

---

# Modernes Datenmanagement: Linked Open Data und die offene Bibliothek

Jens Mittelbach, SLUB Dresden

## Zusammenfassung:

Bibliotheken stehen im Spannungsfeld zwischen sich verändernden Erwartungen der potentiellen Benutzer/innen, technologischem Wandel, Sparzwängen und steigenden Kosten für Inhalte und Infrastrukturen. Um fortzubestehen, müssen sie sich neu erfinden. Eine Möglichkeit ist es, das Paradigma Offenheit auszufüllen, wonach u.a. die eigenen Kernprodukte entsprechend geltenden Webstandards für jeden frei zugänglich zur Verfügung gestellt werden. Um dies zu erreichen, müssen neue Infrastrukturen aufgebaut und neue Bündnisse eingegangen werden. Der Beitrag geht auf einen Vortrag auf dem Symposium Bibliotheks- und Informationsmanagement der Hochschule der Medien Stuttgart im Dezember 2014 zurück.

## Summary:

Libraries are entangled in a mesh of changing demands of potential users, technological advancement, budget cut downs and rising costs for content and infrastructure. In order to survive, they have to reinvent themselves. One option is to adopt the paradigm of openness. This means, among other things, offering the library's core products free of charge to anyone who can make use of them. This can only be done by implementing new infrastructures and forging new alliances. The article is based on a talk given at a symposium of library and information management in December 2014 at the Hochschule der Medien at Stuttgart.

Zitierfähiger Link (DOI): <http://dx.doi.org/10.5282/o-bib/2015H2S61-73>

Autorenidentifikation: Mittelbach, Jens: GND 1045090247,

ORCID: <http://orcid.org/0000-0002-2222-7172>

## 1. Das Geschäftsfeld von Bibliotheken und das Paradigma der Offenheit

Im Zeitalter der digitalen Informationsversorgung stehen Bibliotheken unter einem in der Geschichte des Bibliothekswesens beispiellosen Legitimationszwang. Als öffentlichen Einrichtungen, die einen Großteil ihrer Finanzierung aus der öffentlichen Hand erhalten, mag Bibliotheken ihre Situation noch nicht vollends bewusst geworden sein. Gleichwohl ist sie mittel- bis längerfristig existenzbedrohend. Dafür sprechen die allorts zu verzeichnenden Kürzungen der Personalstellen wie auch der zunehmende Druck im Bereich der Erwerbungsbudgets. Solche Sparauflagen sind Reaktionen der Politik, die zunehmend Schwierigkeiten hat, Bibliotheken einen genuinen kulturellen Wert beizumessen. In England findet z.B. der Kollaps des Systems der öffentlichen Bibliotheken bereits statt: Augenscheinlich führten hier seit 2011 ausgesprochen archaische Gepflogenheiten im Bibliotheksmanagement und technologisch-mediale Rückwärtsgewandtheit zur Schließung von

über 300 Häusern.<sup>1</sup> Radikaler Wandel ist Bibliotheken anzuempfehlen, damit sie im Universum der Informationsmärkte auch in Zukunft noch eine Entität darstellen können.<sup>2</sup>



Abb. 1: *Das Image der Bibliotheken in der Popkultur: Bücher, Würde, Gravitas – und der Staub von Jahrhunderten.* Bildnachweis: „library books“ von timetrax, unter CC BY-SA 2.0, auf Flickr<sup>3</sup>

Denn dieses Universum ist von Rahmenbedingungen gekennzeichnet, die sich rasant ändern. Der technologische Wandel beschleunigt sich und die Vielfalt der Informationskanäle wächst unablässig. Digitale Medien, die im Gegensatz zu den in Bibliotheken noch immer mit einer gewissen Vorliebe behandelten analogen Medien sofort zugänglich sind, bestimmen mehr und mehr die Benutzungsgewohnheiten und Benutzungserwartungen der Informationssuchenden. Die im Zuge des Open-Access-Trends zunehmende freie Verfügbarkeit von wissenschaftlicher Literatur lässt Bibliotheken in der Rolle als Zugangsanbieter an Bedeutung verlieren. Kommerzielle Informationsversorger, die im Gefolge der großen IT-Konzerne auftreten oder auch mit ihnen eine Personalunion bilden, haben eine Macht entwickelt, die Bibliotheken außer jede Konkurrenz stellt. Mobile Informationstechnik ist zum gesellschaftlichen Statussymbol geworden, während Bibliotheken als technologieängstlich und innovationsfern gelten und von ihrem in der Populärkultur mit Verlässlichkeit tradierten Image der Foliantensammlung, dem Verstaubtheit und Verschrobenheit inhärent sind<sup>4</sup>, nicht loskommen.

