

Nadine Walger

Sammlung von Netzpublikationen erreicht nächste Stufe

Die Deutsche Nationalbibliothek erweitert ihre Sammelaktivitäten um wissenschaftliche Open-Access-Publikationen

Das am 22. Juni 2006 erweiterte Gesetz über die Deutsche Nationalbibliothek (DNBG) verpflichtet die Deutsche Nationalbibliothek (DNB) zur Sammlung, Erschließung, Verzeichnung und Langzeitarchivierung von unkörperlichen Medienwerken (Netzpublikationen). Dieser Sammelauftrag ist angesichts der stetig ansteigenden Zahl an Publikationen, die in Deutschland in öffentlichen Netzen publiziert werden, kein leichtes Unterfangen und wäre ohne automatisierte Ablieferungsverfahren, Standardisierungsbemühungen sowie eine leistungsfähige technische Infrastruktur nicht möglich.

Pragmatisches Vorgehen

Sammeln in Stufen

Um die Massen an sammelpflichtigen Netzpublikationen zu bewältigen, hat sich die DNB von Anfang an für ein Sammeln in mehreren Stufen entschieden.¹ In einer ersten Stufe konzentrierten sich die Aktivitäten vornehmlich auf elektronische monografische Verlagspublikationen und Hochschulschriften, meist Dissertationen. Parallel dazu wurden automatische Prozesse und Workflows entwickelt, mit deren Hilfe online erschienene Hörbücher und Noten, E-Paper-Ausgaben, Zeitschnitte von Webseiten sowie nach und nach auch Artikel aus elektronischen Zeitschriften aus dem Verlags- und Wissenschaftssektor importiert werden konnten.

Aktuell sind über vier Millionen Netzpublikationen archiviert; weitere Millionen von digitalen Medienwerken, die in den gesetzlichen Sammelauftrag gehören, stehen noch aus. Die DNB wagt in ihrem Jahresbericht 2016 einen Blick in die Zukunft und

schätzt den Bestand an archivierten Netzpublikationen für das Jahr 2021 auf insgesamt etwa 26 Millionen.²

Um dieses strategische Ziel zu erreichen, verstärkt die DNB ihre Sammelaktivitäten im Verlags- und Wissenschaftsbereich.³ Der Fokus der zweiten Sammelstufe liegt unter anderem auf wissenschaftlichen Open-Access-Publikationen, die nach Maßgabe des DNBG an die DNB abzuliefern sind.

In der Vergangenheit hat es sich bewährt, bei größeren Sammlungsvorhaben mit verschiedenen Kooperationspartnern zusammenzuarbeiten, so zum Beispiel im E-Book-Bereich mit Vertriebsplattformen sowie mit der Marketing- und Verlagservice des Buchhandels GmbH (MVB), im Zeitschriften Sektor mit Software- und Verlagsdienstleistern. Um die komplexe Aufgabe der Sammlung von Open-Access-Publikationen zu bewältigen, hat die DNB im Jahr 2015 Kontakt zu den Betreibern der wissenschaftlichen Spezialsuchmaschine Bielefeld Academic Search Engine (BASE) der Universitätsbibliothek Bielefeld aufgenommen und konnte den bereits seit zehn Jahren international agierenden Metadaten-Aggregator für eine Zusammenarbeit gewinnen.

Neue Verfahren, Workflows und Prozesse

Angesichts der ablieferungspflichtigen Massen, die die DNB bereits heute bewältigen muss, wird es unumgänglich sein, die Speicherkapazitäten weiter auszubauen. Eine erste Anpassung erfolgte im Jahr 2016 „entsprechend des prognostizierten Wachstums im Bereich der elektronischen Publikationen“.⁴ Daneben gilt es, die technische Infrastruktur sowie automatische Geschäftsprozesse und Workflows für einen modularisierten und parallel verarbeitenden Import von Netzpublikationen anzupassen.

Erfolgsfaktor
Kooperation

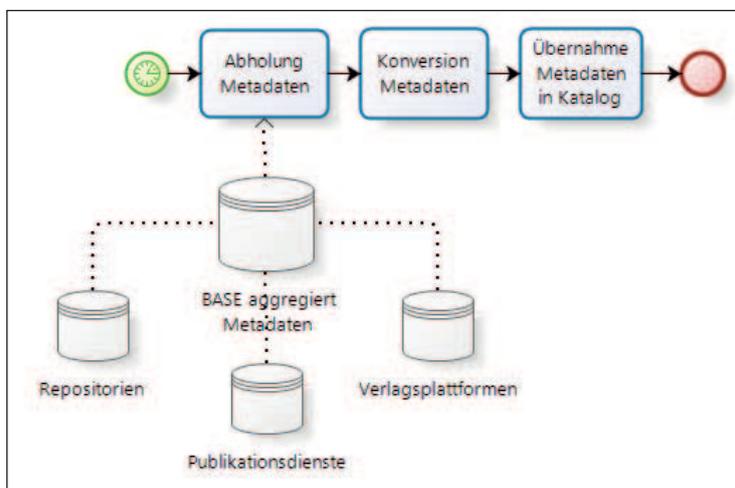
Projekt SamOA

Unbestritten ist in diesem Zusammenhang, dass für die zweite Sammelstufe erweiterte Anforderungen an die bisherigen automatischen Ablieferungsverfahren gestellt werden, für deren Umsetzung technische Lösungen entwickelt werden müssen. Das im März 2016 gestartete Projekt SamOA – Sammlung von Open-Access-Publikationen⁵ unterstützt dieses Vorhaben. Es ist im Fachbereich Informationsinfrastruktur der DNB angesiedelt und auf eine Laufzeit von zunächst drei Jahren angelegt.

