

# Retrokonversion

THOMAS TABERY, WERNER HOLBACH

## Die Retrokonversion der Ostasienkataloge der Bayerischen Staatsbibliothek

Foto: privat



Thomas Tabery

Foto: privat



Werner Holbach

Die Retrokonversion der konventionellen Ostasienkataloge der Bayerischen Staatsbibliothek hat Anfang 2012 ihren erfolgreichen Abschluss gefunden. Damit sind sämtliche chinesische, japanische und koreanische Titel der Bayerischen Staatsbibliothek über deren Online-Katalog, über den Verbundkatalog des Bibliotheksverbunds Bayern (Gateway Bayern) und über WorldCat elektronisch nachgewiesen. Rund 166.000 Katalogkarten vor 1990 erschienener ostasiatischer Titel wurden bearbeitet und ca. 90.000 Datensätze von einem externen Dienstleister sowohl in Originalschrift als auch in Umschrift elektronisch erfasst. Der Artikel beschreibt das gewählte Verfahren, die wesentlichen Etappen und die besonderen Herausforderungen der Retrokonversion dieser Spezialkataloge, mit der die elektronische Erfassung der konventionellen Alphabetischen Kataloge der Bayerischen Staatsbibliothek insgesamt abgeschlossen werden konnte.

In early 2012 the Bavarian State Library completes the retro conversion of the conventional catalogs of its East Asian collection. All Chinese, Japanese and Korean holdings are now researchable through the online catalog, through the Bavarian Union Catalog (»Gateway Bayern«) and through WorldCat. Nearly 166,000 catalog cards of East Asian holdings published before 1990 were processed and nearly 90,000 data records provided by an outside service were created both in the original form and in transcription. This article describes which procedures were followed at various phases of the project and the particular challenges presented by the retro conversion of these special catalogs. The project represents the final stage of the conversion of the Bavarian State Library's conventional alphabetical catalogs.

### 1. KATALOGKONVERSION AN DER BAYERISCHEN STAATSBIBLIOTHEK

Die Einführung elektronischer Katalogsysteme in den 1980er Jahren hatte eine geteilte Kataloglandschaft zur Folge: Ältere Werke waren in konventionellen Katalogen wie hand- und maschinenschriftlichen Band-, Blatt- und Kartenkatalogen verzeichnet, Neuerwerbungen wurden dagegen ausschließlich elektronisch nachgewiesen. Im Zuge der digitalen Revolution der letzten zwei Jahrzehnte und des damit einhergehenden grundlegenden strukturellen Wandels im Bereich der wissenschaftlichen Information und Kommunikation werden digitale, online verfügbare Informationen in zunehmendem Maße als die – vermeintlich – relevanten Informationen betrachtet und nicht online recherchierbare Bestände nur noch sehr eingeschränkt wahrgenommen. Für Bibliotheken resultiert daraus die Aufgabe, einen durchgehenden elektronischen Nachweis ihrer Bestände zur Verfügung zu stellen und ihre historischen Kataloge in elektronische Nachweissysteme zu überführen.

Auf dem Gebiet der Retrokonversion sind im Bayerischen Bibliotheksverbund insbesondere seit den Empfehlungen des Landeskonzepts »Erhaltung, Archi-

vierung und Aussonderung von Druckschriften in Bayern«<sup>1</sup> von 1998, die einen vollständigen maschinenlesbaren Nachweis der Bibliotheksbestände fordern, erhebliche Fortschritte erzielt worden.<sup>2</sup> An der Bayerischen Staatsbibliothek (BSB) begann die Konversion der konventionellen Kataloge bereits fast zwei Jahrzehnte früher: Seit Anfang der 1980er Jahre wurden zunächst der sogenannte Bandkatalog (Nachweiszeitraum 1501 bis 1840, ca. 650.000 Titelaufnahmen),<sup>3</sup> später der Alphabetische Katalog der Jahre 1953 bis 1981 (sog. IFK-Katalog, ca. 1,4 Mio. Titelaufnahmen)<sup>4</sup> und schließlich der Alphabetische Katalog der Jahre 1841 bis 1952 (sog. Quartkatalog, ca. 1,6 Mio. Titelaufnahmen) dem jeweiligen Stand der Technik und den spezifischen Katalogmerkmalen entsprechend elektronisch erfasst.<sup>5</sup>

Damit war seit 2006 der überwiegende Teil der Bestände der BSB elektronisch nachgewiesen.<sup>6</sup> Übrig geblieben war als umfangreiche Sondersammlung einzig die Ostasiensammlung, die nach wie vor zu einem großen Teil nur in Form von Kartenkatalogen erschlossen war.

Mit der Retrokonversion der konventionellen Ostasienkataloge wurde 2009 begonnen. Die Erfassung der chinesischen Kataloge ist seit Februar 2011, die der japanischen Kataloge seit Anfang 2012 abgeschlossen.<sup>7</sup> Der vorliegende Artikel beschreibt das gewählte Verfahren, die wesentlichen Etappen und die besonderen Herausforderungen der Konversion dieser Spezialkataloge.

