

# Wikipedia als Teil einer freien bibliothekarischen Informationsinfrastruktur\*

Jakob Voß

Wikimedia Deutschland e.V.

[jakob.voss@wikimedia.de](mailto:jakob.voss@wikimedia.de)

**Abstract:** This paper explores the usage of Wikipedia in the context of digital libraries. After a short description of Wikipedia and its main characteristics (wiki principle, free content, self-organization), commonalities of Wikipedia and digital libraries are shown. Possibilities to combine both are explained in general with connection methods and in detail with the cooperation between German Wikipedia and German National Library (DDB). For libraries Wikipedia can play an important part on the way to what is called *library 2.0* – a more interactive and collaborative rediscovery of libraries in the digital age. It is explained how authority files play an important role in connecting Wikipedia and digital libraries and finally their integration into the evolving Semantic Web.

**Categories (JITA Classification System of Library and Information Science)**

**I.** Information treatment for information services

**HZ.** Information sources, supports, channels: *None of these, but in this section.*

**Keywords:** Wikipedia, Digital Libraries, integration, information infrastructure

## 1. EINLEITUNG

Für viele Internetnutzer ist nicht die nächste Bibliothek sondern Wikipedia der erste Ausgangspunkt für Recherchen. Innerhalb von nur fünf Jahren hat sich die freie Online-Enzyklopädie zu einer der 20 meistbesuchten Webseiten weltweit entwickelt – in Deutschland werden sogar nur die Seiten von Google, Ebay, Yahoo und Spiegel Online häufiger aufgerufen.<sup>1</sup> Im Gegensatz zu kommerziellen Diensten, wie zum Beispiel Google Print, ist Wikipedia jedoch offen und frei. Dies bedeutet nicht nur, dass sich jeder konstruktiv an der Verbesserung und Erweiterungen ihrer Inhalte beteiligen kann, sondern vor allem, dass alle Inhalte als so genannte Freie Inhalte zur Weiternutzung zur Verfügung stehen. Ziel der Wikipedia ist nicht die Steigerung von Gewinnen oder Verdrängung von anderen Informationsmitteln, sondern wie ihr Gründer Jimbo Wales in einem Interview sagte, der freie Zugang zum gesamten menschlichen Wissen für jede Person.<sup>2</sup> Schon damit sollte deutlich sein, dass es starke Überschneidungen der Ziele und Interessen von Wikipedia und Bibliotheken gibt, und beide eher als mögliche Kooperationspartner, denn als Konkurrenten zu betrachten sind. Im Folgenden sollen existierende und zukünftige Möglichkeiten der Kooperation aufgezeigt und dargestellt werden wie Wikipedia sinnvoll in die bibliothekarische Informationsinfrastrukturen eingebunden werden und damit zur Informationsversorgung beitragen kann. Diese Arbeit entstand auf Basis eines am 22.3.2006 auf dem Bibliothekartag 2006 in Dresden gehaltenen Vortrags mit dem Titel ‚Wikipedia als Bestandteil digitaler Bibliotheken‘.

## 2. WIKIPEDIA

Wikipedia umfasst inzwischen weltweit mehr als 4 Millionen Artikel in über 200 Sprachen. Die deutschsprachige Wikipedia ist mit über 400.000 Artikeln nach der englischsprachigen Wikipedia die

---

\* Diese Arbeit steht unter den Bedingungen der freien Lizenzen GFDL und CC-BY-SA 2.5 bereit.

<sup>1</sup> <[http://www.alexa.com/site/ds/top\\_sites?cc=DE&ts\\_mode=country&lang=none](http://www.alexa.com/site/ds/top_sites?cc=DE&ts_mode=country&lang=none)> (15.6.2006).

<sup>2</sup> „Imagine a world in which every single person on the planet is given free access to the sum of all human knowledge. That's what we're doing.“ (Jimbo Wales im slashdot-Interview vom 28.7.2004).

zweitgrößte.<sup>3</sup> Ihre Inhalte werden von freiwilligen Autoren gemeinschaftlich in einem Wiki erstellt, korrigiert und ausgebaut. Das heisst, das nach dem ‚Wiki-Prinzip‘ jeder direkt (auch ohne Anmeldung) Änderungen und Erweiterungen vornehmen kann. Alle Bearbeitungen werden protokolliert und sind einsehbar, so dass sie ihrerseits von anderen Mitarbeitern verbessert werden können. Näheres zu den Besonderheiten der Schreibprozesse in Wikipedia und den Ursachen ihres Erfolges siehe in (Voss, 2006b). Eine praktische Einführung in Wikipedia gibt (Fiebig, 2005).

