anzukommen. Unterwegs erfahren Sie in kurzen Vorträgen Interessantes über einige Firmen, die am Kongress vertreten sein werden, und können sich mit anderen BIS-Mitgliedern austauschen. Fahrt und Mittagessen sind im Preis von 75 Franken inbegriffen. Melden Sie sich jetzt an, die Platzzahl ist beschränkt: www.bis.ch.



Congrès BIS à Lugano: Un programme attractif dans une région ensoleillée

Katia Röthlin, collaboratrice de projets,
Secrétariat BIS

«Bibliothèques & formation», tel est le thème du prochain congrès BIS qui se tiendra à Lugano du 3 au 6 septembre 2014. Réservez dès maintenant cette date!

Après le congrès 2012 qui s'est déroulé à Constance – une réussite! – BIS invite ses membres au Tessin pour le congrès de septembre 2014. Quelque 400 personnes se retrouveront dans le sud ensoleillé de la Suisse, pour continuer à se former et pour échanger. Vous y êtes cordialement invités. Le Palazzo dei Congressi donne directement sur le lac et offre des conditions optimales pour notre congrès.

De nombreuses entreprises ont d'ores et déjà réservé des places pour l'exposition et présenteront les dernières nou-

veautés du monde des bibliothèques. Outre des visites dans plusieurs bibliothèques de la région de Lugano, comme la bibliothèque du nouveau Tribunal pénal fédéral à Bellinzona, le programme comprendra également des conférences. On ne négligera pas non plus la convivialité, notamment lors de la cérémonie d'ouverture et de la soirée

festive du jeudi: vous vous rendrez à Grotti en bateau, un bon tuyau, même pour les Tessinois. Le vendredi après-midi, BIS tiendra son assemblée générale. Et pour celles et ceux qui veulent se rendre au Tessin en toute décontraction, ils pourront voyager avec nous à bord du train spécial «Churchill» (voir encadré).

Avec le «Churchill»

Le prochain congrès BIS commence déjà «en chemin»: voyagez avec nous à bord du train spécial «Churchill», l'après-midi du 3 septembre 2014, de Olten ou Zurich jusqu'au Tessin. A Bellinzona, nous visiterons la nouvelle bibliothèque du Tribunal pénal fédéral avant de continuer en direction de Lugano. La légendaire flèche rouge, construite à l'occasion de l'exposition nationale de 1939, a notamment servi en 1946 pour la visite de Winston Churchill en Suisse. D'où son surnom. Vous avez maintenant la possibilité unique de voyager à bord de ce train spécial et d'arriver au congrès frais et dispos. En cours de route, vous pourrez écouter de brefs exposés sur certaines entreprises qui seront présentes au congrès et échanger informations et impressions avec d'autres membres BIS. Le trajet et le repas de midi sont proposés au prix de 75 francs. Inscrivez-vous dès maintenant, le nombre de places est limité: www.bis.ch.

Umfrage *arbido*

Herbert Staub, Präsident BIS

Im Rahmen ihrer Bachelorarbeit an der HTW Chur hat die Informationswissenschaftlerin Sarah Carbis im Sommer die BIS-Mitglieder zu *arbido* befragt. 202 Mitglieder aus der Deutschschweiz und 67 aus der Romandie haben teilgenommen. Positiv bewerteten die Befragten den Erscheinungsrhythmus, die Zweisprachigkeit, die Abstracts, die interdisziplinären Themen und die Printausgabe. Veränderungen wurden

gewünscht beim Inhalt: keine monothematischen Hefte, mehr Informationen zur Verbandstätigkeit, nicht nur Fachartikel. Klarer Optimierungsbedarf zeigte sich bei Layout und Lesefreundlichkeit. Der VSA hatte seine Mitglieder bereits 2009 zu *arbido* befragt.