### 2. DIE OSTASIENBESTÄNDE UND -KATALOGE DER BAYERISCHEN STAATSBIBLIOTHEK

Die Ostasiensammlung der Bayerischen Staatsbibliothek umfasst rund 320.000 originalsprachige Bände in den Sprachen Chinesisch, Japanisch und Koreanisch. Die chinesische Sammlung, die zu einer der weltweit bedeutendsten außerhalb Chinas zählt und deren Anfänge bis in die erste Hälfte des 17. Jahrhunderts zurückreichen, umfasst gegenwärtig etwa 130.000 gedruckte chinesische Titel in ca. 230.000 Bänden und über 3.000 chinesische Handschriften.<sup>8</sup> Die japanische Sammlung zählt ca. 50.000 japanische Titel in ca. 90.000 Bänden, obwohl hier die Erwerbung erst nach dem Zweiten Weltkrieg einsetzte.<sup>9</sup> Die kleine koreanische Sammlung von knapp 3.000 Titeln entstand

frühere Retrokonversionsprojekte an der BSB

ebenfalls erst nach dem Zweiten Weltkrieg. Da sie vollständig elektronisch katalogisiert ist, wird im Weiteren nicht näher auf sie eingegangen.

Die Nachweissituation der Sinica und Japonica bis zum Beginn der Katalogkonversion im Jahre 2009 stellte sich wie folgt dar:

Obwohl die Bayerische Staatsbibliothek seit ihren Anfängen Literatur aus China sammelte, begann die systematische Erschließung der Ostasiatica in Katalogen erst nach 1945. Allerdings ließen die finanziellen und personellen Möglichkeiten der Bibliothek zunächst nur die Katalogisierung des laufenden Neuzugangs zu. Der reiche Altbestand konnte erst in den Jahren 1979–1982 durch die Unterstützung der Deutschen Forschungsgemeinschaft im Rahmen eines Sonderprojekts katalogisiert werden. Dessen Ergebnis – ein Kartenkatalog im Umfang von 60.000 Katalogkarten, der die Sinica bis Katalogisierungsjahr 1983 sowohl unter dem Verfasser als auch unter dem Sachtitel geordnet nachweist – erschien 1984 auch in gedruckter Form.<sup>10</sup> Nach der Drucklegung des L.sin.-Katalogs<sup>11</sup> wurde der L.sin.-Supplementkatalog angelegt, der die Sinica der Katalogisierungsjahre 1984–1989 verzeichnet und ca. 40.000 Katalogkarten umfasst.<sup>12</sup>

Alle vor 1990 katalogisierten japanischen Titel waren ebenfalls ausschließlich in Kartenkatalogen erschlossen worden. Ursprüngliche Planungen, auch den japanischen Kartenkatalog gedruckt zu publizieren, waren nicht realisiert worden.

Seit 1990 werden die CJK-Titel (d. h. Titel in Chinesisch, Japanisch und Koreanisch) der BSB elektronisch katalogisiert: Von 1990 bis Mitte 2007 wurden sie in der RLIN-Datenbank der Research Libraries Group (RLG) erfasst (insgesamt etwa 43.000 Titelaufnahmen), nach der Integration von RLIN in WorldCat<sup>13</sup> ab Ende 2007 bis Ende 2008 direkt in WorldCat (etwa 5.000 Titelaufnahmen). Folgende zwei Argumente hatten den Ausschlag für den RLG Union Catalog gegeben: Zum einen konnten CJK-Titel in RLIN zu dieser Zeit bereits originalschriftlich erfasst werden, zum anderen war in RLIN schon damals eine beträchtliche Anzahl an originalschriftlichen Titelaufnahmen vorhanden, die für die Katalogisierung genutzt werden konnten. Eine Erfassung in Originalschrift ist deshalb so bedeutsam, weil aufgrund der Besonderheiten der ostasiatischen Sprachen und Schriften Titel im Katalog nur dann eindeutig zu identifizieren sind, wenn sie nicht nur in Umschrift, sondern auch in Originalschrift verzeichnet werden. Für das Chinesische trifft dies in besonderer Weise zu.

Seit Anfang 2009 werden CJK-Titel direkt im B3Kat, der Verbunddatenbank des Bibliotheksverbands Bay-