## 2.1 Inhalte

Da es sich bei Wikipedia um eine allgemeine Enzyklopädie handelt, finden sich in ihr Artikel aus einer Vielzahl verschiedener Themengebiete. Die einzelnen Artikel sollen allerdings so geschrieben sein, dass sie auch für fachfremde Leser verständlich sind (eine Anforderung, die in der deutschsprachigen Wikipedia salopp als ‚Oma-Test‘ bezeichnet wird) Auch muss eine gewisse Relevanz vorhanden sein; gerade darüber, was als relevant zu betrachten sei, gibt es allerdings regelmäßig heftige Diskussionen zwischen den Wikipedianern.

Von den vorhandenen Artikeln behandeln gut 22% einzelne Personen, etwa 8,5% der Artikel sind so genannte Begriffsklärungsseiten zur Auflösung von Homonymen und Polysemen. Eine weitere große Gruppe von Artikeln behandelt Länder und Orte – so sind beispielsweise bereits für alle 12.336 selbständigen Gemeinden Deutschlands und alle 2.737 Schweizer Gemeinden eigene Artikel vorhanden. Knapp 8% aller Artikel sind mit maschinenlesbaren geographischen Koordinaten versehen, die es ermöglichen, ihre Positionen automatisch in Karten einzublenden und über die Koordinaten nach Karten zu suchen.<sup>4</sup> Neben Körperschaften – darunter über 150 Bibliotheken – gibt es auch Artikel zu einzelnen literarischen Werken und Publikationen.

Zur Verbesserung der Qualität von Wikipedia-Artikeln werden diese vermehrt mit Literatur- und Quellenangaben versehen. Derzeit wird versucht, die noch relativ uneinheitlichen Ansetzungen durch detaillierte Ansetzungsregeln zu vereinfachen, so dass die einzelnen Bestandteile automatisch extrahiert werden können.<sup>5</sup> Auf diese Weise wäre es möglich, von den Literaturangaben mittels openURL direkt auf Bibliotheksbestände zu verweisen. Für ISBN-Nummern gibt es schon seit längerer Zeit einen einfachen Linkresolver, der in die Software der Wikipedia (MediaWiki) integriert ist.<sup>6</sup> Neben ISBN-Nummern werden die Nummern von Datensätzen in PubMed und RFC-Dokumenten automatisch verlinkt und verweisen direkt auf die entsprechenden Dokumente. Eine allgemeine Methode zur Verlinkung von Wikipedia-Artikeln und Einträgen in anderen Datenbanken bilden die so genannten Datenbanklinks.<sup>7</sup> Sie werden mit Hilfe von Vorlagen im Artikelquelltext in der Form `{{Kürzel|Identifikator}}` eingetragen, wobei das Kürzel je nach Datenbank unterschiedlich ist. Ein Beispiel für Datenbanklinks ist die weiter unten beschriebene Verknüpfung mit der Personennamendatei. Die Wikipedia-Artikel verschiedener Sprachen sind untereinander durch ‚Interlanguage-Links‘ verknüpft, für die eine spezielle Syntax existiert.<sup>8</sup>

Neben der relativ dichten Verlinkung der Artikel durch Hyperlinks werden die Einträge in Wikipedia durch ein System aus so genannten Kategorien erschlossen. Kategorien können als freie Schlagwörter sowohl Artikeln als auch anderen Kategorien zugeordnet werden; letzteres erzeugt hierarchische Verknüpfungen zwischen Kategorien, wodurch das Kategoriensystem, einen gemeinschaftlich erstellten Thesaurus bildet (Voss, 2006a). Weitere Metadaten lassen sich aus verschiedenen Arten von Vorlagen extrahieren, die in Form von Infoboxen<sup>9</sup> tabellarisch die wichtigsten Daten zum Gegenstand eines Artikels enthalten – dies hat sich beispielsweise bei Artikeln zu Städten und Gemeinden durchgesetzt, wo u. A. Einwohnerzahl, Fläche und Gemeindegemeinschaftsnummer aufgenommen werden.

## 2.2 Freie Inhalte und Arbeitsweise

<sup>3</sup> <<http://stats.wikimedia.org/EN/>> (15.6.2006).

<sup>4</sup> <[http://de.wikipedia.org/wiki/Wikipedia:WikiProjekt\\_Georeferenzierung](http://de.wikipedia.org/wiki/Wikipedia:WikiProjekt_Georeferenzierung)> (15.6.2006).

<sup>5</sup> <<http://de.wikipedia.org/wiki/Wikipedia:Literatur/Ansetzungsregeln>> (15.6.2006).

