

Renate Gömpel, Ulrike Junger, Elisabeth Niggemann

Veränderungen im Erschließungskonzept der Deutschen Nationalbibliothek

Erweiterter
Sammelaufrag

Mit Inkrafttreten des Gesetzes über die Deutsche Nationalbibliothek (DNBG) am 23. Juni 2006 ist ein neuer Auftrag zu den bisherigen Aufgaben hinzugekommen: Die Sammlung, Erschließung, Archivierung und Bereitstellung von Online-Publikationen. Die Bedingungen, unter denen diese gesammelt werden, und die Einschränkungen, die sich durch automatisierte Sammelverfahren ergeben, werden in der Pflichtablieferungsverordnung – PflAV vom 17. Oktober 2008 konkretisiert.

Bisher kamen bis zu 20.000 Online-Publikationen jährlich in den Bestand der Deutschen Nationalbibliothek (DNB). Da mittlerweile die erforderlichen technischen Verfahren so weit entwickelt sind, dass künftig große Mengen an Online-Publikationen verarbeitet werden können, ist für das laufende Jahr mit einem enormen Mengenzuwachs zu rechnen. Es ist zu erwarten, dass das Volumen an Online-Publikationen auch zukünftig weiter anwachsen wird, es erreicht bereits heute eine völlig neue Größenordnung im Vergleich zu den traditionellen Medien: Im August 2009 ist die 13-millionste .de-Domain bei der deutschen Domainverwaltung DENIC registriert worden. Pro Jahr wächst die deutsche Top Level Domain (TLD) um rund eine Mio. Einträge. Die Musikindustrie schätzt die Menge der Musik-Onlinepublikationen auf über eine Mio., Tendenz stark steigend! Bei den Book-on-Demand-Verlagen wird über Produktionszahlen im sechsstelligen Bereich gesprochen. Auf den von DINI (Deutsche Initiative für Netzwerkinformationen e. V.) gelisteten deutschen Dokumenten- und Publikationsservern lagern ebenfalls Online-Publikationen im sechsstelligen Bereich.

Enormer
Zuwachs bei
Netzpublikationen
zu erwarten

Das massive Anwachsen der Online-Publikationen hat bislang jedoch nicht zu einem Rückgang bei den traditionellen Medien geführt. Auch hier sind in den vergangenen Jahren die Zugangszahlen immer noch gestiegen und haben im Jahr 2009 wieder die Marke von 300.000 Einheiten überschritten. Darüber hinaus sieht sich die DNB

vor der Notwendigkeit, nutzerorientiert auch auf Einzeldokumentebene (Zeitschriftenaufsatz, CD-Track) zu erschließen.

Diesen gestiegenen Anforderungen stehen nicht im gleichen Maße gestiegene Personalressourcen gegenüber. Es ist klar absehbar, dass allein die Bearbeitung der Online-Publikationen aufgrund ihrer schiereren Menge nicht mit den etablierten Arbeitsverfahren und Werkzeugen der Erschließung bewältigt werden kann. Aber nicht nur für die Sammlung und Erschließung von Online-Publikationen muss über neue Wege nachgedacht werden, sondern es müssen auch Konsequenzen für die Bearbeitung der traditionellen Medien gezogen werden, um zeitliche Freiräume zu schaffen, die für die Entwicklung, Einführung und Bewertung neuer Verfahren eingesetzt werden können.

»Neue Wege« bedeutet zunächst, dass für die Bearbeitung der Online-Publikationen überwiegend automatisierte Verfahren eingesetzt werden, sowohl für den Ingest als auch für die Indexierung und Erschließung, die Bereitstellung und Langzeitarchivierung. Das gilt sowohl für gezielt gesammelte Netzpublikationen als auch für das Web-Harvesting.