- 1 Vgl. z.B. Clark, Nick: The great British library betrayal: Closures bring national network to brink of 'absolute disaster', reveals official inquiry. In: The Independent, 17. Dezember 2014, <http://www.independent.co.uk/arts-entertainment/books/news/the-great-british-library-betrayal-closures-have-brought-national-network-to-brink-of-absolute-disaster-reveals-official-inquiry-9931965.html> (05.06.2015).
- 2 Vgl. Johnson, Larry, u. a.: NMC Horizon Report: 2014 Library Edition, Austin, Texas: The New Media Consortium, 2014, <http://cdn.nmc.org/media/2014-nmc-horizon-report-library-EN.pdf> (05.06.2015), S. 28-29.. <https://www.flickr.com/photos/timetrax/376152628/>
- 3 <https://www.flickr.com/photos/timetrax/376152628/>
- 4 Das lässt sich u.a. anhand populärer Film- und Fernsehproduktionen wie „Star Wars, Episode II: Attack of the Clones“ (vgl. <http://reel-librarians.com/2013/03/26/the-jedi-librarian/> [05.06.2015]) oder dem Tatort „Kalter Engel“, in dem eine Unibibliothek mit schnippisch-strengem (und leider inkompetentem) Personal eine Nebenrolle spielt, belegen.

Für die überwältigende Mehrzahl der Informationssuchenden sieht die Welt dagegen heute anders aus. Diese Welt ist ein virtueller Flickenteppich der unterschiedlichsten Apps, Plattformen und Dienste, auf dem jede/r Nutzer/in seinen Platz findet. Jeder dieser Plätze ist grundsätzlich individuell gestaltbar; sie alle präsentieren sich gleichwohl, von Trends und Moden bestimmt, vergleichsweise uniform im Rahmen massenkultureller Konventionen.



Abb. 2: Die Welt von heute für die technologie-affinen Digital Natives.  
Bildnachweis: „Ball, Kugel, Netzwerke, Internet“ von geralt, unter CC0, auf Pixabay<sup>5</sup>

Bibliotheken kommen auf den wenigsten dieser Plätze überhaupt vor. Zum Teil liegt das natürlich am Grad der Sichtbarkeit der bibliothekarischen Angebote (zu allererst der immensen Daten- bzw. Metadatenschätze). In einer Keynote-Rede auf der SWIB 2014 beklagte Eric Miller, dass Coffee-Shops im Internet sichtbarer seien als Bibliotheken.<sup>6</sup> Zu einem guten Teil sind es aber die Angebote selbst, die nicht zeitgemäß, unattraktiv, irrelevant und damit im Grunde überflüssig sind oder es doch zumindest zu werden drohen. Bibliotheken verlieren an Legitimation in einem von Technologie beherrschten Informationsuniversum – es sei denn, sie wandeln sich, erfinden sich neu und setzen auf neuartige Paradigmen.

Ein solches neues – oder neu ausgedeutetes – Paradigma ist die konsequente Verfolgung von Offenheitsprinzipien als Grundsatz bibliothekarischer Arbeit bzw. grundlegendes bibliothekarisches Selbstverständnis.<sup>7</sup> Dabei geht es einerseits darum, als Institutionen für offene Informationsinfrastrukturen in dem Sinne einzutreten, wie es der Ausschuss für Wissenschaftliche Bibliotheken

5 <https://pixabay.com/de/ball-kugel-netzwerke-internet-142738/>

6 Miller, Eric: Moving from MARC: How BIBFRAME moves the Linked Data in Libraries conversation to large-scale action. SWIB14, Bonn, 3. Dezember 2014, Abschn. Act 3: Visibility, 37:38, <http://www.scivee.tv/node/63285> (05.06.2015).

7 Vgl. Lohmeier, Felix; Mittelbach, Jens: Offenheit statt Bündniszwang. In: Zeitschrift für Bibliothekswesen und Bibliographie 61 (2014), H. 4/5, S. 209-214, <http://dx.doi.org/10.3196/1864295014614554>; Pre-Print verfügbar unter <http://nbn-resolving.de/urn:nbn:de:bsz:14-qucosa-157772> (05.06.2015).

und Informationssysteme der DFG hinsichtlich der Förderstrategien der Forschungsgemeinschaft postuliert:

„Alle vorgeschlagenen Maßnahmen zielen darauf ab, den möglichst offenen Zugang zu wissenschaftlich relevanter Information, zu Forschungsdaten sowie zu Arbeits- und Kommunikationsplattformen zu optimieren. Dabei setzt sich die DFG konsequent für das Paradigma des Open Access ein.“<sup>8</sup>

Offenheit als Prinzip geht aber über Open Access und Open Infrastructure noch hinaus und umfasst Lehre und Forschung im Ganzen, auch auf einer ethischen, philosophischen Ebene. Es geht also darum, dass Bibliotheken sich im Paradigma der Offenheit als Anwälte einer offenen Gesellschaft und einer offenen Kultur verstehen und alle ihre Aktivitäten darauf ausrichten.

Um das zu verwirklichen, müssen Bibliotheken viel mehr Energie in den Aufbau branchengerechten Know-Hows in den Bereichen Informationstechnologien, Wissenschaftskommunikation, Marktökonomie und Informationswissenschaft stecken.