<sup>6</sup> <<http://de.wikipedia.org/wiki/Spezial:Booksources>> (15.6.2006).

<sup>7</sup> <<http://de.wikipedia.org/wiki/Wikipedia:Datenbanklinks>> (15.6.2006).

<sup>8</sup> <<http://de.wikipedia.org/wiki/Wikipedia:Internationalisierung>> (15.6.2006).

<sup>9</sup> <<http://de.wikipedia.org/wiki/Wikipedia:Infoboxen>> (15.6.2006).

Der wesentliche Unterschied der Wikipedia zu herkömmlichen Nachschlagewerken besteht neben dem Wiki-Prinzip darin, dass es sich ausdrücklich um eine *freie* Enzyklopädie handelt. Frei ist dabei im Sinne freier Inhalte zu verstehen, deren Nutzung keinen engen urheberrechtlichen Schranken unterworfen ist. Stattdessen wird durch entsprechende Lizenzen jedem das Recht zur Nutzung, Veränderung und Verbreitung eingeräumt. Im Falle der Wikipedia ist dies die GNU Free Documentation License (GFDL), die an die GNU General Public License (GPL) für Freie Software angelehnt ist. Beide Lizenzen enthalten eine ‚Copyleft‘-Klausel, nach der die Nutzung lizenzierter Inhalte daran geknüpft ist, dass daraus entstehende Werke wiederum unter der selben Lizenz freigegeben werden müssen. Dadurch wird sichergestellt, dass die Inhalte der Wikipedia immer frei bleiben werden.

Die Freiheit der Inhalte geht einher mit einer selbstorganisierten Arbeitsweise, bei der sich jeder direkt einbringen kann. Soziales Vertrauen ermöglicht ähnlich wie in sozialen Gemeinschaften und bei anderen Formen der Freiwilligenarbeit die Zusammenarbeit (Frost, 2006). Der Grad der Beteiligung ist in Wikipedia zwar sehr ungleich verteilt, aber es ist relativ einfach möglich, sich Reputation zu verschaffen (Voss, 2006b). Wie unter Anderem (Sanger, 2004) kritisiert, basiert diese Reputation allerdings nicht auf einem beispielsweise durch akademische Auszeichnungen erworbenen Expertenstatus sondern allein auf der Beurteilung der konkreten Arbeit in Wikipedia.

Die selbstorganisierte Arbeitsweise sorgt für Motivation bei den Wikipedianern und ist durch das weitgehende Fehlen von Hierarchien ungleich flexibler als Entscheidungsstrukturen in institutionellen Organisationen wie zum Beispiel Bibliotheken. Auf der anderen Seite können Mitarbeiter durch die ungewohnte Freiheit überfordert werden und das Konfliktpotential ist höher, da Regelungen zur Verteilung von Aufgaben und Kompetenzen selbst ausgehandelt werden müssen. Auch gestaltet sich die Führung und Planung als außerordentlich schwierig oder gar unmöglich, weil alle Beteiligten freiwillig dabei sind und relativ spontan agieren. Diese Mentalitätsunterschiede sollten bei der Zusammenarbeit mit Wikipedia stets bedacht werden.

### 3. Wikipedia und Bibliotheken

Mit dem Anspruch, dass Wissen der Welt zu sammeln, zu ordnen und verfügbar zu machen, hat Wikipedia von ihrer Zielsetzung einiges mit Bibliotheken gemeinsam. Bei ihrer Arbeit an Artikeln werden Bibliotheken von viele Wikipedia-Autoren genutzt. Artikel zu Themen aus dem bibliothekarischen Bereich werden seit Mitte 2004 im ‚Portal Bibliothek, Information, Dokumentation‘<sup>10</sup> gesammelt (Danowski und Voss, 2004). Auch der Verein Wikimedia Deutschland, der zur Förderung Freien Wissens insbesondere die Wikipedia unterstützt, arbeitet mit Bibliotheken zusammen. Mitte Juni 2006 fand an der Niedersächsische Staats- und Universitätsbibliothek die erste Wikipedia-Academy statt, um vor allem Wissenschaftler und Studenten die Mitarbeit in Wikipedia näher zu bringen. Zusammen mit dem Göttinger Digitalisierungszentrum wurde auch das erste Digitalisierungsprojekt des Vereins durchgeführt, bei dem das aus dem 16. Jahrhundert stammende Mathematiklehrbuch ‚Drei Register Arithmetischer ahnfeng zur Practic‘ des erzgebirgischen Rechenmeisters Andreas Reinhard digitalisiert und anschließend in der freien Quellensammlung Wikisource zur weiteren Bearbeitung veröffentlicht wurde.<sup>11</sup> Eine Affinität der Wikipedia zu Bibliotheken ist anscheinend auch in anderen Sprachversionen vorhanden: so nennen sich die Administratoren der spanischsprachigen Wikipedia ‚bibliotecarios‘.

