

Cornelia Diebel

Netzpublikationen – Sammlung, Archivierung und Bereitstellung in der Deutschen Nationalbibliothek

Einleitung

Die Sammlung von Pflichtexemplaren in Bibliotheken existiert seit dem 16. Jahrhundert und ist auch in Deutschland seit dem 17. Jahrhundert bekannt. Wurden die Pflichtexemplare zunächst als Abgeltung für Druckprivilegien oder für Zensurzwecke genutzt, ist seit Anfang des 19. Jahrhunderts der Hauptzweck die möglichst lückenlose Sammlung des geistigen Schaffens eines definierten Einzugsbereichs. Anfangs bezog sich die Sammlung von Pflichtexemplaren hauptsächlich auf das veröffentlichte Schrifttum eines Landes, im Laufe der technischen Innovationen wurden auch die neuen Medienträger in den Sammelauftrag integriert.¹⁾ Für Nationalbibliotheken existiert damit seit langer Zeit der Auftrag, die nationale Literaturproduktion zu sammeln, zu katalogisieren, langfristig zu erhalten und dadurch das jeweilige nationale Kulturerbe zu bewahren.

Um das Ziel zu erreichen, Gegenwärtiges zu sammeln, bezieht die Pflichtabgabe von gedrucktem Kulturgut in vielen Ländern inzwischen elektronische Publikationen mit ein; denn die Bedeutung und die Funktion der Sammlung würden verloren gehen, würde das geistige Schaffen nicht umfassend durch Einbeziehung elektronischer Medien dokumentiert werden.

Elektronische Publikationen haben vielfältige Erscheinungsformen und stellen Bibliotheken im Allgemeinen, aber auch Pflichtexemplarbibliotheken vor besondere Herausforderungen. Neben dem Bereich der elektronischen Publikationen, die weiterhin den körperlichen Medien zuzuordnen sind wie CDs oder DVDs, gehören auch die sogenannten Netzpublikationen zum Bereich der elektronischen Publikationen. Netzpublikationen sind elektronische Publikationen, die nicht auf einem physischen Datenträger, sondern in der Regel über

das Internet angeboten werden. Die Übernahme dieses neuen Medientyps in den Bestand der Bibliothek muss organisiert und dauerhaft betrieben werden. Im Bereich der traditionellen Materialien existieren ausdifferenzierte, an bibliothekarischen Standards orientierte Bearbeitungsprozesse, deren Übertragbarkeit und Nützlichkeit im Bereich der Netzpublikationen infrage gestellt werden müssen. Die grundsätzliche Frage nach adäquaten Geschäftsprozessen für diesen Objekttyp stellt sich im gesamten bibliothekarischen Umfeld. Durch den im Bereich der Neuzugänge besonders auf Langfristigkeit und Langzeitverfügbarkeit der gesammelten Medien ausgerichteten Sammelauftrag der Pflichtexemplarbibliotheken hat die Frage nach tragbaren Geschäftsprozessen für diese einen anderen Schwerpunkt. In der Folge werden sich Teile von Geschäftsprozessen für Netzpublikationen in Pflichtexemplarbibliotheken von denen in anderen Bibliotheken unterscheiden.

Sammlung

Die Sammlung von Netzpublikationen in der Deutschen Nationalbibliothek (DNB) erfolgt seit Anfang der 2000er-Jahre, damals auf freiwilliger Basis in Zusammenarbeit mit größeren, wissenschaftlichen Verlagen und in enger Kooperation mit der deutschen Hochschulcommunity.