Abb. 3: Bibliotheken als Verfechter von Offenheitsprinzipien.  
Bildnachweis: „Sorry, We’re Open“ von tara hunt, unter CC BY-SA 2.0, auf Flickr<sup>9</sup>

Erst dann können sie sich aus der Umklammerung oligopolistisch organisierter Software- und Content-Anbieter befreien und die innovationsfeindlichen Kartellstrukturen bekämpfen, die die Entwicklung bedarfsgerechter Services seit Jahrzehnten lähmen. Erst dann können Bibliotheken ihre Mittäterschaft bei der Profitmaximierung der Wissenschaftsverlage aufkündigen, die dazu führt, dass letztere mit Zeitschriften jährlich 8 Mrd. US-Dollar Gewinn erwirtschaften, die zu 89 % aus dem

8 Deutsche Forschungsgemeinschaft, Ausschuss für Wissenschaftliche Bibliotheken und Informationssysteme: Die digitale Transformation weiter gestalten – Der Beitrag der Deutschen Forschungsgemeinschaft zu einer innovativen Informationsinfrastruktur für die Forschung, Positionspapier der Deutschen Forschungsgemeinschaft, Bonn, 2012, S. 4, <http://nbn-resolving.de/urn:nbn:de:hbz:5:2-60225> (05.06.2015).

9 <https://www.flickr.com/photos/missrogue/3353012785/>

Abonnementgeschäft mit wissenschaftlichen Bibliotheken (73 %) und anderen Wissenschaftseinrichtungen (16 %) stammen.<sup>10</sup> Erst dann können Bibliotheken und Hochschulen dem Debakel ein Ende machen, dass sie ihre eigenen Daten und Metadaten von der Produktion bis zum Management mehrfach bezahlen und am Ende noch nicht einmal besitzen. Erst wenn all dies geschehen ist, werden Bibliotheken als Sachwalter der Offenheit auf allen Ebenen wahrhaft und glaubhaft im Dienste der Nutzerinnen und Nutzer, im Dienste der wissenschaftlichen Standards sowie im Dienste des kulturellen Erbes stehen – und gleichzeitig werden sie dann eine genuine, durch niemanden streitig zu machende Aufgabe im System der Wissenschaftsinfrastrukturen erfüllen können.

Institutionelle Anwaltschaft ist der Auftrag, der sich aus dem Paradigma zur Durchsetzung allumfassender Offenheit und einer offenen Kultur für Bibliotheken ergibt. Die verschiedenen Ebenen, auf denen Offenheit durch die operativen, strategischen und politischen Aktivitäten der Bibliotheken gefördert werden sollen, lassen sich so visualisieren<sup>11</sup>:

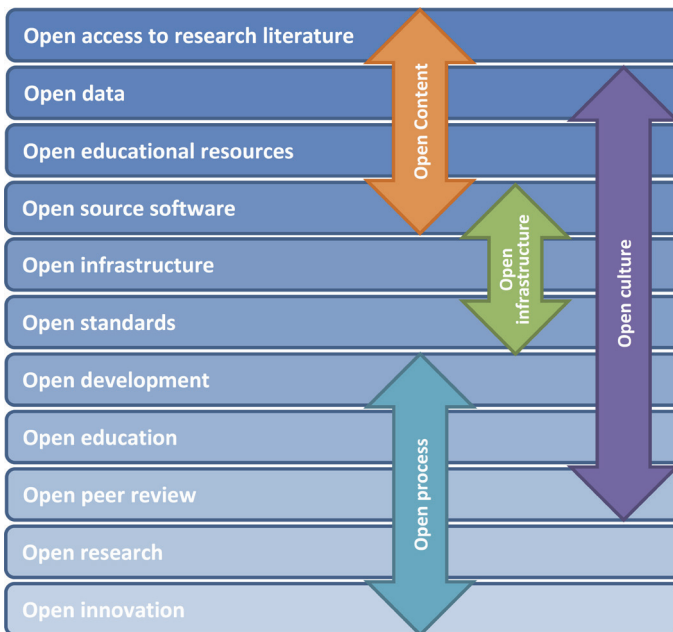


Abb. 4: Offenheit auf allen Ebenen.  
Bildnachweis: „The range of ‘Opens’“ von Lilian van der Vaart u.a.; unter CC BY-SA 3.0

10 Morrison, Heather: Economics of scholarly communication in transition. In: First Monday 18 (2013), H. 6, <http://dx.doi.org/10.5210/fm.v18i6.4370>.

11 Illustration entnommen aus van der Vaart, Lilian, u. a.: e-InfraNet: ‘Open’ as the default modus operandi for research and higher education, e-InfraNet, 2013, <http://www.surf.nl/en/knowledge-and-innovation/knowledge-base/2013/policy-paper-e-infranet-open-as-the-default-modus-operandi-for-research-and-higher-education.html> (05.06.2015), S. 11.

Im Rahmen des Offenheitsparadigmas werden freie Inhalte erzeugt, offene, wissenschaftsdienliche Infrastrukturen aufgebaut und offene wissenschaftliche Prozesse begünstigt.<sup>12</sup> Das bedeutet für Bibliotheken, ihre Metadaten zur freien Nachnutzung zur Verfügung zu stellen oder – z.B. im Zuge der Retrodigitalisierung gemeinfreier Werke – selbst offene Daten ohne neue Lizenzbeschränkungen zu erzeugen. Es bedeutet auch, sich von den lang gehegten Vorbehalten gegen Open-Source-Software frei zu machen und eigene Anwendungen konsequent quelloffen zu programmieren.