Der BIS-Vorstand bedankt sich bei Sarah Carbis für die äusserst wertvolle Arbeit. An seiner Retraite im Dezember wird der Vorstand die Umfrageergeb-

nisse konkret auswerten und allfällige Optimierung von *Arbido* skizzieren. Anfang 2014 wird mit dem VSA als Mitherausgeber ein möglicher Relaunch diskutiert. Eine Umsetzung gemeinsam beschlossener Änderungen wäre per 2015 realisierbar.

Mehr dazu: Die Bachelorarbeit wird demnächst in der Reihe «Churer Schriften zur Informationswissenschaft» veröffentlicht. www.htwchur.ch/informationswissenschaft

Enquête *arbido*

Herbert Staub, Präsident BIS

Sarah Carbis, spécialiste en information documentaire, a, dans le cadre de son travail de bachelor à la HTW Coire, réalisé auprès des membres BIS une enquête concernant *arbido*, en été 2013. 202 membres de Suisse alémanique et 67 de Suisse romande ont participé à cette enquête. Les personnes interrogées ont considéré positifs le rythme de parution, le bilinguisme, les abstracts, les thèmes interdisciplinaires et l'édi-

tion print. Elles souhaitent également des changements au niveau du contenu: pas de cahier monothématique, plus d'informations sur les activités de l'association, pas uniquement des articles spécialisés. Quant au layout et à la lisibilité, ils doivent être clairement optimisés. L'AAS avait déjà interrogé ses membres en 2009 concernant *arbido*.

Le Comité BIS remercie Sarah Carbis pour son travail extrêmement précieux. Lors de sa retraite de décembre, le Comi-

té évaluera concrètement les résultats de cette enquête et esquissera les éventuelles mesures d'optimisation. Au début 2014, l'AAS, co-éditrice, discutera d'une possible nouvelle version. Une mise en œuvre des changements décidés en commun serait réalisable pour 2015.

Plus d'informations: le travail de bachelor sera publié prochainement dans la série «Churer Schriften zur Informationswissenschaft». www.htwchur.ch/informationswissenschaft.

Antonio Spadafora (1942 – 2013)

Robert Barth

Am 29. Juni ist der ehemalige Direktor der Biblioteca Cantonale von Locarno, Antonio Spadafora, verstorben. In der Deutschschweiz und der Romandie weniger bekannt, gehörte Spadafora im Tessin während vier Jahrzehnten zu den prägenden intellektuellen Persönlichkeiten mit zahlreichen Beiträgen in Presse, Radio und Fernsehen. Die von ihm organisierten Konferenzen «Scienza e Società» führten so berühmte Persönlichkeiten und Nobelpreisträger wie Karl Popper, Rita Levi-Montalcini, Paul Feyerabend oder John Carew Eccles nach Locarno. Die französische Aufklärung (besonders Rousseau und Diderot) und die Zivilgesellschaft (Alexis de Tocquevilles Demokratie in Amerika) waren die Themen, die ihn ein Leben lang beschäftigten.

Wissenschaftliche Arbeit hielt Spadafora aber nicht davon ab, in der Gesellschaft praktisch zu wirken: Er plante die 1989 eröffnete Bibliothek von Locarno im Palazzo Morettini, die Luciano Canfora in seinem Nachruf auf Spadafora



im Mailänder «Corriere della Sera» zu Recht als «Schmuckstück und Modell» bezeichnet hat. Während 20 Jahren hat Spadafora sie auch geleitet. Diese Institution nahm das vorweg, was wir heute mit «Bibliothek als dritter Ort» bezeichnen: ein einladender Ort der Begegnung mit vielfältigem Medienbestand und reichem kulturellem Angebot. Spadafora war auch der Spiritus Rector des bis heute vorbildlichen Tessiner Bibliotheksgesetzes von 1991. Und in den Neunzigerjahren initiierte er die Tessiner Bibliothekarenkurse, die in Zusammenarbeit mit dem BBS durchgeführt wurden. Sie haben eine ganze Generation von Tessiner Bibliothekaren geprägt. Und nicht zuletzt stand er dem Tessiner Bibliotheksverbund vor, dem Systema bibliotecario ticinese.