Im folgenden Abschnitt wird zunächst die Zusammenarbeit mit Der Deutschen Bibliothek dargestellt, um danach auf grundsätzliche Möglichkeiten der Zusammenarbeit von Wikipedia und Bibliotheken einzugehen.

#### 3.1. Kooperation mit Der Deutschen Bibliothek

Durch einen Bericht über die Erschließung von Wikipedia-Artikeln mit Personendaten<sup>12</sup> (Voss, 2005) kam es zu einem Gespräch zwischen Vertretern der deutschsprachigen Wikipedia und Der Deutschen Bibliothek (DDB). Auf dem 94. Deutschen Bibliothekartag 2005 in Düsseldorf wurde eine Kooperation vereinbart, um die Personenartikel in Wikipedia mit Personendatensätzen der

<sup>10</sup> <<http://de.wikipedia.org/wiki/Portal:BiD>> (15.6.2006).

<sup>11</sup> <[http://de.wikisource.org/wiki/Drei\\_Register\\_Arithmetischer\\_ahnfeng\\_zur\\_Practic](http://de.wikisource.org/wiki/Drei_Register_Arithmetischer_ahnfeng_zur_Practic)> (15.6.2006).

<sup>12</sup> Jakob Voss: *Wikipedia sammelt Metadaten*. In: INETBIB. 14.2.2005  
<<http://www.ub.uni-dortmund.de/listen/inetbib/msg26842.html>> (15.6.2006).

Personennamendatei (PND) zu verknüpfen. Die Personennamendatei ist mit etwa 2,7 Millionen Datensätzen, davon etwa 1,1 Millionen individuellen Personen (Stand Frühjahr 2006), eine der weltweit größten Normdateien für Personen und Personennamen weltweit. Jeder Datensatz der PND besitzt eine PND-Nummer, mit der Personen unabhängig von ihrem Namen eindeutig identifiziert werden können. Die Normdatei wird in einer Vielzahl von Bibliotheken und zunehmend auch von Museen und anderen Institutionen zur Erschließung eingesetzt.

Die Kooperation bestand darin, dass zunächst die Personenartikel der Wikipedia mit PND-Nummern versehen wurden, um einen Abgleich der Daten in Wikipedia und PND zu erreichen. Diese Arbeit wurde maßgeblich durch ein von Christian Thiele entwickelt Webinterface vereinfacht, das für Personenartikel ohne PND-Nummer versucht, die passende Nummer zu ermitteln und anschließend das Eintragen in den entsprechenden Wikipedia-Artikel erleichtert. Die Eintragung einer PND-Nummer in einen Wikipedia-Artikel geschieht mittels Datenbanklinks (siehe oben). Beim Bearbeiten eines Artikels wird an entsprechender Stelle ein Auszeichnung der Form `{{PND|118578537}}` eingetragen, wobei 118601121 die PND-Nummer ist. Bis zum 4.6.2006 sind auf diese Weise bislang 19.977 der rund 100.000 Personenartikel in Wikipedia mit PND-Nummer versehen. Die aus der Wikipedia-Datenbank erstellbare Konkordanz von Wikipedia-Artikeln und PND-Nummern ist frei verfügbar.<sup>13</sup> Fehlerhafte oder unvollständige Angaben in der PND werden von Wikipedianern gesammelt und der PND-Redaktion an der DDB gesammelt zur Verfügung gestellt.<sup>14</sup> Die in Wikipedia eingetragenen PND-Nummer erzeugen automatisch einen Verweis auf den Katalog der Deutschen Bibliothek, so dass für die Leser und Autoren von Wikipedia ein direkter Nutzen vorhanden ist (Czech und Pfeiffer, 2006). Geplant ist, den Verweis auch auf andere Bibliotheken bzw. Bibliotheksverbünde auszuweiten, wozu bereits ein Prototyp eines Linkresolvers entwickelt wurde (Abbildung 1).