## 2. Daten, Daten, Daten

Sich Entwickler-Communities in Open-Source-Projekten anzuschließen bzw. selbst solche Projekte zu initiieren kann – entgegen landläufiger Annahmen und eindringlicher Warnrufe von Anbietern proprietärer Software sowie auch von namhaften Vertretern der Bibliotheksbranche<sup>13</sup> – für Bibliotheken durchaus sehr sinnvoll sein und einen besseren *Return of Investment* zeitigen, als es auf den etablierten Wegen möglich ist. Das Kerngeschäft von Bibliotheken, daran wird sich auch im Zeitalter der digitalen Gesellschaft nichts ändern, ist ja das Sammeln von Objekten bzw. Daten und das Erschließen dieser Objekte durch Metadaten. Die Instrumente, die dafür zur Verfügung stehen, sind – trotz Veränderungen am oligopolistisch bestimmten Markt in den letzten Jahren<sup>14</sup> – defizitär. Zwar bestimmen cloud-basierte Technologien und die Erweiterung des Wahrnehmungshorizontes auf Nicht-Print-Medien die Entwicklung auf diesem Gebiet.<sup>15</sup> Nach wie vor ist Software für das Management bibliothekarischer Daten jedoch allzu häufig zu wenig modular und serviceorientiert (im Sinne von SOA) aufgebaut. Vollen programmatischen Zugriff auf Daten und Funktionen erlauben nur wenige Systeme, stattdessen dominiert der Trend zu Hyperintegration und damit eine *One-size-fits-all*-Philosophie, die die individuellen Wünsche einzelner Bibliotheken unberücksichtigt lassen muss. Vor allem aber der Fokus auf die allgemeinen Anforderungen zum Umgang mit den im Bibliothekswesen historisch gewachsenen Datenformaten und auf die Abbildung schwer zu überwindender abstrakter Prozess-Konzepte verstellt den Weg hin zu modernem Datenmanagement und zur Entfaltung des vollen Potentials bibliothekarischer Daten. Datenhaltung in der Cloud oder in zentralisierten Datenpools bedeutet nicht automatisch das Aufbrechen der Daten-Silos, sondern zunächst einmal nur die Erweiterung der Silo-Abmessungen. Nach wie vor sind bibliothekarische Daten im *Deep Web* verborgen und damit dort unsichtbar, wo sich die aktiven und potentiellen Nutzer/innen von Bibliotheken heute aufhalten: im vom Suchmaschinen erschlossenen Internet.<sup>16</sup> Zudem bleiben hochstrukturierte und qualitativ hochwertige Daten untereinander unverbunden und unsemantisch, solange nicht endlich Linked-Data-Technologien auch im Bibliothekswesen breite Anwendung finden.

12 Siehe hierzu auch Pohl, Adrian: Bibliotheken: Wir öffnen Daten. Zum Stand der Entwicklung einer offenen Dateninfrastruktur. In: o-bib 1 (2014), S. 45-55, <http://dx.doi.org/10.5282/o-bib/2014H1545-55>, hier S. 47, der auch Hardware in sein Offenheitskonzept einbezieht.

13 Vgl. z.B. Ceynowa, Klaus: Content ist king, context is queen. In: B.I.T. online 18 (2015), H. 2, S. 182-185, <http://www.b-i-t-online.de/heft/2015-02-interview-ceynowa.pdf> (05.06.2015), hier S. 184, der Befürwortern von Offenheit eine „Open-Data-Gutmenschattitüde“ unterstellt.

14 Siehe zur Entwicklung der Industrie Breeding, Marshall: Library Systems Report. Operationalizing innovation. In: American Libraries Magazine, Mai 2015, <http://americanlibrariesmagazine.org/2015/05/01/library-systems-report/> (05.06.2015).

15 Ebd.

16 Vgl. Miller (wie Anm. 6) 33:05.



Domäne und des hohen Strukturiertheitsgrades der vorliegenden Daten sehr viel einfacher haben, als beispielweise eine Firma, die sich auf das Indexieren von Webseiten spezialisiert hat. Bereits 2011 hat der Wissenschaftsrat entsprechende Empfehlungen formuliert und darauf hingewiesen, dass mit der Adoption von Linked-Data-Technologien „Verbundkataloge in ihrer konventionellen Form zumindest als bibliographische Datenressourcen an Bedeutung verlieren“ werden.<sup>18</sup> Aufgrund der vorhandenen Metadaten-Expertise sollten Bibliotheken eigentlich Vorreiter und Wegbereiter bei der Etablierung des *Semantic Web* sein. Bislang ist allerdings im Bibliothekswesen auf diesem Sektor – abgesehen von mutigen Vorstößen einiger weniger Protagonisten bzw. Einrichtungen – zu wenig, und insbesondere zu wenig praktische Aktivität zu verzeichnen. Das ist nicht zuletzt deshalb so, weil Linked-Data-Ansätze sehr stark von Offenheitskriterien abhängen, gerade aber Offenheit vielen Bibliotheken suspekt ist: „Unfortunately, many librarians – especially on the management level – are not aware of the importance of applying web standards and publishing open data.“<sup>19</sup>