Ziel des Projekts ist die Anpassung des Workflows für die automatische Pflichtablieferung von wissenschaftlichen Open-Access-Publikationen von in Deutschland betriebenen Repositorien und Publikationsdiensten mit Hilfe der aggregierten Datensammlung der Suchmaschine BASE. Die DNB sieht sich bei diesem Vorhaben explizit als nachweisende Institution und Langzeitarchiv, jedoch nicht als Betreiberin einer Publikationsplattform für Zweitveröffentlichungen.

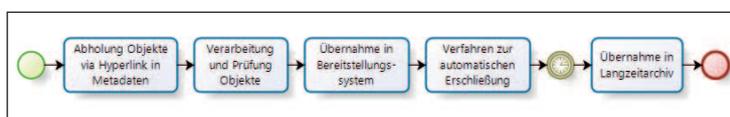
DNB als nachweisende Institution und Langzeitarchiv

Im Projekt SamOA kommt erstmals ein für die Übernahme von Netzpublikationen ergänztes zweistufiges Verfahren zur Anwendung: Beschreibende Metadaten zur Publikation werden nicht, wie bisher, zusammen mit dem digitalen Objekt abgeliefert, sondern getrennt voneinander. In einem ersten Schritt werden die von BASE aggregierten und normalisierten bibliografischen Metadaten über das OAI-Protocol for Metadata Harvesting (OAI-PMH) in den Katalog der DNB übernommen. Ein Mehrwert für Nutzerinnen und Nutzer besteht darin, dass viele Volltexte über einen externen Hyperlink in den BASE-Metadaten direkt und frei genutzt werden können.



Workflow Schritt 1

Der zweite Projektschritt beinhaltet die Übernahme der digitalen Objekte in die Speichersysteme der DNB. Der Objektimport erweist sich hier als Spezialfall, da die Objekte von BASE selbst nicht lokal gespeichert werden, sondern auf den jeweiligen Repositorien und Plattformen vorgehalten sind. Das derzeit technische Konzept der DNB sieht daher vor, den Objektimport über die Hyperlinks in den BASE-Metadaten zu realisieren. Im Idealfall sind diese Hyperlinks volltextverbunden und die DNB kann nach Maßgabe des DNBG auf das jeweilige unkörperliche Pflichtexemplar zugreifen und dessen Langzeitarchivierung sowie Langzeitverfügbarkeit sicherstellen.



Workflow Schritt 2

Herausforderungen und Perspektiven

Neben der Anpassung der technischen Workflows und Prozesse ergeben sich aus dem Projekt weitere Herausforderungen für die Pflichtablieferungsverfahren insgesamt, so zum Beispiel für eine automatisierte Verwaltung von Dubletten, Embargofristen oder verschiedenen Versionen von Publikationen, wie Preprint, Postprint, Verlags-PDF und Ähnlichem.

Eine Besonderheit der BASE-Metadaten ist, dass sie Lizenzinformationen zur Nachnutzung der Publikationen über das Urheberrecht hinaus enthalten. Die Angabe von Nachnutzungslizenzen und Rechtshinweisen ist für die DNB insbesondere zum Zweck der Sammlung und Langzeitarchivierung wichtig. Die DNB entwickelt derzeit Lösungen für die Übernahme und Bereitstellung dieser Angaben in den Katalog. Die hierbei gesammelten Erfahrungen und Projektergebnisse sind sicher auch für die Weiterentwicklung der Metadatenformate (beispielsweise MARC 21 oder NISO JATS) zur Pflichtablieferung nachnutzbar.

Ein positiver Nebeneffekt des Kooperationsprojekts mit BASE ist, dass die Nachweissituation und

Lizenzangaben und Rechtedokumentationen

damit die Sichtbarkeit von häufig stark verteilt vorliegenden Open-Access-Publikationen verbessert wird. Mit Unterzeichnung der Kooperationsver-

einbarung zwischen der DNB und BASE im April 2017 hat das Projekt SamOA einen wichtigen ersten Meilenstein erreicht.

Anmerkungen

- 1 Vgl. Diebel, Cornelia: Netzpublikationen - Sammlung, Archivierung und Bereitstellung in der Deutschen Nationalbibliothek. In: Dialog mit Bibliotheken 27 (2015) 1, S. 25. <<http://nbn-resolving.de/urn:nbn:de:101-2015100136>>
- 2 Vgl. Deutsche Nationalbibliothek (Hrsg.) Zwei Null Eins Sechs - Jahresbericht 2016. Frankfurt am Main, 2016, S. 5. <<http://nbn-resolving.de/urn:nbn:de:101-2017061405>>
- 3 Vgl. ebenda.
- 4 Vgl. ebenda, S. 35.
- 5 <<http://www.dnb.de/DE/Wir/Projekte/Laufend/samOA.html>>