Auch wenn Spadafora als Liberaler dem Tessiner Freisinn nahestand, so blieb er Italiener. Er stammte aus dem kalabrischen Cosenza, doktorierte an der Universität Pavia, arbeitete kurze Zeit als Lehrer an einem Mailänder Jugendgefängnis und kam Ende der Sechzigerjahre ins Tessin, wo er am Gymnasium und Lehrerseminar von Locarno Philosophie und Pädagogik unterrichtete. Erst anschliessend stiess er als Quereinsteiger zum Bibliothekswesen.

Wer die Bibliotheksgeschichte kennt, kommt nicht um einen Vergleich mit Antonio Panizzi herum: Ebenfalls aus Italien stammend, gelang es diesem, dem britischen Bibliothekswesen wichtige Impulse zu geben. Spadafora seinerseits hat die Organisation, die gesetzlichen Grundlagen und die Ausbildung der Tessiner Bibliotheken massgeblich beeinflusst.

Die letzten anderthalb Jahre mit seiner Krebserkrankung hat Antonio Spadafora mit bemerkenswerter Würde und der ihm eigenen Ironie getragen. Er blieb platonischer Philosoph bis zuletzt.

Zeitschel OS15000
Schöner kann man nicht scannen.
Besser auch nicht.



NEUHEIT!

Zeitschel OS 15000 Advanced Plus
- mit motorischer Buchwippe und selbstöffnender Glasplatte

Zeitschel OS 15000 Advanced
- mit manueller Buchwippe und selbstöffnender Glasplatte

Zeitschel OS 15000 Comfort
- mit Comfort Buchwippe, ohne Glasplatte

Der Zeitschel OS 15000 besticht durch sein aussergewöhnliches Design und liefert exzellente Daten. Er ist dank der bewährten Software Omniscan 12 sehr einfach zu bedienen. Mit einer Auflagefläche von maximal 460x360mm und einer Auflösung von 600 dpi bei 42 Bit Farbtiefe lassen sich auch grossformatige Bücher und Dokumente problemlos digitalisieren.

digitalisieren ocr plotten scannen duplizieren mikrofilm kopieren
scannen vektorisieren ocr digitalisieren archivieren ocr mikrofilm

SUPAG
Informations-Management

Spichtig und Partner AG
Rietstrasse 15, 8108 Dällikon
Tel. 044 844 29 39
mail@supag.ch, www.supag.ch

TRIALOG

*Die Unternehmensberatung
für Bibliothek, Archiv und
Wissensorganisation*

Neue Ideen für neue Herausforderungen.

Wir helfen dabei.

*Trialog AG, Holbeinstr. 34, 8008 Zürich
Tel. 044 261 33 44, Fax 044 261 33 77
trialog@trialog.ch - www.trialog.ch*

[a[r[b|i]d]o]

Jetzt Metadaten 2014 anfordern!

Tel. +41 31 300 63 89
inserate@staempfli.com



ARTProfil GmbH

alles für die Bestandserhaltung

Qualitätsprodukte der Firma Schempp®
Museumskarton Grossformate, Museumsglas
Fotoarchivierung, Klebebänder, Montageecken

NEU: EB-Wellkarton 4,5 mm / Tyvek® / Solanderboxen
Museumskarton ungepuffert, 2,2mm 118,9 x 168,2 cm

Fordern Sie detaillierte Informationen unverbindlich an
oder kontaktieren Sie uns im Web.