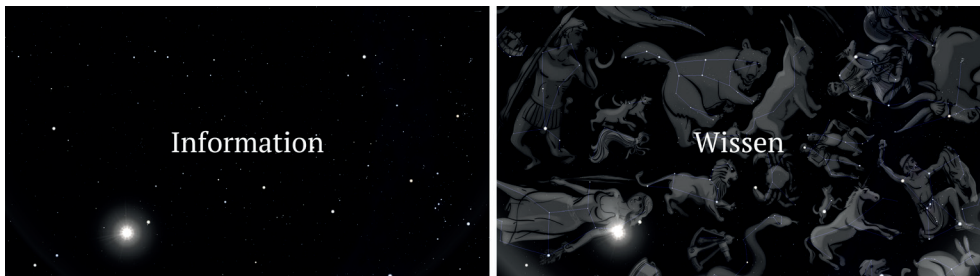


Abb. 6: Unverknüpfte im Gegensatz zu verknüpfte Daten. Bildnachweis: Screenshots aus Stellarium, unter GNU GPL

#### 4. Das Management bibliothekarischer Daten

Bibliothekarisches Datenmanagement ist mehr denn je eine Herausforderung. Mit dem Aufkommen von *Web-Scale Discovery* gilt es, immer mehr Daten aus vielen verschiedenen Datenquellen miteinander zu integrieren. Die bislang zur Verfügung stehenden Instrumente erlauben eine solche Integration nur ungenügend. Gerade die Deduplizierung und das Zusammenführen von Metadaten, die sich auf dieselben Objekte beziehen, ist in der Praxis nach wie vor ein ungelöstes Problem. Darüber hinaus bleiben Wünsche bei der Anreicherung mit Normdaten und weiterführenden Informationen sowie bei der Hierarchisierung von Metadaten entsprechend der FRBR-Prinzipien offen – alles Voraussetzungen für die Semantisierung von Daten und für eine bessere Datenpräsentation, die den Nutzerinnen und Nutzern ein intuitiveres Herangehen beim Retrieval, ein Explorieren von Wissensräumen erlauben würden. Bei der Datenintegration geht es nicht um die Herstellung eines kleinsten gemeinsamen Nenners – der gewöhnliche Effekt der landläufigen Datennormalisierung –, denn dies bedeutet Reduktion, wo Transformation und Extension gefragt sind. Ein weiteres Desiderat ist es, dass Datenintegration von Domänenexperten geleistet werden können muss. Nicht

18 Wissenschaftsrat: Empfehlungen zur Zukunft des bibliothekarischen Verbundsystems in Deutschland, 28. Januar 2011, <http://www.wissenschaftsrat.de/download/archiv/10463-11.pdf> (05.06.2015), S. 32.

19 Pohl, Adrian: Discovery Silos vs. the Open Web. Open Bibliography and Open Bibliographic Data, 23. Juni 2013, <http://openbiblio.net/2013/06/23/discovery-silos-vs-the-open-web/> (05.06.2015).



Programmierer, sondern Bibliothekare sollen die Integrationswerkzeuge bedienen können. Es sind Sharing-Funktionen wünschenswert, die es erlauben, Daten-Mappings und Transformationsworkflows über Institutionsgrenzen hinweg nachzunutzen, so dass einmal getane Arbeit auch anderen zugute kommt. Zu guter Letzt müssen die Daten auf einfache Weise – per Knopfdruck – als Linked Open Data publiziert werden können, damit sie von anderen Institutionen, und nicht zuletzt von Web-Entwicklern aller Couleur genutzt – und somit verbreitet – werden können.