Für die Erarbeitung und Etablierung automatischer Erschließungsverfahren für Online-Publikationen hat die DNB das interne Projekt PETRUS (Prozessunterstützende Software für die digitale Deutsche Nationalbibliothek) aufgelegt. Zurzeit wird an den folgenden vier Themen gearbeitet:

- der automatischen Vergabe von DDC-Sachgruppen,
- der automatischen Übernahme von inhaltserschließenden Daten bei so genannten Parallelausgaben,
- der automatischen Verknüpfung von Personennamen in Dokumenten mit Normdatensätzen aus der Personennamendatei (PND) sowie
- der automatischen Beschlagwortung.

Nähere Informationen zum Projekt PETRUS finden Sie auf den vorhergehenden Seiten in diesem Heft.

Intellektuelle
Erschließung von
Netzpublikationen
ist mengenmäßig
nicht mehr reali-
sierbar

Automatisierte
Verfahren sind
notwendig

Die DNB stellt sich den skizzierten Herausforderungen mit einer Veränderung und Neuausrichtung ihres Erschließungskonzepts.

Sie wird auch weiterhin alle Dokumente mit formal- und inhaltserschließenden Daten versehen. Dabei strebt die DNB an, den Umfang an maschinell oder halbautomatisch erzeugten Daten zu erhöhen und solche Daten, die auch aus anderen Quellen stammen können, konsequent an ihre Datenbezieher weiterzugeben.

Seit Anfang 2010 kein Autopsieverfahren für Netzpublikationen

Die DNB liefert seit Anfang 2010 keine intellektuell und per Autopsie erstellten Formal- und Inhaltserschließungsdaten für Online-Publikationen mehr.¹⁾ Es werden – wo vorhanden – Fremddaten weitergegeben, und es erfolgt eine automatische Erschließung²⁾. Die Netzpublikationen durchlaufen einen gestuften Geschäftsgang zur Erschließung.

Die Metadaten zu den Netzpublikationen werden aus unterschiedlichen Quellen erzeugt: Für geharvestete Netzpublikationen ohne vorhandene Metadaten werden Metadaten aus dem Volltext generiert; von Verlagen, Bibliotheken oder Urhebern gemeldete oder bereitgestellte Metadaten werden übernommen.

Die in den Metadaten genannten Personennamen werden automatisch mit der PND abgeglichen und es werden automatisch neue Normdatensätze angelegt, sofern der Abgleich kein Ergebnis erbringt. Wie erwähnt, werden im Projekt PETRUS entsprechende automatisch unterstützte Abgleichs- und Verknüpfungsverfahren entwickelt.

Alle Netzpublikationen sollen mit einer DDC-Sachgruppe versehen werden; dies soll zukünftig allerdings automatisch erfolgen. Sind bei Parallelausgaben weitere Sacherschließungsdaten (wie Schlagwörter, DDC-Notationen) vorhanden, so werden diese zu den Metadaten der Online-Publikationen transferiert und ebenfalls ausgeliefert.



Digitalisate brauchen Publikum

Damit Ihre wertvollen Digitalisate nicht ungesehen bleiben, kümmern wir uns um den Weg zum Nutzer. Seit mehr als 14 Jahren entwickeln und gestalten wir erfolgreiche Internetauftritte, sorgen für die richtige Präsentation digitaler Inhalte und entwickeln vielschichtige webbasierte Instrumente für die Online-Arbeit.

Konzeption // Beratung // Gestaltung // Redaktion
 Programmierung // Hosting // Social Media

3pc
 Neue Kommunikation

3pc.de · info@3pc.de
 3-point concepts GmbH

Für die traditionellen Medien wird die DNB verstärkt Fremddaten nachnutzen und diese gegebenenfalls anreichern. Grundsätzlich werden die traditionellen Medien wie gewohnt in vollem Umfang formal nach RAK (bzw. nach dem Umstieg nach RDA) erschlossen einschließlich der erforderlichen Normdatenverknüpfungen. Von dieser Regelung ausgenommen werden Belletristik und Kinder- und Jugendliteratur. Für diese Bereiche werden Fremddaten genutzt, die nicht intellektuell auf Regelkonformität überprüft, für die aber eine intellektuelle Qualitätssicherung für die Personennamen und die Titelfassungen gewährleistet werden soll.

