

Kathrin Jockel

Eine Million Buchseiten digital

Bestandsschutz für rund 27 Mio. Medieneinheiten

In den Magazinen der Deutschen Nationalbibliothek (DNB) lagern nach 100 Jahren Sammeltätigkeit über 27 Mio. Medieneinheiten – jedes Stück ein potenzieller Schatz für künftige Generationen. Jedem Werk muss daher der bestmögliche Bestandsschutz zukommen: Dazu trägt ganz wesentlich bei, dass alle Exemplare in geschlossenen Magazinen unter optimalen klimatischen Bedingungen aufbewahrt werden und dass die DNB als Präsenzbibliothek die Werke nicht aus dem Haus gibt.

Gleichwohl reichen die besten Lagerbedingungen nicht aus, um auch die Medien dauerhaft zu schützen, die produktionsbedingt, wie etwa durch die Verwendung holzschliffhaltigen Papiers, nur begrenzte Zeit haltbar sind. Dann können Verfahren wie die der Massensäuerung oder der Papierspaltung greifen, um dem Zerfall Einhalt zu gebieten.

Schließlich aber gibt es – im Kontext der 27 Mio. sind es relativ wenige – Exemplare, bei denen die Maßnahmen der physischen Bestandserhaltung nicht mehr greifen, weil das Papier schon zu brüchig oder der Buchblock beschädigt ist. Solche Werke sind für den Lesesaal gesperrt, werden also nicht mehr an Benutzerinnen und Benutzer ausgegeben. Jede Benutzung würde sie nachhaltig schädigen oder gar für die Zukunft unbrauchbar machen. Die DNB hat nun begonnen, diese besonders gefährdeten Bücher zu digitalisieren. In einer Pro-

Dem Zerfall Einhalt gebieten durch Digitalisierung

jektlaufzeit von zunächst zwei Jahren sollen rund 8.000 Bücher mit einer Million Seiten in die digitale Form überführt werden.

Wie sehen die gedruckten Werke aus, die digitalisiert werden? Nahezu alle Werke enthalten lose Seiten, zum Teil ist das Papier an den Rändern angegriffen, sodass Ecken fehlen. Der Buchblock ist – da die Werke in der Regel nur wenig genutzt wurden – relativ stabil, bei vielen Bänden hat sich aber der Buchrücken gelöst oder ist gebrochen. In diesem Zustand können die Bücher nicht mehr an Benutzerinnen und Benutzer ausgegeben werden – als Vorlage für die Digitalisierung sind sie jedoch gleichwohl gut geeignet. Die Schädigungen wurden hervorgerufen, weil der Zahn der Zeit an den Materialien nagt. Sie können aber auch beispielsweise durch eine häufige Benutzung in der Bibliothek entstanden sein.

Die 8.000 Werke, die nun digitalisiert werden, sind kein in sich geschlossener Bestand, keine zusammenhängende Sammlung, kein Magazinsegment, sondern kommen quer aus den 27 Mio. Medienwerken des Gesamtbestandes, wobei der Schwerpunkt naturgemäß auf älteren Werken liegt, die aufgrund des sauren Papiers zerfallen. Die Gemeinsamkeit der 8.000 Bücher – und zunächst sind es nur Bücher, die bearbeitet werden, Zeitschriftenhefte werden erst im Rahmen einer möglichen Projektverlängerung behandelt werden – ist ihr mangelhafter konservatorischer Zustand. Thematisch spiegeln sie den Bestand der DNB: So wie im Magazin Kochbuch neben Dissertation, graue Literatur neben Comic steht, so stammen auch die 8.000 Werke aus allen Publikationen, die in Deutschland erscheinen und daher von der DNB gesammelt werden.

Nahezu der komplette Bestand der zu digitalisierenden Werke ist urheberrechtlich geschützt, auch das spiegelt die Situation des Bibliotheksbestandes wider. Auf eine individuelle Rechtlklärung wird in diesem Projekt verzichtet, da die Wahrscheinlichkeit, dabei ein gemeinfreies Werk zu identifizieren, nicht sehr groß ist und der Aufwand in keiner vertretbaren Relation zum Ergebnis stünde. Das

8.000 Werke werden digitalisiert



Problemstellung:
Urheberrecht

bedeutet, dass vom gedruckten Werk – weil urheberrechtlich geschützt – zwar eine digitale Kopie angefertigt, nicht aber eine Texterkennung mittels OCR durchgeführt werden kann. Es bedeutet außerdem, dass die Digitalisate nicht weltweit frei genutzt werden können, sondern dass Benutzerinnen und Benutzer sie ausschließlich in den Lese-sälen in Leipzig und Frankfurt am Main ansehen können. Sollte zu einem späteren Zeitpunkt eine Rechtlklärung möglich werden, so lässt sich der Status des Werks ohne großen Aufwand ändern, um das Digitalisat frei anzuzeigen. Auch eine OCR-Bearbeitung kann nachträglich erfolgen.

gedruckten Werkes eine Sekundärform erstellt und dabei alle relevanten Daten übernimmt, die für das gedruckte Werk spezifischen Angaben löscht und den neu angelegten Datensatz für die Sekundärform um digitalisatspezifische Angaben ergänzt. Das Ergebnis ist ein halbfertiger Datensatz, der später beim Import des Digitalisats angepasst und fertiggestellt wird. Damit existiert für jedes Buch bereits vor der Digitalisierung ein Datensatz mit einer Identifikationsnummer (IDN) für seine digitale Kopie. Die IDN wird im Rahmen der Digitalisierung – das ist der nächste Schritt im Workflow – beispielsweise für die Dateibezeichnungen verwendet.