## 5. Infrastrukturen und Bündnisse

Dafür sind neuartige Werkzeuge und Infrastrukturen nötig. Aufgebaut werden können und müssen sie von den Bibliotheken selbst bzw. unter deren enger Beteiligung, denn sie sind es, die ihre lokalen Daten am genauesten kennen und entsprechende Anforderungen an das Datenmanagement formulieren können. Dazu müssen sie natürlich Ressourcen widmen – in Zeiten sinkender Etats, mangelhafter Ausstattung mit qualifizierten Entwicklern und immer höherer Erwartungen an die IT-Abteilungen der Einrichtungen ein möglicher Einwand gegen die *Do-it-yourself-Variante*. Bedenkt man jedoch, dass Bibliotheken es vielerorts gewöhnt sind, erheblichen Aufwand in die Anpassung von verkapselten, unzugänglichen und nicht für individuelle Anpassung konzipierten *Black-Box-Softwaresystemen* zu stecken, scheint es nicht mehr so abwegig, solche Forderungen zu stellen. Statt die Defizite proprietärer Software durch *Workarounds* beheben zu wollen, die in den einzelnen Häusern fragil programmiert werden und die bei jedem Software-Update zu kollabieren drohen, sollten Bibliotheken ihre Entwicklerkapazitäten doch lieber in vielversprechende Lösungen investieren, deren Quellcodes offen liegen und die einen hohen Grad an Flexibilität ermöglichen. Statt zum Teil hohe Software-Lizenz- bzw. Anschaffungskosten in Kauf zu nehmen und sich dann einsam mit der Implementierung nicht vorgesehener individueller Funktionen abzumühen, wäre es doch zumindest im Hinblick auf die Werkzeuge, die zur Wahrnehmung der Kernaufgaben von Bibliotheken dienen, sinnvoller, dass die Einrichtungen ihre Kräfte bündeln und personell und finanziell zur Arbeit überregionaler Entwickler-Communities beisteuern. Statt nur den eigenen Tellerrand zu sehen, sollten Bibliotheken doch offene Infrastrukturen zur Schaffung offener Inhalte aufbauen – zu ihrem eigenen Nutzen und zum Nutzen ihrer Kunden. Stehen solche Infrastrukturen zunehmend zur Verfügung, werden auch kleinere Einrichtungen wie z.B. die eine oder andere Fachhochschulbibliothek, die keine eigenen IT-Kapazitäten hat, davon profitieren und damit gleichzeitig ihren Beitrag zu Offenheit und Open Data leisten können.

Diese Art von Kooperation wäre für Bibliotheken neuartig und sicher auch mutig. Sie würde es ermöglichen, in einer nach wie vor von erstarrten Verbundstrukturen charakterisierten Bibliothekslandschaft die Fokussierung auf die regionale Ebene zu überwinden, den engen Rahmen der damit verbundenen politischen *Raison d'Être* zu sprengen und den Blick zu weiten auf eine zunehmend entgrenzte, dezentrale, von Netzwerkstrukturen gekennzeichnete Welt. In dieser Welt können, unabhängig von Gebietsabsteckungen, seien sie nun regional oder national, strategische Bündnisse geschmiedet werden. Hier ist man nicht schicksalhaft an Partner gekettet, vielmehr werden diese Partner themen- und aufgabenbezogen gewählt. Bei konsequenter Verfolgung von Offenheitsprinzipien steht in dieser Welt auch Entwicklungspartnerschaften mit privatwirtschaftlichen Unternehmen nichts im Wege. Gerade das Bündnis mit gezielt identifizierten kommerziellen

Partnern kann Innovation herbeiführen und Entwicklungsschübe verursachen. Nachhaltigkeit ist dann immanent, wenn gleichzeitig ein Know-How-Transfer hin zu den Bibliotheken stattfindet.

## 6. D:SWARM – eine Open-Source-Datenmanagement-Plattform

Ein Beispiel für ein Bündnis dieser neuen Art ist D:SWARM<sup>20</sup>, ein kooperatives Entwicklungsprojekt der Sächsischen Landesbibliothek – Staats- und Universitätsbibliothek Dresden (SLUB) und Avantgarde Labs GmbH<sup>21</sup>.

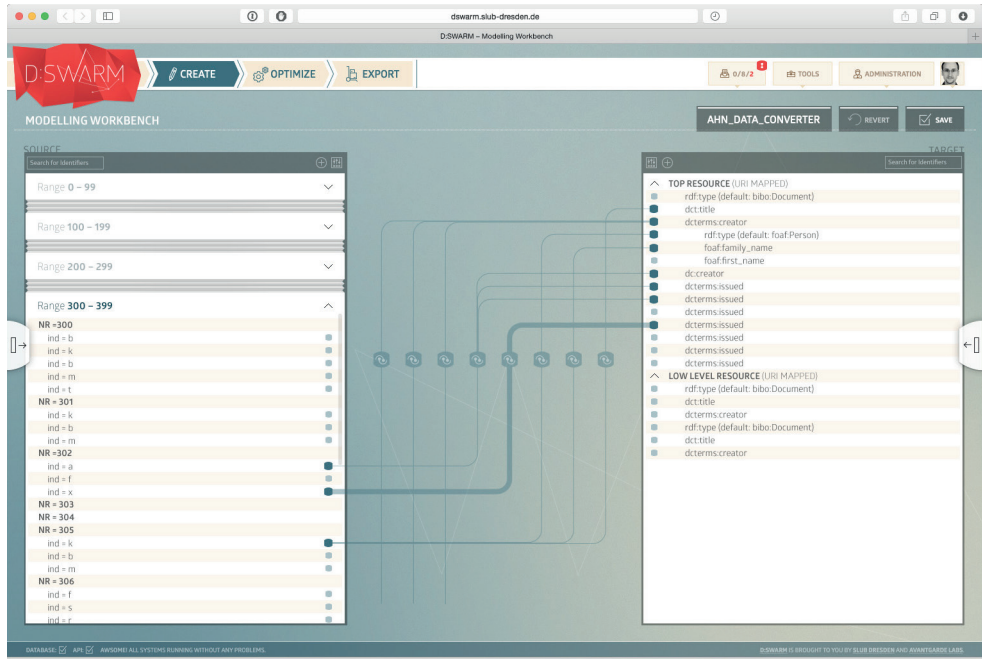


Abb. 7: D:SWARM-Webapplikation zur Definition von Datenmapping- und Transformationsroutinen. Bildnachweis: D:SWARM-Mock\_Up von Thomas Jung, SLUB Dresden