■ [Literatur von und über Karl Marx](#)  in verschiedenen Katalogen

↓ <http://wikimeta.de/pnd/?pnd=118578537>

**Literatursuche mit Normdaten** Hilfe

PND-Nummer:    ohne Javascript

<input checked="" type="checkbox"/> Wikipedia	<b>1 Treffer</b> [[Deutschland deutscher]] [[Philosoph]], [[Ökonom]] und [[Journalist]] (1818-1883)
Freie Online-Enzyklopädie	
<input checked="" type="checkbox"/> Die Deutsche Bibliothek	<b>2926 Treffer</b> Marx, Karl, 1818-1883, Dt. Philosoph und Politiker
Katalog Der Deutschen Bibliothek	
<input checked="" type="checkbox"/> Gemeinsamer Verbundkatalog	<b>2251 Treffer</b> Marx, Karl (1818-1883 ; Dt. Philosoph und Politiker)
Gemeinsamer Verbundkatalog	
<input checked="" type="checkbox"/> VD17	kein Treffer
Verzeichnis der im deutschen Sprachraum erschienenen Drucke des 17. Jahrhunderts	
<input checked="" type="checkbox"/> Zeitschriftendatenbank	<b>7 Treffer</b> Marx, Karl, 1818-1883, Dt. Philosoph und Politiker
Zeitschriftendatenbank	
<a href="#">Deutsches Musikarchiv</a>	

Suche | Hilfe  powered by WIKIMEDIA Toolserver

**Abbildung 2:** Prototyp eines Linkresolvers mittels PND-Nummern aus Wikipedia-Artikeln

Neben PND-Nummern können in Zukunft auch die Datensätze der Gemeinsamen Körperschaftsdatei (GKD) und der Schlagwortnormdatei (SWD) zur Erschließung von Wikipedia-Artikeln und Verlinkung auf Bibliotheksbestände verwendet werden. Konkrete Pläne bestehen dazu noch nicht, allerdings haben schon einige Wikipedianer erste Experimente unternommen.<sup>15</sup> Auch eine Ausweitung auf andere Sprachen ist wünschenswert. Im Projekt Virtual International Authority File (VIAF) sollen PND und Library of Congress Name Authority File (LCNAF) zusammengeführt werden, so dass auch andere Sprachen von der Verknüpfung profitieren werden.

### 3.2. Verknüpfungsmöglichkeiten

<sup>13</sup> <<http://wikimeta.de/personendaten/staging-area/20060604-pnd.tab>> (15.6.2006).

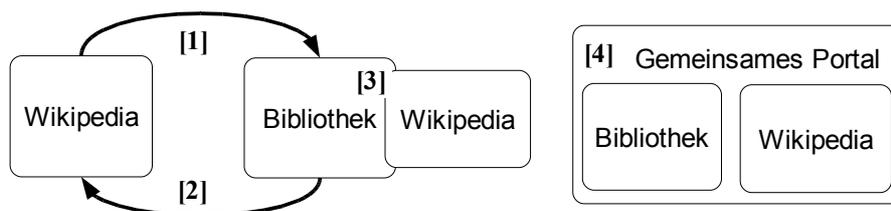
<sup>14</sup> <<http://de.wikipedia.org/wiki/Wikipedia:PND/Fehlermeldung>> (15.6.2006).

<sup>15</sup> <[http://de.wikipedia.org/wiki/Vorlage\\_Diskussion:SWD](http://de.wikipedia.org/wiki/Vorlage_Diskussion:SWD)> (15.6.2006).

Strukturell betrachtet gibt es vier grundsätzliche Möglichkeiten der Verknüpfung von Wikipedia und (digitalen) Bibliotheken (Abbildung 2).

1. Verlinkung von Wikipedia auf Bibliotheken (z.B. über Normdaten auf Kataloge)
2. Verlinkung von Bibliotheken auf Wikipedia (z.B. Linksammlungen und Kataloge)
3. Integration von Wikipedia-Inhalten in Bibliotheken (z.B. Fachportale und Kataloge)
4. Integration von Bibliotheken und Wikipedia in einem Portal (z.B. Suchmaschine)

Zusätzlich wäre zumindest theoretisch die Einbindung von Bibliotheken in Wikipedia möglich. Da Wikipedia jedoch nicht anstrebt eine Bibliothek, sondern eine Enzyklopädie zu sein und da keine dynamischen Inhalte wie zum Beispiel Suchformulare in Wikipedia möglich sind, ist eher die Zusammenführung unter einem gemeinsamen Portal wahrscheinlich. Eine solche Zusammenführung könnte auch von einem Drittanbieter wie zum Beispiel einer Suchmaschine betrieben werden.