Bereits in damaligen Workflows zur Annahme, Speicherung und Nutzung waren Methoden der automatischen Informationsverarbeitung etabliert, die Be- und Verarbeitung von Netzpublikationen war aber stark an traditionellen bibliothekarischen Standards, wie z. B. der umfassenden Verzeichnung der Objekte nach dem gültigen Katalogisierungsregelwerk orientiert. Damalige bereitgestellte Schnittstellen zur Ablieferung der elektronischen

Langzeitverfügbarkeit bei Pflichtexemplarbibliotheken
Thema

Sammlung zunächst auf freiwilliger Basis

Pflichtabgabe elektronischer Publikationen

Erweiterung des Sammelauftrags 2006

Objekte erforderten manuelle Interaktion durch bibliothekarisches Personal, Ablieferungsformulare waren sehr ausführlich und für nichtbibliothekarische Expertinnen und Experten kaum verständlich und nachvollziehbar.

Im Jahr 2006 wurde der gesetzliche Sammelauftrag der Deutschen Nationalbibliothek um den Aspekt der Sammlung von unkörperlichen Medienwerken, also Netzpublikationen, erweitert.²⁾ Diesen Zeitpunkt nutzte die DNB dafür, ihre bisherigen Verfahrensweisen in Bezug auf Netzpublikationen zu analysieren und zu überprüfen. Bereits 2006 war absehbar, dass die Anzahl der online publizierten Objekte, die durch die DNB zu sammeln sind, sehr hoch werden wird. Ausgehend von dieser Einschätzung wurde deutlich, dass die Aufgabe der Sammlung, Verzeichnung, Bereitstellung und dauerhaften Archivierung von Netzpublikationen nur mit umfassend automatisierten Prozessen zu erledigen sein wird, zumal die Zahl der zu sammelnden traditionellen Materialien absehbar nicht zurückgehen würde.

Automatisierte Verfahren und Prozesse notwendig

Die Entscheidung, bei der Sammlung, Archivierung und Bereitstellung von Netzpublikationen zukünftig auf automatisierte Verfahren und Prozesse zu setzen, stand damit fest und wird seitdem konsequent umgesetzt. Zusätzlich zum Bekenntnis zur Arbeit mit automatischen Workflows war bereits 2006 klar, dass es sinnvoll ist, die Aufgabe – und damit zusammenhängend den Aufbau der benötigten technischen Infrastruktur – in Etappen einzuteilen und nicht zu versuchen, alle anstehenden Aufgaben zeitgleich zu erledigen. Die zweite Linie, der die DNB seitdem folgt, ist die des schrittweisen Vorgehens. Konkret bedeutet dies, dass die Sammlung sukzessive sowohl um zu sammelnde Objektgruppen als auch um technische Verfahren erweitert wird.

Schrittweises Vorgehen

Im Gegensatz zu einigen anderen Nationalbibliotheken hat die DNB beim Aufbau der Sammlung von Netzpublikationen nicht mit der Webarchivierung begonnen, sondern mit der Sammlung von monografischen Objekten, vor allem Online-Hochschulschriften und zusätzlich E-Books, vornehmlich aus dem kommerziellen Verlagsbereich. Dieses Vorgehen lässt sich durch die besondere Nähe der DNB zum Verlagsbuchhandel begründen, denn die DNB wurde 1912 als Deutsche Bücherei auf Initia-

tive des Börsenvereins des Deutschen Buchhandels gegründet.

Der nächste Entwicklungsschritt war die Erweiterung der Sammlungsmöglichkeiten auch für fortlaufende Objekte wie E-Journals und E-Paper. Es folgte dann auch der Einstieg in die Webarchivierung, die inzwischen mit einem selektiven Ansatz umgesetzt wurde. Zuletzt ist die Sammlung der herunterladbaren Hörbücher umgesetzt worden.