Das Ziel der Entwicklung ist eine Open-Source-Lösung für das Management aller Arten von bibliothekarischen Daten. D:SWARM ist dabei als Middleware konzipiert, die sich architektonisch zwischen bestehende Datenmanagementsysteme (z.B. lokale ILS, Dokument-Repositories, IDM-Systeme usw.) und Daten-Präsentationssysteme (z.B. Discovery-Frontends) einfügt. Auf diese Weise können unter vorläufiger Beibehaltung vorhandener Infrastrukturen und der daran geknüpften Workflows sämtliche bibliothekarischen Datenintegrations- und Datenmodellierungsprozesse durchgeführt werden, wobei Daten verlustfrei in ein flexibles, graphenbasiertes Datenmodell überführt werden. Die Definition von Extraktion-Transformation-Load-Routinen (also ETL-Prozessen) erfolgt dabei

20 Webseite des Projektes: [www.dswarm.org](http://www.dswarm.org) (05.06.2015).

21 Webseite der Firma: [www.avantgarde-labs.de](http://www.avantgarde-labs.de) (05.06.2015).

mithilfe einer grafischen, intuitiv zu bedienenden Benutzeroberfläche, so dass Datentransformation zukünftig direkt von denjenigen geleistet werden kann, die über das dafür notwendige Domänenwissen verfügen. Die Daten werden nach der Extraktion aus den heterogenen Datenquellen und einem Mapping auf das interne Zielschema (inklusive einer Reihe von Basistransformationen, für die das Metafacture-Framework benutzt wird<sup>22</sup>) innerhalb des Graphen verschiedenen weiteren Transformations-, Deduplizierungs- und Anreicherungsprozessen unterzogen, so dass die Datenqualität und -homogenität signifikant steigt. Integraler Bestandteil der Lösung ist die Möglichkeit, die transformierten Daten nicht nur beliebigen Präsentationssystemen in variablen Output-Formaten zur Verfügung zu stellen, sondern sie auf Wunsch auch als Linked Data zu publizieren. Außerdem ist geplant, Transformationsroutinen und Mappings innerhalb einer Community verfügbar zu machen (zu „teilen“). Auf diese Weise können andere Einrichtungen, die ähnliche Datenmanagement-Aufgaben zu bewältigen haben, von der Arbeit ihrer Kolleginnen und Kollegen profitieren, aber auch ihrerseits einen Beitrag leisten. Eine verbesserte Datenqualität, die Singularisierung von Informationen als Entitäten entsprechend der Linked-Data-Prinzipien und die Erhöhung des semantischen Verknüpfungsgrades ermöglicht schließlich die Präsentation der Daten in neuartigen Frontends, die Discovery und Wissensexploration erst wahrhaft möglich machen.<sup>23</sup>

Das Projekt startete im Juni 2013 als Teilprojekt eines aus EFRE-Mitteln geförderten Gemeinschaftsprojektes, an dem, neben den bereits genannten Partnern, die Universitätsbibliothek Leipzig und die Forschungsgruppe Agile Knowledge Engineering and Semantic Web (AKSW) des Instituts für Angewandte Informatik der Universität Leipzig beteiligt waren. Seit Ende der Förderphase im September 2014 wird die Entwicklung in Zusammenarbeit mit der Universitätsbibliothek Leipzig, der Universitätsbibliothek Dortmund, der Europauniversität Viadrina und dem Metafacture-Team in einem offenen Prozess weiter vorangetrieben. Interessenten sind eingeladen, die Webapplikation zu testen<sup>24</sup> und sich an der Entwicklung zu beteiligen.

Wenn solche Beispiele Schule machen, bibliothekarisch relevante Open-Source-Projekte in Zukunft mehr Gravitation entwickeln und Bibliotheken sich darüber hinaus ganz offensiv dem Thema Offenheit verschreiben, kann die Branche in Abgrenzung und Ergänzung zu kommerziellen Informationsanbietern ihre Legitimation zurückerobern, ihre Relevanz in der Kultur- und Bildungslandschaft beweisen und zukunftsfähig werden.

## Literaturverzeichnis

- Breeding, Marshall: Library Systems Report. Operationalizing innovation. In: American Libraries Magazine, Mai 2015, <http://americanlibrariesmagazine.org/2015/05/01/library-systems-report/> (05.06.2015).

22 Siehe [github.com/culturegraph/metafacture-core/wiki](https://github.com/culturegraph/metafacture-core/wiki) (05.06.2015).