**Abbildung 2:** Möglichkeiten der Verknüpfung von Wikipedia und Bibliotheken

Die Verknüpfung wird vor allem durch offene Schnittstellen und Standards befördert. So existiert zwar mit Z39.50 ein Protokoll zur Abfrage von Bibliotheken und anderen Informationssystemen – die Anwendung ist allerdings komplex und viele Einrichtungen stellen keinen offenen Zugang bereit. Möglicherweise werden auf XML basierende Weiterentwicklungen wie die ZING-Initiative<sup>16</sup> diese Situation verbessern. Ein anderer Standard ist das OAI-Protokoll<sup>17</sup> zur Sammeln und Verteilen von Metadaten, das sich im Bereich des Open Access bereits durchgesetzt hat. Noch einfacher sind so genannte Search-Plugins, die relativ leicht zu programmieren sind und in den eigenen Browser oder Suchmaschinen eingebunden werden können. Weitere Ansätze zur Bereitstellung von offenen Schnittstellen zum Information Retrieval sind OpenSearch von Amazon<sup>18</sup> und Google Co-op.<sup>19</sup> Gemeinsam ist diesen Anwendungen, dass das World Wide Web zunehmend als Plattform zur Ausführung von vernetzten Anwendungen genutzt wird – ein Paradigma, das im Zusammenhang mit dem Begriff Web 2.0 an Bedeutung gewinnt. Angelehnt an Web 2.0 wird seit Ende 2005 in der englischsprachigen Bibliothekswelt unter der Bezeichnung Library 2.0 eine grundlegende Neuausrichtung von Bibliotheken im digitalen Zeitalter diskutiert. Sarah Houghton schreibt dazu:

„Library 2.0 simply means making your library's space (virtual and physical) more interactive, collaborative, and driven by community needs. Examples of where to start include blogs, gaming nights for teens, and collaborative photo sites. The basic drive is to get people back into the library by making the library relevant to what they want and need in their daily lives...to make the library a destination and not an afterthought.“<sup>20</sup>

Daran ist prinzipiell nichts neues aber angesichts der technischen Möglichkeiten und konkurrierenden Angeboten im Internet fällt der Grad an Zusammenarbeit und Interaktivität in digitalen Bibliotheken noch eher bescheiden aus. Wikipedia und verwandte Projekte können bei sinnvoller Nutzung als wesentlicher Baustein einer ‚Bibliothek 2.0‘ dienen den offenen Charakter Sozialer Software in Bibliotheken bringen. Ein wesentliches Mittel dafür, dass die Zusammenarbeit nicht in Beliebigkeit ausartet, sondern in die vorhandene Erschließungsarbeit eingebunden werden kann, sind Normdaten.

### 3.3. Die Verwendung von Normdaten zur eindeutigen Referenzierung

<sup>16</sup> <<http://www.loc.gov/standards/sru/>> (15.6.2006).

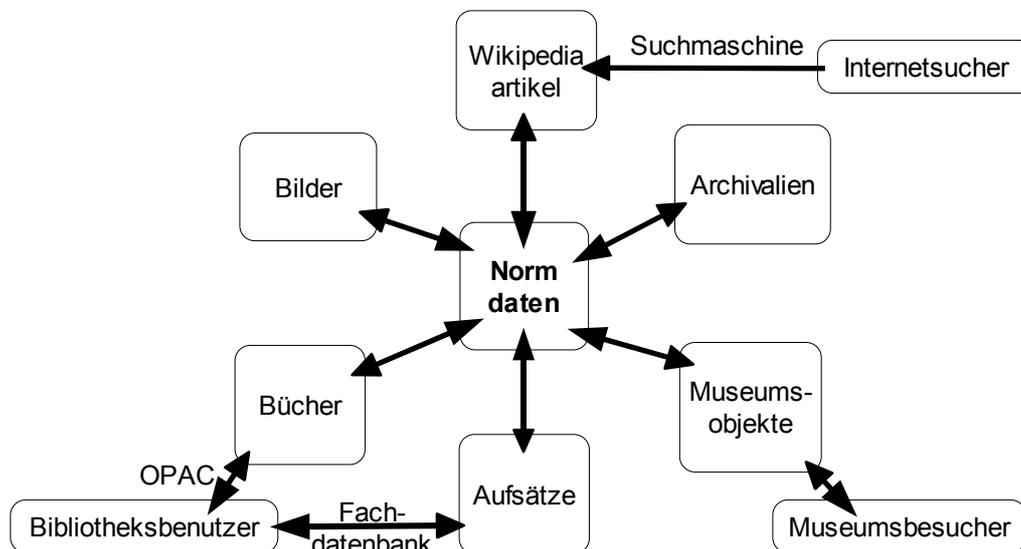
<sup>17</sup> <<http://www.openarchives.org/OAI/openarchivesprotocol.html>> (15.6.2006).