Aktuell sammelt die DNB damit folgende Objektgruppen:

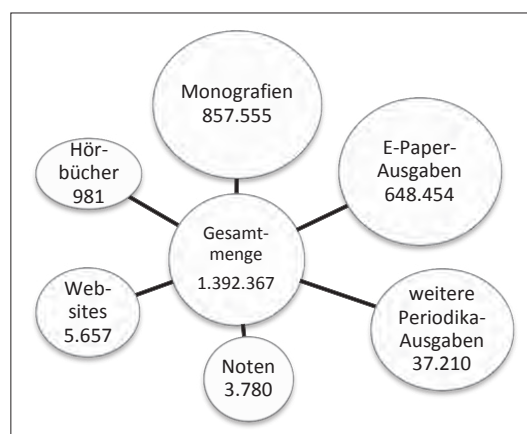
- E-Books
- E-Journals
- E-Paper
- Online-Hochschulschriften
- Noten
- Websites
- herunterladbare Hörbücher

Erweiterung auf fortlaufende Objekte

Zahlen, Daten, Fakten

Der Gesamtbestand der Netzpublikationen in der DNB beträgt am 31.12.2014 1.549.857 Objekte. Die folgende Grafik zeigt eine Übersicht über die Verteilung der oben genannten Objekttypen. Der tägliche Zugang steigt ständig an und liegt zurzeit bei rund 1.400 Objekten täglich. Auch für die kommenden Jahre ist mit einem weiteren Anstieg zu rechnen.

Zugang: 1.400 Objekte täglich



Mengenübersicht nach Objektgruppen, Stand 31.12.2014

Neben der allgemeinen und umfassenden Sammlung von Netzpublikationen ist die Sammlung der

E-Paper

sogenannten E-Paper-Ausgaben besonders erwähnenswert. Zunehmend vertreiben die Zeitungsverlage nicht nur gedruckte Tageszeitungen, sondern stellen zusätzlich layoutgetreue digitale Ausgaben – sogenannte E-Paper – für ihre Abonnenten ins Internet. Seit 2010 sammelt die Deutsche Nationalbibliothek diese E-Paper-Ausgaben von Tageszeitungen.³⁾

Die aktuelle Zeitungssammlung der DNB basiert nahezu vollständig auf E-Paper, nur noch rund 70 Zeitungen in Deutschland haben keine solche Ausgabe. Diese Zeitungen werden als Mikrofilm gesammelt; es erfolgt keine Archivierung von Papierausgaben.

Der aktuell im Archivsystem der DNB belegte Speicherplatz für die Sammlung von Netzpublikationen liegt bei 25 TB. Der Gesamtspeicher für elektronische Objekte ist damit allerdings nicht beschrieben, denn zusätzlich verwaltet die DNB auch eigene digitalisierte Objekte, insgesamt liegt der Speicherbedarf zurzeit bei netto 177 TB. Dies beinhaltet zu knapp 52 Prozent migrierte Audio-CDs, 32 Prozent Digitalisate von Printmedien (Bücher, Zeitschriften, Plakate, Archivalien), 15 Prozent Netzpublikationen (E-Books, E-Paper, Hörbücher) und 1 Prozent digitalisierte Inhaltsverzeichnisse.

Automatische Verfahren

Die Entwicklung der automatischen Verfahren erforderte eine weitere grundlegende Ausrichtung, die bis heute gilt. Die Entscheidung, zunächst nur Objekte zu sammeln, bei denen die Produzenten in der Lage sind, bibliografische Metadaten mitzuliefern, wird in Zukunft zugunsten der automatischen Erstellung von bibliografischen Metadaten infrage zu stellen sein. Die aktuelle Festlegung wurde allerdings besonders im Jahr 2010 relevant, denn zu diesem Zeitpunkt entschied sich die DNB, keine intellektuelle Erschließung der Netzpublikationen mehr durchzuführen, sondern auch hier auf die Entwicklung und Nutzung von maschinellen Prozessen zu setzen. Mit diesem Hintergrund lassen sich zum heutigen Zeitpunkt folgende drei technische Rahmenbedingungen formulieren, die einer automatisierten Verarbeitung von Netzpublikationen in der DNB zugrunde liegen:

- Nutzung eines standardisierten Metadatenformats
- Ablieferung bekannter Dateiformate
- Bereitstellung der Objekte über standardisierte Schnittstellen.