ARTProfil GmbH, Bahnhofstrasse 75, 8887 Mels
Tel. 081 725 80 70 Fax 081 725 80 79
office@artprofil.ch www.artprofil.ch



metasystems

la communication globale

- Des solutions personnalisées en communication
- Conception, multimedia et nouvelles technologies
- Un réseau de professionnels de l'information

Chemin des Rosiers 2B • 1701 Fribourg • 079 660 62 66 • gilliozs@metasystems.ch



neu – nouveau

Alles aus einem Hause – komplette Sortimente von Oekopack, KLUG Conservation und Secol.

Un seul interlocuteur pour les gammes complètes d'Oekopack, KLUG Conservation et Secol.

Zur Erhaltung von Originalen entwickeln und fabrizieren wir seit mehr als 25 Jahren Produkte aus Papier und Karton, garantiert ISO 9706

- Schachteln aus Vollkarton, Mappen, Umschläge, u.v.m. von **Oekopack**
- Karton, Papiere, Boxen aus Wellkarton, etc. von **KLUG Conservation**
- Polyesteretaschen v.a. für Fotobestände von **Secol**



Depuis plus que 25 ans nous développons et fabriquons des produits en papier et carton, garanti ISO 9706 pour la conservation d'originaux

- boîtes, chemises à rabats, chemises simples, cartables, etc. d'**Oekopack**
- carton, papiers, boîtes en carton cannelé, etc. de **KLUG Conservation**
- pochettes en polyester de **Secol** essentiellement pour fonds photographiques

Oekopack Conservus AG

Lattigen, 3700 Spiez, Tel. 033 655 90 55/60, Fax 033 655 90 51, www.oecag.ch, info@oecag.ch

Ihr Spezialist für digitale Archive

Dienstleistungen und Lösungen für ein effektives Dokumenten-Management

Unser Dienstleistungsspektrum im Überblick:

- Scannen von Dokumenten und Büchern, von Kleinst- bis über A0 Format
- Scannen von großformatigen Dokumenten und Zeichnungen bis zu 1,2 Meter Breite und 12 Meter Dokumentenlänge in Farbe, Graustufe und s/w
- Scannen von sämtlichen Arten von Mikrofilmen, Fichen und Jackets, sowie Rollfilmen, Negative / Glasnegative / Dias
- Konvertierung digitaler Daten auf Mikrorollfilm und Fiche sowie Dokumentenverfilmung
- Digitale Archive für technische, kaufmännische und katasterspezifische Dokumente/Pläne Softwareentwicklung und Anpassung
- Individualsoftware



Tecnocor ACC AG

Arsenalstrasse 51
CH-6010 Kriens
Telefon: +41 (0)41 440 74 22
Mobil: +41 (0)79 340 46 06
Fax: +41 (0)41 440 85 84
E-Mail info@tecnocor.ch
www.tecnocor.ch

AIW – ArchiImage Writer

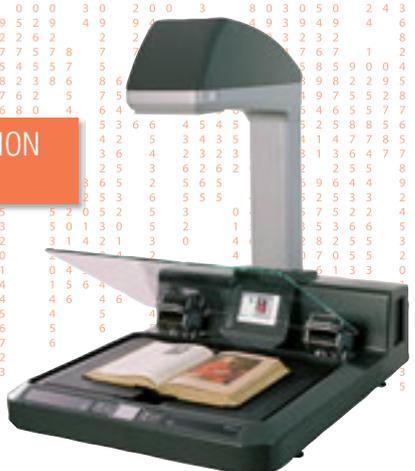
schreibt digitale Daten als analoge Informationen auf Mikrofilm.