<sup>18</sup> <<http://opensearch.a9.com/>> (15.6.2006).

<sup>19</sup> <<http://www.google.com/coop>> (15.6.2006).

<sup>20</sup> <[http://librarianinblack.typepad.com/librarianinblack/2005/12/library\\_20\\_disc.html](http://librarianinblack.typepad.com/librarianinblack/2005/12/library_20_disc.html)> (15.6.2006).

Normdaten dienen der eindeutigen Referenzierung von Objekten und Konzepten und sind Teil der bibliothekarischen Infrastruktur. Die Hauptaufgabe von Normdaten ist die Auflösung der Homonymproblematik, d.h. Es werden Mehrdeutigkeiten und Unklarheiten aufgrund gleicher Namen vermieden. Damit ermöglichen Normdaten einen wesentlichen Mehrwert, den Volltextsuchmaschinen nicht bieten können. Die gemeinsame Verwendung von Normdaten ermöglicht es zudem, verschiedene Datensätze eindeutig aufeinander abzubilden: Über Hyperlinks können Internetsucher so beispielsweise von Wikipedia-Artikeln über die gemeinsam verwendeten PND-Nummern zu Katalogen in Bibliotheken, Museen und Archiven geleitet werden (Abbildung 3).



**Abbildung 3:** Normdaten ermöglichen vielfältige Verknüpfungen

Die Verknüpfung von Wikipedia-Artikeln mit Normdateien kann im Falle der Wikipedia auch als Mapping zwischen zwei kontrollierten Vokabularen angesehen werden. Die Verwendung von Wikipedia zur Erschließung ist bereits an verschiedener Stelle vorgeschlagen worden (Voss, 2006a; Hepp, 2006); im Rahmen des Semantic Web könnte Wikipedia bei einem Problem helfen, das in der Forschung zur Zeit unter der Bezeichnung ‚Identity and Reference‘ diskutiert wird (Halpin, 2006).<sup>21</sup> Zwar gibt es schon eine wachsende Zahl von Ontologien, die Klassen und Relationen festlegen (z.B. Dublin Core, MODS, FOAF etc.), aber es mangelt an Identifikatoren für eine Vielzahl von konkreten Objekten und Konzepten. Es ist zwar vorgesehen, dass Objekte im Semantic Web grundsätzlich durch URIs referenziert werden, aber wer diese URIs festlegen soll und wie sich Informationen zu vergebenen URIs ermitteln lassen sollen, ist eher unklar.<sup>22</sup> Für die Normdatei der Library of Congress existiert bereits ein URI-Namensraum nach dem von der National Information Standards Organization (NISO) verwalteten info-URI Schema:<sup>23</sup> So lässt sich beispielsweise Bertrand Russel bzw. sein Datensatz im Library of Congress Name Authority File (LCNAF) durch die URI `info:lcna/n79056054` identifizieren. Allerdings besteht keine einfache Möglichkeit, den Datensatz über die URI aufzurufen, weshalb sich das gesamte URI-System bislang nicht wirklich durchgesetzt hat. Ein pragmatischer Vorschlag (Pepper, 2006) für die Lösung dieser Probleme unter dem Namen ‚Public Resource Identifiers‘ (PRI) besteht darin, als URI ganz einfach eine URL zu verwenden, die speziell zum Zweck der Identifikation angelegt wurde und beim Aufruf im Webbrowser eine Beschreibung des identifizierten Objektes liefert. Dies ist bereits für alle Wikipedia-Artikel der Fall. Auch die Normdateien der DDB können bereits über eine URL angesprochen werden, jedoch wäre eine offizielle, dauerhafte Form bspw. <http://authorities.ddb.de/pnd/118601121> der bisherigen URL <http://dispatch.opac.ddb.de/DB=4.1/PPN?PPN=118601121> vorzuziehen.

<sup>21</sup> Siehe dazu auch den Workshop *Identity, Reference, and the Web* (IRW2006) auf der World Wide Web Conference 2006 (WWW2006) <<http://www.ibiblio.org/hhalpin/irw2006/>> (15.6.2006).

<sup>22</sup> Vorhandenen Erschließungssysteme aus dem Bibliotheks- und Dokumentationsbereich werden von Seiten der Semantic-Web-Community weitgehend ignoriert, obwohl sie gerade im Bereich Identity und Reference von Nutzen wären. Allerdings ist daran auch das Bibliotheks- und Informationswesen – zum eigenen Nachteil – nicht ganz unschuldig.

<sup>23</sup> <<http://info-uri.info/>> (15.6.2006).