Verlage und andere Produzenten und Ersteller von Netzpublikationen sind in der Nutzung von Metadatenformaten frei und können, wenn sie es denn möchten, eigene Strukturen entwickeln oder variable Tabellenformate wie csv-Dateien zur Verwaltung ihrer Daten einsetzen. Für die DNB war aber bereits zu Beginn der Sammlung klar, dass nur weit verbreitete und standardisierte Metadatenformate für die Ablieferung an die DNB genutzt werden können. Die Integration proprietärer und/oder selten genutzter Metadatenformate ist aus Aufwandssicht bei einer potenziell fünfstelligen Zahl von Anbietern nicht zu leisten und muss scheitern.

Aufgrund dieser Festlegung werden heute mehrere weit verbreitete Metadatenformate für die Ablieferung an die DNB akzeptiert. Dabei kann es auch zu unterschiedlichen Standards für verschiedene Objekt- oder Medientypen kommen, denn nicht für alle Medientypen gelten die gleichen Standards.

Für den Bereich der Ablieferung monografischer Medien (E-Books, E-Thesis, herunterladbare Hörbücher) werden gegenwärtig drei Metadatenformate von der DNB akzeptiert. Die hauptsächliche Ablieferung erfolgt im Metadatenformat ONIX 2.1.⁴⁾ Dieses Format ist bei deutschen Verlagen, aber auch international, das meistgenutzte Metadatenformat für bibliografische Informationen. Weiterhin ist die Ablieferung im bibliothekarischen Standardmetadatenformat MARC21⁵⁾ möglich. Zusätzlich zu diesen international genutzten und verbreiteten Metadatenformaten kann die Ablieferung über das Metadatenformat XMetaDissPlus⁶⁾ erfolgen. Dieses Metadatenformat ist eine eigenständige Entwicklung in Deutschland und kommt historisch gesehen aus den Anfängen der Ablieferung von E-Thesis (Online-Hochschulschriften); es wurde in der deutschen Hochschullandschaft ursprünglich als MetaDiss Ende der 1990er-Jahre entwickelt. Nach verschiedenen Weiterentwicklungsschritten kann es heute umfassend für die Ablieferung von Netzpublikationen eingesetzt werden und ist nach wie vor

Nutzung von standardisierten Metadatenformaten

Keine intellektuelle Erschließung

in der Hochschulschriftenablieferung das gängigste Metadatenformat.

Im Bereich der Ablieferung von Inhalten, die zu E-Journals gehören, existiert kein ähnlich einheitliches und weitverbreitetes Metadatenformat, auf das die Ablieferung spezialisiert werden könnte. Im deutschen Hochschulbereich kann XMetaDissPlus auch für die Ablieferung von Inhalten zu E-Journals genutzt werden, was auch zunehmend zum Einsatz kommt. Eine Lösung in MARC21 wurde in Zusammenarbeit mit dem Springer-Verlag entwickelt und kann generell für die Ablieferung von Artikeln oder Heften zu E-Journals verwendet werden. Mit Blick auf die Verlagslandschaft lässt sich aber feststellen, dass diese beiden Lösungen nicht so weit verbreitet sind, dass damit alle relevanten Ablieferergruppen und Verlage in die Lage versetzt werden könnten, ihrer Ablieferungspflicht nachzukommen. Aus diesem Grund strebt die DNB an, die Ablieferung über das JATS-Format⁷⁾ (vormals NLM Archiving and Interchange DTD) und über das Format von CrossRef⁸⁾ zu ermöglichen. Bei beiden Metadatenstandards existiert eine gewisse Verbreitung, und die Ablieferung hierüber vergrößert die Zahl der möglichen Ablieferer in Deutschland. Grundsätzlich kann die Ablieferung von Inhalten zu E-Journals über ganze Zeitschriftenhefte, aber auch über die Ablieferung einzelner Zeitschriftenartikel erfolgen. Um die Zuordnung der einzelnen Teile zum E-Journal zu erhalten, wird im Rahmen der Ablieferung eine Verbindung über die Auswertung eines eindeutigen Identifiers, der bei jeder Lieferung enthalten sein muss, gewährleistet.