23 Siehe Mittelbach, Jens: Wissensexploration mit bibliothekarischen Daten: Ein EFRE-Projekt zur Entwicklung einer LOD-basierten Datenmanagement-Plattform. 103. Deutscher Bibliothekartag, Bremen, 4. Juni 2014, <http://prezi.com/xgn-kazs6f3ep/datenmanagement-plattform-der-slub-dresden/> (05.06.2015), und Mittelbach, Jens; Glaß, Robert: A Library Data Management Platform Based on Linked Open Data. SWIB14, Bonn, 2. Dezember 2014, <https://speakerdeck.com/swib14/a-library-data-management-platform-based-on-linked-open-data> (05.06.2015).

24 Unter [demo.dswarm.org](http://demo.dswarm.org), Dokumentation auf [github.com/dswarm/dswarm-documentation/wiki](https://github.com/dswarm/dswarm-documentation/wiki), Backlog jira.slub-dresden.de (alle: 05.06.2015).

- Ceynowa, Klaus: Content ist king, context is queen. In: B.I.T. online 18 (2015), H. 2, S. 182-185, <http://www.b-i-t-online.de/heft/2015-02-interview-ceynowa.pdf> (05.06.2015).
- Clark, Nick: The great British library betrayal: Closures bring national network to brink of 'absolute disaster', reveals official inquiry. In: The Independent, 17. Dezember 2014, <http://www.independent.co.uk/arts-entertainment/books/news/the-great-british-library-betrayal-closures-have-brought-national-network-to-brink-of-absolute-disaster-reveals-official-inquiry-9931965.html> (05.06.2015).
- Deutsche Forschungsgemeinschaft, Ausschuss für Wissenschaftliche Bibliotheken und Informationssysteme: Die digitale Transformation weiter gestalten – Der Beitrag der Deutschen Forschungsgemeinschaft zu einer innovativen Informationsinfrastruktur für die Forschung, Positionspapier der Deutschen Forschungsgemeinschaft, Bonn, 2012, <http://nbn-resolving.de/urn:nbn:de:hbz:5:2-60225> (05.06.2015).
- Gallagher, Sean: How Google and Microsoft taught search to „understand“ the Web. In: Ars Technica, 7. Juni 2012, <http://arstechnica.com/information-technology/2012/06/inside-the-architecture-of-googles-knowledge-graph-and-microsofts-satori/> (05.06.2015).
- Johnson, Larry, u. a.: NMC Horizon Report: 2014 Library Edition, Austin, Texas: The New Media Consortium, 2014, <http://cdn.nmc.org/media/2014-nmc-horizon-report-library-EN.pdf> (05.06.2015).
- Lohmeier, Felix; Mittelbach, Jens: Offenheit statt Bündniszwang. In: Zeitschrift für Bibliothekswesen und Bibliographie 61 (2014), H. 4/5, S. 209-214, <http://dx.doi.org/10.3196/1864295014614554>; Pre-Print verfügbar unter <http://nbn-resolving.de/urn:nbn:de:bsz:14-qucosa-157772> (05.06.2015).
- Miller, Eric: Moving from MARC: How BIBFRAME moves the Linked Data in Libraries conversation to large-scale action. SWIB14, Bonn, 3. Dezember 2014, Abschn. Act 3: Visibility, <http://www.scivee.tv/node/63285> (05.06.2015).
- Mittelbach, Jens: Wissensexploration mit bibliothekarischen Daten: Ein EFRE-Projekt zur Entwicklung einer LOD-basierten Datenmanagement-Plattform, gehalten auf dem 103. Deutscher Bibliothekartag, Bremen, 4. Juni 2014, <http://prezi.com/xgnkzsf3ep/datenmanagement-plattform-der-slub-dresden/> (05.06.2015).
- Mittelbach, Jens; Glaß, Robert: A Library Data Management Platform Based on Linked Open Data, gehalten auf der SWIB14, Bonn, 2. Dezember 2014, <https://speakerdeck.com/swib14/a-library-data-management-platform-based-on-linked-open-data> (05.06.2015).

- Morrison, Heather: Economics of scholarly communication in transition. In: First Monday 18 (2013), H. 6, <http://dx.doi.org/10.5210/fm.v18i6.4370>.
- Pohl, Adrian: Bibliotheken: Wir öffnen Daten. Zum Stand der Entwicklung einer offenen Dateninfrastruktur. In: o-bib 1 (2014), S. 45-55. <http://dx.doi.org/10.5282/o-bib/2014H1S45-55>.
- Pohl, Adrian: Discovery Silos vs. the Open Web. Open Bibliography and Open Bibliographic Data, 23. Juni 2013, <http://openbiblio.net/2013/06/23/discovery-silos-vs-the-open-web/> (05.06.2015).
- van der Vaart, Lilian, u. a.: e-InfraNet: 'Open' as the default modus operandi for research and higher education, e-InfraNet, 2013, <http://www.surf.nl/en/knowledge-and-innovation/knowledge-base/2013/policy-paper-e-infranet-open-as-the-default-modus-operandi-for-research-and-higher-education.html> (05.06.2015), S. 11.
- Wissenschaftsrat: Empfehlungen zur Zukunft des bibliothekarischen Verbundsystems in Deutschland, 28. Januar 2011, <http://www.wissenschaftsrat.de/download/archiv/10463-11.pdf> (05.06.2015).