Datenbank als
Projektgrundlage

Link zu diesem Datensatz	http://d-nb.info/364026782
Titel	Die Märchentante : Märchen u. Erzählungen / Klara Schott
Person(en)	Schott, Klara
Ausgabe	41.-43. Taus. Neue Ausg.
Verleger	Weseli : Düms
Erscheinungsjahr	[1913]
Umfang/Format	127 S. : [Ill.] ; 8
Parallele Ausgabe(n)	Online-Ausgabe: Schott, Klara: Die Märchentante
Anmerkungen	In Fraktur
Leipzig	Signatur: 1915 B 9003 Exemplar nicht verfügbar

Link zu diesem Datensatz	http://d-nb.info/1014155843
Titel	Die Märchentante [Elektronische Ressource] : Märchen u. Erzählungen / Klara Schott
Person(en)	Schott, Klara
Ausgabe	41.-43. Taus. Neue Ausg., [Online-Ausg.]
Verleger	Weseli : Düms
Erscheinungsjahr	[1913]
Umfang/Format	127 S. : [Ill.]
Sekundärausgabe	Online-Ausg. : Leipzig ; Frankfurt : Dt. Nationalbibliothek 2011. Online-Ressource
Parallele Ausgabe(n)	Druckausg. : Schott, Klara: Die Märchentante
Anmerkungen	In Fraktur
Online-Zugriff	Medium öffnen

Grundlage für den Geschäftsgang des Digitalisierungsprojektes ist eine Datenbank der beschädigten Medien, die seit vielen Jahren geführt wird. Auf ihr basieren die verschiedenen Bestandserhaltungsmaßnahmen der DNB. Zunächst werden die Objekte aus dem Magazin ausgehoben und buchbinderisch vorbereitet, d. h., sie werden aus Schutzhüllen ausgepackt, eventuell vorhandene Klammern werden entfernt. Neu identifizierte beschädigte Werke, entdeckt beispielsweise durch aktive Benutzung oder durch retrospektive Arbeiten im Magazin, können jederzeit in den Geschäftsgang integriert werden.

Die Digitalisate werden – noch vor ihrer Erstellung – als Parallelausgaben der gedruckten Werke im Katalog verzeichnet. Dabei hilft ein sogenannter »Digitalisierungsknopf« im Katalogisierungssystem der DNB, der automatisch aus dem Datensatz des

Für die Digitalisierung der 8.000 Bücher hat als Dienstleister im Rahmen einer EU-weiten Ausschreibung der Harald Fischer Verlag den Zuschlag erhalten. Die Objekte werden im Leipziger Haus der DNB digitalisiert, d. h., der Dienstleister arbeitet mit einer Scanstation vor Ort. Es entsteht eine TIFF-Datei je Buchseite, gescannt wird durchgehend in 300 ppi mit 24bit-Farbtiefe. Die Scanparameter wurden so gewählt, dass die spätere Darstellung des Werkes für den Nutzer so authentisch wie möglich ist und gleichzeitig möglichst gute Voraussetzungen für die Langzeitarchivierung geschaffen werden. Der Dienstleister erstellt zu dem Digitalisat technische Metadaten nach einem vorgegebenen XML-Schema (METS, MIX, MODS) und stellt die Digitalisate der DNB über ein Monitoringsystem

zur Qualitätssicherung zur Verfügung. DNB-Mitarbeiterinnen und -Mitarbeiter des Qualitätssicherungsteams kontrollieren stichprobenartig die Scans sowie die zugehörigen Metadaten. Die Qualitätssicherung findet laufend und parallel zum Produktionsbetrieb statt, um Fehler umgehend erkennen und beheben lassen zu können. Nach einem Zeitfenster von acht Tagen werden die Dateien automatisch weiterverarbeitet, um keinen Zeitverzug entstehen zu lassen. Die Digitalisate samt der Metadaten werden zur DNB übertragen,



von wo aus automatisch der Import in ihre Systeme startet: Die Dateien des Digitalisats werden in das entsprechende Repository eingespielt, und eine Importroutine vervollständigt den Katalogdatensatz des digitalisierten Werkes.

Recherchiert ein Nutzer im Katalog der DNB und stößt auf ein gedrucktes Werk, das aufgrund des konservatorischen Zustands nicht mehr ausgeliehen werden kann und das bereits digitalisiert ist, findet er eine Verknüpfung auf den Katalogeintrag der digitalen Form und kann von dort aus das Digitalisat öffnen. Dies ist jedoch nur möglich, sofern er sich in einem der Lesesäle der DNB befindet.

Ende August 2012 waren bereits 3.500 Werke mit 445.000 Seiten digitalisiert. Dafür werden etwa 4 TB Speicherplatz benötigt. 15 Werke kommen täglich neu hinzu.

Mit dem Digitalisierungsprojekt werden gesperrte Werke in digitaler Form wieder nutzbar. Nicht mehr Zugängliches wird wieder zugänglich. Die Inhalte der Publikationen werden dauerhaft gesichert. Nach der Digitalisierung werden die beschädigten Medien in Papierhüllen verpackt und langfristig der Benutzung entzogen, da dann ein Digitalisat zur Verfügung steht. Trotz Beschädigung, erhalten die Originale somit den bestmöglichen Schutz.

Fazit