Netzpublikationen kommen in den unterschiedlichsten Dateiformaten vor. Neben offenen Dateiformaten wie PDF oder EPUB sind auch geschlossene Formate wie das von Amazon verwendete AZW- oder das I-Book-Format von Apple auf dem Markt erhältlich. Die DNB beschränkt sich bei der Sammlung von Netzpublikationen im Sinne des schrittweisen Vorgehens aktuell auf die Sammlung von Netzpublikationen im PDF- oder im EPUB-Format. Um weitere Dateiformate zu sammeln, müssen einige weitere Voraussetzungen erfüllt werden. Auf der einen Seite muss die aktuelle Nutzbarkeit in den Lesesälen der DNB gesichert werden, hierzu muss die entsprechende Reader-Software zur Verfügung stehen. Auf der anderen Seite müssen alle Objekte

im Sinne der Langzeitarchivierung (LZA) analysiert und geprüft werden, nur so kann eine dauerhafte Qualität in der Bereitstellung sichergestellt werden. Bei der Übernahme der Netzpublikationen in das Bereitstellungssystem und im nächsten Schritt in das LZA-System werden im Workflow verschiedene technische Prüfungen durchgeführt, die als Risikoanalyse die Objekte beschreiben. Risikoanalyse ist hier in Bezug auf die Sicherstellung der zukünftigen Szenarien der Nutzung zu verstehen.⁹⁾

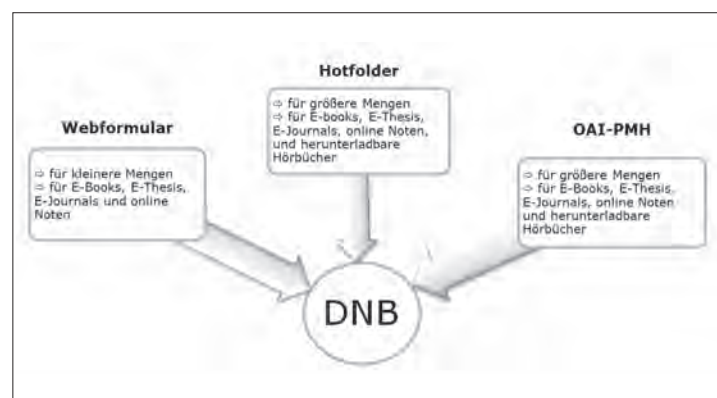
Die Zustellung der traditionellen körperlichen Pflichtexemplare läuft heute wie damals in der Regel über die Postzustellung. In Analogie könnte man überlegen, ob nicht die Zustellung von Netzpublikationen über E-Mails funktionieren könnte und ob dies als Lösung für die Übermittlung der zu sammelnden Objekte anzusehen ist.

Tatsächlich hat die DNB in den Anfangsjahren der Sammlung Netzpublikationen auch über ein E-Mail-Verfahren entgegengenommen. Dabei hat sich allerdings herausgestellt, dass trotz IT-gestützter Verfahren immer noch ein hoher Prozentsatz Handarbeit nötig war, um die Objekte zu archivieren, die Metadaten in den Katalog zu stellen und die Nutzung zu ermöglichen. Diese Vorgehensweise hat die DNB im Laufe der Zeit abgelöst und bietet dafür andere, vollständig automatisierte Schnittstellen an. Diese Schnittstellen wurden unter unterschiedlichen Aspekten entwickelt und können somit von unterschiedlichen Gruppen von Ablieferern verwendet werden. Die neu entwickelten Schnittstellen bedienen die Bedürfnisse der Ablieferer in gleicher Weise wie die Bedürfnisse der DNB.

Bereitstellung der Objekte über standardisierte Schnittstellen

Ablieferung bekannter Dateiformate

JATS-Format



Übersicht über die Ablieferungsschnittstellen

Eine der Schnittstellen ist für die Ablieferung von kleinen Mengen von Netzpublikationen geeignet. Hierfür wird ein Webformular eingesetzt, bei dem die Metadaten erfasst werden müssen und das Objekt hochgeladen oder per URL abgeholt wird.