#### 4. FAZIT

Wikipedia und Bibliotheken sind beide Einrichtungen, die es sich zum Ziel gesetzt haben, Wissen zu sammeln, zu ordnen und verfügbar zu machen. Während Bibliotheken publizierte Informationen sammeln und erschließen, werden die Inhalte der Wikipedia direkt und gemeinsam von ihren Lesern erstellt. Dabei herrscht eine selbstorganisierte Arbeitsweise vor, die tausende von Mitarbeitern zur freiwilligen Beteiligung motiviert; trotz dieser scheinbar chaotischen, wenig planbaren Entwicklung, hat es Wikipedia zu erstaunlichem Umfang und Beliebtheit gebracht. Wesentliche für Wikipedia ist auch, dass ihre Inhalte unter den Bedingungen der GFDL frei verwendet werden können, und somit für verschiedene Anwendungen bereitstehen. Wie anhand grundsätzlicher Möglichkeiten dargestellt wird, können Bibliotheken und Wikipedia sowohl miteinander verlinkt als auch in neue Systeme integriert werden. Der seit etwa einem Jahr praktizierte Einsatz der PND ist ein hier ein erfolgreiches Beispiel. Als Bindeglied bei der technischen Verknüpfung von Wikipedia und Bibliotheken dienen vor allem Normdaten, die zur Identifizierung von Objekten und Konzepten zusammen mit Wikipedia-Artikeln auch im Rahmen des Semantic Web von Bedeutung sein werden. Die Freigabe weiterer Daten und die Bereitstellung von offenen Schnittstellen wird es Nutzern zunehmend ermöglichen, nach dem Prinzip einer ‚library 2.0,‘ sich aktiv in die Verknüpfung von Wikipedia, Bibliotheken und anderen Informationsquellen einzubringen und damit eine freien Informationsinfrastruktur weiterzuentwickeln, in der Bibliotheken und Wikipedia weiterhin eine wichtige Rolle besitzen werden.

#### 5. LITERATUR

Czech, Jürgen; Pfeiffer, Barbara (2006): *Wikipedia und Personennamendatei*. Bibliothekartag 2006.

Danowski, Patrick; Voss, Jakob (2004): Bibliothek, Information und Dokumentation in der Wikipedia. *Information: Wissenschaft & Praxis*, 55, 2004 <<http://eprints.rclis.org/archive/00002566/>> (15.6.2006)

Fiebig, Henriette (2005): *Wikipedia – Das Buch*. Directmedia, 2005.

Free Software Foundation (2002): GNU Free Documentation License. <<http://www.gnu.org/copyleft/fdl.html>> (15.6.2006).

Frost, Ingo (2006): *Zivilgesellschaftliches Engagement in virtuellen Gemeinschaften. Eine systemwissenschaftliche Analyse des deutschsprachigen Wikipedia-Projektes*. Herbert Utz Verlag, 2006 (Diplomarbeit an der Universität Osnabrück).

Halpin, Harry (2006): *Identity, Reference, and Meaning on the Web*. Workshop at the Word Wide Web Conference 2006 <<http://www.ibiblio.org/hhalpin/irw2006/hhalpin.html>> (15.6.2006).

Hepp, Martin; Bachlechner, Daniel; Siorpaes, Katharina (2006): *Harvesting Wiki Consensus - Using Wikipedia Entries as Ontology Elements*. SemWiki2006 Workshop <<http://www.heppnetz.de/files/SemWiki2006-Harvesting%20Wiki%20Consensus-LNCS-final.pdf>> (15.6.2006).

Pepper, Steve (2006): *The Case for Published Subjects*. Workshop at the Word Wide Web Conference 2006 <<http://www.ibiblio.org/hhalpin/irw2006/spepper2.pdf>> (15.6.2006).

Sanger, Larry (2004): *Why Wikipedia Must Jettison Its Anti-Elitism*. Kuro5hin, 31. Dezember 2004 <<http://www.kuro5hin.org/story/2004/12/30/142458/25>> (15.6.2006).

Voss, Jakob (2005): *Metadata with Personendaten and beyond*. Proceedings of the first Wikimania Conference, 2005. <<http://meta.wikimedia.org/wiki/Transwiki:Wikimania05/Paper-JV2>> (15.6.2006).

Voss, Jakob (2006a): *Collaborative Thesaurus tagging the Wikipedia way*. 10. April 2006 <<http://arxiv.org/abs/cs/0604036>> (15.6.2006).

Voss, Jakob (2006b): *Gemeinschaftliche Schreibprozesse in Wikipedia*. Tagungsband der Clio-Online-Tagung .hist 2006 (erscheint Juli 2006).