- schreibt 16/35mm Mikrofilme (Rollfilme)
- wahlweise Hoch- / Querformat durch einfaches Drehen des Kamerakopfes
- s/w, Graustufen und Farbe
- Einzel-, Doppel- oder Vierfachbild pro Frame
- Annotationen, Indexierung, Zähler per Software definierbar.
- unterstützt über 100 Fileformate
- 10MP Graustufen-TFT (A2) oder
- 6MP Farb-TFT (A2)
- Bewährte Technologie für die Langzeitarchivierung



1 9 0 4 7 1 4 9 4 2 3 1 4 6 5 1 6 1 1 2 3 4 2 1 1 1 1 2 9 5 1 1
0 0 0 1 0 5 0 2 0 0 0 0 0 0 3 0 2 0 0 3 8 0 3 0 5 0 2 4 3 0 0
9 2 0 7 9 8 9 3 9 9 6 9 5 2 9 4 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9
2 0 1 0 3 7 2 6 2 2 9 2 2 6 2 8 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
7 9 1 6 6 7 7 7 7 5 7 8 7 2 7 5 7 8 7 2 7 5 7 8 7 2 7 5 7 8
5 7 5 5 8 5 5 9 5 4 5 5 7 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
8 1 1 4 8 7 8 8 8 8 2 3 8 6 8 6 8 6 8 6 8 6 8 6 8 6 8 6 8 6
7 5 7 4 7 6 7 7 7 6 2 5 7 5 7 5 7 5 7 5 7 5 7 5 7 5 7 5 7 5
6 4 2 2 5 6 5 6 8 4 6 6 8 0 4 6 4 6 4 6 4 6 4 6 4 6 4 6 4 6
5 3 6 6 4 5 4 5 5 3 6 6 4 5 4 5 5 3 6 6 4 5 4 5 5 3 6 6 4 5
4 2 5 3 4 3 5 5 3 6 6 4 5 4 5 5 3 6 6 4 5 4 5 5 3 6 6 4 5
3 6 6 4 2 3 2 3 2 5 3 6 2 6 2 3 6 6 4 2 3 2 3 2 5 3 6 2 6 2
3 6 5 2 5 6 5
3 5 6 3 2 5 5 2 0 5 2 5 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 3 5 0 3 5 0 1 3 0 5 2 1 3 0 4 3 0 4 3 0 4 3 0 4 3 0 4
1 4 1 2 3 1 4 2 1 3 0 4 3 0 4 3 0 4 3 0 4 3 0 4 3 0 4 3 0
4 3 4 0 4 1 0 4 1 5 6 4 6 4 6 4 6 4 6 4 6 4 6 4 6 4 6 4
4 1 4 1 0 4 4 1 5 6 4 6 4 6 4 6 4 6 4 6 4 6 4 6 4 6 4
5 5 4 1 5 6 4 6 4 6 4 6 4 6 4 6 4 6 4 6 4 6 4 6 4 6 4
2 4
5 4 5 4 5 4 5 4 5 4 5 4 5 4 5 4 5 4 5 4 5 4 5 4 5 4 5
6 5 6 5 6 5 6 5 6 5 6 5 6 5 6 5 6 5 6 5 6 5 6 5 6 5 6
7 6 7 6 7 6 7 6 7 6 7 6 7 6 7 6 7 6 7 6 7 6 7 6 7 6 7
8
9 9

YOUR INFORMATION MANAGER



SECUR'ARCHIV

Leader sur le marché de la dématérialisation de document en Suisse depuis 20 ans, Secur'Archiv numérise vos livres, registres, documents anciens et fragiles afin d'assurer leur pérennisation. Un système de caméras à faible charge thermique et lumineuse dépourvue d'UV se charge du travail tout en préservant vos documents les plus précieux.

Marktführer im Gebiet der Entmaterialisierung von Dokumenten in der Schweiz seit 20 Jahren, Secur'Archiv scannt Ihre Bücher, Aufzeichnungen, alte und empfindliche Dokumente um ihre Nachhaltigkeit zu gewährleisten. Die Arbeit wird durch ein Kamera-System mit niedriger thermischer Belastung und ohne UV-Strahlung getätigt, um Ihre kostbarsten Dokumente zu bewahren.



BERN
Tel : 031 853 1156

GENEVE
Tel : 022 827 8025

Secur'Archiv
www.securarchiv.ch
e-mail: info@securarchiv.ch