Für sogenannte Massenablieferer stehen zwei Varianten zur Verfügung, bei denen auf eine manuelle oder intellektuelle Interaktion verzichtet werden kann. Eine sogenannte Push-Schnittstelle bietet die Möglichkeit, dass vom Ablieferer ausgehend die Objekte im Paket mit den Metadaten an einen von der DNB bereitgestellten Server übermittelt werden können.¹⁰ Zusätzlich existiert eine Pull-Schnittstelle, die über das von der Open Archiv Initiative bereitgestellte Protokoll für Metadatenharvesting¹¹ (OAI-PMH) funktioniert. Hierbei wird OAI-PMH dazu genutzt, Metadaten bei einem Client oder Repository anzufragen und über einen in den Metadaten bereitgestellten Link (Transfer-URL) das Objekt beziehungsweise eine Gruppe von Objekten abzuholen.

Die Schnittstellen werden von den Schedulingtools der DNB regelmäßig auf neu abgelieferte Objekte hin untersucht. Werden neue Objekte gefunden, werden diese von einem in der DNB entwickelten Importprozess verarbeitet. Hierbei kann es sehr unterschiedliche und individuelle Ablieferungsrythmen geben, die für jeden eigenen Task angelegt werden können. Hierzu findet in der Regel eine Absprache mit dem Ablieferer statt.

Mit aktuell über 700 regelmäßigen Abholungsprozessen wird die Push-Schnittstelle am häufigsten genutzt, um die Pflichtablieferung an die DNB zu erfüllen. Hierfür existieren bei der digitalen Verlagsauslieferung von E-Books bei Dienstleistern im kommerziellen Sektor ähnliche Verfahren, sodass diese Schnittstelle den Bedürfnissen des kommerziellen E-Book-Bereichs sehr entgegenkommt. Die OAI-PMH-Schnittstelle wiederum ist besonders im wissenschaftlichen und universitären Bereich verbreitet und wird vor allem für die Pflichtablieferung von Online-Hochschulschriften aller Art benutzt. Im Januar 2015 konnten über 65 aktive und regelmäßige Tasks gezählt werden, mit steigender Tendenz.

Allen Schnittstellen gemeinsam ist die Verarbeitung, die der Übernahme der Objekte folgt. Der technische Workflow stellt auf der einen Seite die

Extraktion, Konversion, Anreicherung und Einspielung der Metadaten in den Katalog der DNB sicher. Auf der anderen Seite werden die Objekte auf ein Repository übertragen und stehen von dort aus der Nutzung zur Verfügung.

Nach der Übernahme der Objekte in das Repository der DNB werden die Objekte zusätzlich in ein Langzeitarchiv (LZA) übernommen, das außerhalb der Räume der DNB in Göttingen betrieben wird.

Ergänzende automatische Verfahren, Bereitstellung und Nutzung

Neben der automatisierten Sammlung und Ablieferung sorgen weitere automatische Verfahren für eine Anreicherung der Daten im Katalogsystem der DNB. Nach dem Import der Metadaten wird im Katalogsystem ein Abgleich gestartet, dessen Aufgabe es ist, parallel erschienene Printausgaben zu den Netzpublikationen ausfindig zu machen und diese wechselseitig zu verknüpfen. Mit dieser Verknüpfung wird den Nutzenden ermöglicht zu entscheiden, welche Variante sie benutzen möchten. Zusätzlich werden neben der reinen Verknüpfung auch Daten zwischen den Katalogisaten ausgetauscht, z. B. inhaltserschließende Elemente; damit wird der inhaltliche Zugang zu den abgelieferten Netzpublikationen erweitert. Finden sich beim Abgleich der Metadaten keine Verknüpfungen zur Gemeinsamen Normdatei (GND)¹², werden bei Personennamen solche angelegt und die Namen in der GND ergänzt und damit auch hier der standardisierte Zugang über die Normdatei ermöglicht.¹³