Für die Inhaltserschließung wird das jetzt gültige Konzept (Sachgruppenvergabe für alle Dokumente, verbale Sacherschließung für Werke der Bibliografierihe A, klassifikatorische Feinerschließung mit DDC für die Bibliografierihen A, B, H) dahingehend verändert, dass Schulbücher seit Anfang 2010 nur noch eine Sachgruppe bekommen.

In der Reihe C verzeichnete Karten werden seit Anfang 2010 nach den DDC-Sachgruppen gegliedert und verbal erschlossen.

Bei den DDC-Sachgruppen gab es folgende Veränderungen, die ebenfalls seit Anfang 2010 wirksam sind: Die Sachgruppe 640 wurde umbenannt in »Hauswirtschaft und Familienleben«, die Sachgruppe 670 in »Industrielle und handwerkliche Fertigung«. Neu eingeführt wurde die Sachgruppe 333.7 »Natürliche Ressourcen, Energie und Umwelt«. Damit wurde einem lange gehegten Wunsch Rechnung getragen, Werke zu dieser wichtigen Thematik in einer eigenen Sachgruppe auszuweisen. Im Zuge der Umsetzung des neuen Bibliografiekonzeptes der DNB sind die Sachgruppen »Y/Abgeschlossene Veröffentlichungen« und »Z/Berichtigungen« entfallen.

Belletristische Werke sollen künftig mit Genrebegriffen versehen werden; dabei werden die so

genannten VLB-Warengruppen zugrunde gelegt. Es ist geplant, im 2. Quartal 2010 mit der Erzeugung und Auslieferung entsprechender Daten zu beginnen. Deutschsprachige Originalliteratur wird weiterhin außerdem motivisch mit Schlagwörtern erschlossen.

Ebenfalls ab dem 2. Quartal 2010 wird die DNB in der intellektuellen verbalen Erschließung den neuen Standard umsetzen, wie er durch die Überarbeitung der RSWK-§§ 13 - 15 vorgegeben ist. Das heißt, dass im Regelfall der Inhalt eines Dokumentes in einer Schlagwortfolge wiedergegeben wird. Die intellektuell vergebenen Permutationsmuster werden abgeschafft.

Darüber hinaus wird die verbale Erschließung in den naturwissenschaftlich-technischen Fächern in diesem Jahr daraufhin untersucht, ob sie nicht auf eine ressourcensparende Weise modifiziert werden kann. Dies könnte als Einstieg in ein fachliches oder an Literaturgruppen orientiertes Erschließungskonzept dienen.

Die DNB wird so umfassend wie möglich Tonträger und Musikalien intellektuell nach Autopsie erschließen. Sie bereitet darüber hinaus eine Kooperation zur Übernahme von Fremddaten vor und wird die übernommenen Daten weitergeben.

Für alle Neuerscheinungen der deutschen Verlage mit Ausnahme der Belletristik und der Kinder- und Jugendliteratur werden Zusatzinformationen wie z. B. Inhaltsverzeichnisse, Klappentexte, Coverscans etc. zu den Titeln mitgeliefert. Die Zusatzinformationen werden außerdem indiziert und als Sucheinstiege bereitgestellt.

Im Ergebnis strebt die DNB die wesentlich schnellere Bereitstellung von erschließenden Metadaten für sehr viel größere Dokumentmengen an - und hofft auf Zusammenarbeit mit und Rückmeldung von ihren Partnern und Kunden.

Genrebegriffe für belletristische Werke

Verstärkte Nutzung von Fremddaten

Veränderungen bei den DDC-Sachgruppen

Fazit

Anmerkungen

1 Die bibliografischen Daten zu den Online-Publikationen werden in einer neuen Bibliografierihe, der Reihe O, kostenlos bereitgestellt.

2 Der Erschließungsvorgang kann dabei bei einer Verbesserung der eingesetzten Verfahren wiederholt werden. Das bedeutet, dass ggf. mehrfach jeweils verbesserte Metadaten für ein- und dieselbe Publikation geliefert werden.