Weiterhin sorgen eigenständige Prozesse für eine ergänzende automatische inhaltliche Erschließung. Zum einen werden für alle Netzpublikationen Sachgruppen nach der Dewey Decimal Classification (DDC-Sachgruppen) vergeben, die der Gliederung der Deutschen Nationalbibliografie mit etwas über 100 Sachgruppen entspricht. Zum anderen werden in einem weiteren eigenständigen Prozess Schlagwörter aus der bereits oben genannten GND automatisiert vergeben.¹⁴ Diese Anreicherungen der Daten dienen der Sicherstellung eines durchgängigen und einheitlichen inhaltlichen Zugriffs auf alle Objekte der DNB, auch für die nicht mehr intellektuell erschlossenen Netzpublikationen.

Freie Objekte
weltweit über
DNB-Katalog
nutzbar

Über den Katalog der DNB wird der Zugriff auf die Objekte ermöglicht. Netzpublikationen sind über den allgemeinen Portalzugang auffindbar, dafür werden sowohl die abgelieferten als auch die maschinell erzeugten Metadaten indexiert und suchbar gemacht. Der Katalog enthält alle Objekte, die die DNB sammelt und unterscheidet nicht nach Medien- oder Objektarten. Natürlich kann die Suche aber auf Netzpublikationen eingeschränkt werden. Der Zugriff auf die Netzpublikationen selbst entspricht immer den jeweiligen urheberrechtlichen Beschränkungen, denen die einzelnen Objekte unterliegen. Das bedeutet, dass freie Objekte auch weltweit über die DNB frei genutzt werden können. Das ist hauptsächlich für Objekte aus dem wissenschaftlichen und kulturellen Umfeld gegeben. Kommerzielle und damit in der Regel zugriffsbeschränkte Objekte werden zur Nutzung nur in den Lesesälen der DNB in Frankfurt am Main und Leipzig bereitgestellt. Die Nutzung der Objekte in den Lesesälen läuft über bereitgestellte Viewer an bibliothekseigenen Computern, eine Möglichkeit zum Download ist nicht gegeben.

Persistente
Adressierung

Die dauerhafte Adressierung und Zitierbarkeit der Netzpublikationen wird durch die konsequente und durchgängige Vergabe von Uniform Resource Names (URN) sichergestellt.¹⁵⁾ Alle von der Deutschen Nationalbibliothek archivierten Netzpublikationen erhalten eine URN aus dem Namensraum »urn:nbn:de«. Im Gegensatz zu URLs bieten URNs eine eindeutige und beständige Identifikation von Objekten, unabhängig vom Ort der Speicherung. Dies ist ein wichtiger Faktor zur Erhaltung der Langzeitverfügbarkeit. Ändert sich der Speicherort der Publikation, z. B. durch eine Serverumstellung, kann die Zugriffsadresse (URL), die einer URN hinterlegt ist, korrigiert werden. Dadurch behält die URN ihre Gültigkeit und verweist nach wie vor auf die entsprechende Publikation. Das Zitieren einer URN ist somit ein sicherer Weg, auch langfristig auf ein digitales Objekt zu verweisen. Objekte, die

bereits eine URN von anderen Organisationen bekommen haben, behalten diese. Objekte, die keine URN von ihrem Produzenten zugeteilt bekommen haben, bekommen diese aus dem Unternehmensraum der DNB zugeteilt.

Fazit und Ausblick

Die Zahl der bisher bereits von der DNB archivierten Netzpublikationen zeigt, dass die Konzentration auf die Etablierung automatischer Prozesse erfolgreich ist. Es scheint mit Blick auf die beschränkten personellen und finanziellen Ressourcen unrealistisch, dass diese Zahl allein durch konventionelle Erwerbung und Erschließung zu erreichen gewesen wäre. Aus diesem Grund wird die DNB diesen Weg weiter fortführen und ausbauen. Für Medientypen, die bisher noch nicht in angemessenem Umfang gesammelt werden, müssen angedachte und bereits laufende Verfahren vollständig umgesetzt werden, was z. B. für die Steigerung der Ablieferung von E-Journals gilt. Für Medientypen, die bisher noch gar nicht gesammelt werden, müssen Konzepte entwickelt und umgesetzt werden. Ein Beispiel wäre die Sammlung von Musik, die bisher noch nicht erfolgt. Ebenso muss die weitere technische Entwicklung auch im Bereich der E-Books verfolgt und in technische Erweiterungen umgesetzt werden. Hier sind z. B. Entwicklungen bei Apps, HTML-basierten E-Books und auch bei dynamischen Publikationen zu nennen.

Diese Entwicklungen sind nicht nur eine Herausforderung für die Seite der Sammlung, sondern für alle in diesem Aufsatz beschriebenen Aspekte der automatischen Erschließung, der Bereitstellung und nicht zuletzt der Langzeitarchivierung und damit der Erhaltung der Langzeitverfügbarkeit.

Die DNB hat begonnen sich diesen Herausforderungen zu stellen.

Automatische
Prozesse
fortführen und
ausbauen

Anmerkungen

- 1 Beger, Gabriele: Bibliotheksrecht. In: Die moderne Bibliothek. Ein Kompendium der Bibliotheksverwaltung / hrsg. von Rudolf Frankenberger und Klaus Haller. München: Saur, 2004, S. 377 ff.
- 2 Gesetz über die Deutsche Nationalbibliothek erreichbar unter: <<http://www.gesetze-im-internet.de/dnbg/index.html>>
- 3 Jockel, Kathrin; Schneider, Kurt: E-Paper-Sammlung in der Deutschen Nationalbibliothek – Eine Zwischenbilanz. In: Dialog mit Bibliotheken 26 (2014) 1, S. 21–24. – URN: <<http://nbn-resolving.de/urn:nbn:de:101-20140923125>>
- 4 <<http://www.editeur.org/15/Previous-Releases/>>
- 5 <<http://www.loc.gov/marc/>>
- 6 <<http://www.dnb.de/DE/Standardisierung/Metadaten/xMetadissPlus.html>>
- 7 <<http://jats.nlm.nih.gov/>>
- 8 <<http://www.crossref.org/>>
- 9 vgl. Schmitt, K.; Hein, S.: Risk Management for Digital Long-Term Preservation Services IPRES 2013. In: Proceedings of the 10th International Conference on Preservation of Digital Objects. 3.–5. September Lisbon – Portugal. ed. Jose Borbinha, Michael Nelson, Steve Knight, p. 314–317. – <http://purl.pt/24107/1/iPres2013_PDF/iPres2013-Proceedings.pdf>
- 10 Zur Objektstruktur vgl.: Spezifikation von Transferpaketen und deren Übertragung an die Deutsche Nationalbibliothek mittels eines Hotfolders. – URN: <<http://nbn-resolving.de/urn:nbn:de:101-2012101804>>
- 11 <<http://www.openarchives.org/pmh/>>
- 12 <<http://www.dnb.de/gnd>>
- 13 Beyer, Christian; Trunk, Daniela: Automatische Verfahren für die Formalerschließung im Projekt PETRUS. In: Dialog mit Bibliotheken 23 (2011) 2, S. 5–10. – URN: <<http://nbn-resolving.de/urn:nbn:de:101-2012030831>>
- 14 Uhlmann, Sandro: Automatische Beschlagwortung von deutschsprachigen Netzpublikationen mit dem Vokabular der Gemeinsamen Normdatei (GND). In: Dialog mit Bibliotheken, 25 (2013) 2, S. 26–36. – URN: <<http://nbn-resolving.de/urn:nbn:de:101-20140305238>>
- 15 <<http://www.dnb.de/urnservice>>