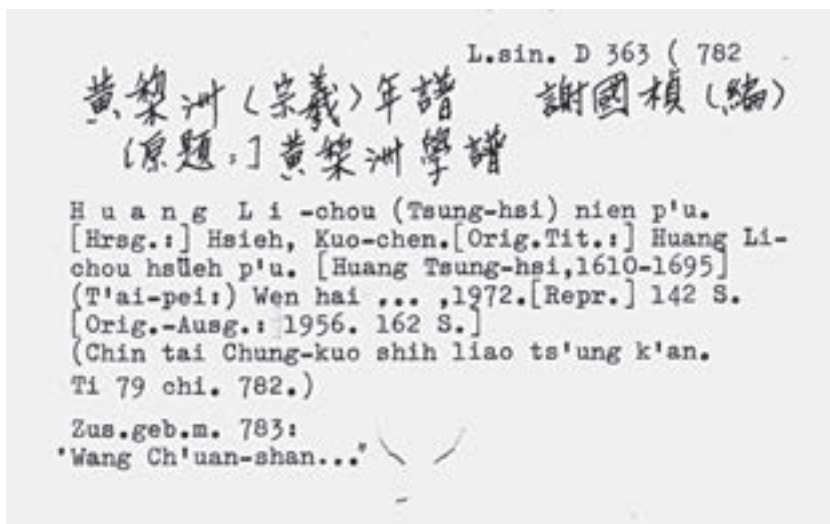


Abb. 1: Katalogkarte aus dem Sinica-Kartenkatalog

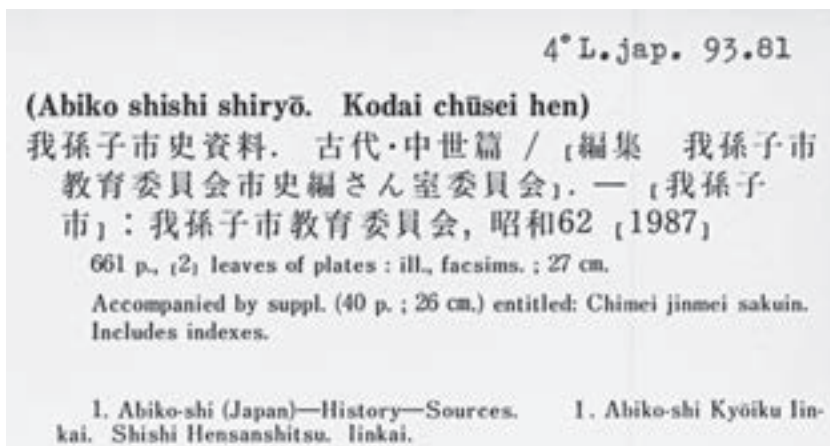


Abb. 2: Katalogkarte aus dem Japonica-Kartenkatalog

ern (BVB) und des Kooperativen Bibliotheksverbunds Berlin-Brandenburg (KOBV), katalogisiert. Die in RLIN und in WorldCat katalogisierten Titel wurden bereits in den Verbundkatalog und in den lokalen Katalog (OPACplus) der Bayerischen Staatsbibliothek migriert. Dies war möglich geworden, nachdem zwei wichtige technische Voraussetzungen für eine Datenmigration erfüllt waren: Zum einen verfügte das Verbundsystem Aleph 500 über eine MARC-Schnittstelle, die einen Import von Titelsätzen im Datenformat MARC21, dem in RLIN und WorldCat verwendeten Katalogisierungsformat, ermöglichte. Zum anderen war die Verbunddatenbank unicodefähig, so dass eine korrekte Übernahme und Speicherung originalschriftlicher Zeichen gewährleistet war.<sup>14</sup>

Sämtliche vor 1990 katalogisierte chinesische und japanische Titel waren nur in Form von Kartenkatalogen nachgewiesen und damit Gegenstand der Retrokonversion.

bisherige Nachweissituation der asiatischen Bestände

### 3. ANALYSE DER KATALOGE UND WAHL DES KONVERSIONSVERFAHRENS

Die zu konvertierenden Ostasienkataloge umfassten insgesamt rund 166.000 Katalogkarten, das exakte Mengengerüst stellte sich dabei wie folgt dar:

	Gesamtzahl der Katalogkarten	davon Nebeneintragungen	davon Verweisungen	Zahl der zu erfassenden Katalogkarten
<b>Sinica-Grundkatalog</b> (Katalogisierungsjahre bis 1983)	60.000	28.000	1.500	32.000
<b>Sinica-Supplementkatalog</b> (Katalogisierungsjahre 1984–1989)	40.000	22.000	500	18.000
<b>Japonica-Grundkatalog</b> (Katalogisierungsjahre bis 1989)	65.000	33.000	1.000	32.000
<b>Altjaponica-Katalog</b> (Katalog alter und seltener japanischer Werke bis Katalogisierungsjahr 1989)	1.000	250	–	750
<b>Gesamt</b>	166.000	83.250	3.000	82.750

Die Kataloge sind alphabetisch nach Verfassern (oder sonstigen beteiligten Personen) und Sachtiteln geordnet. Die Katalogkarten im IFK-Format (7,5 x 12,5 cm) enthalten maschinenschriftliche und handschriftliche Informationen sowohl in chinesischer und japanischer Originalschrift als auch in lateinischer (Um-)Schrift. Für chinesische Titelaufnahmen wurde Wade-Giles, für japanische Titelaufnahmen Hepburn und modifiziertes Hepburn als Umschrift verwendet.

Die Haupteintragung eines Werkes findet sich entweder unter dem Verfasser oder dem Sachtitel. Verfasserwerke haben zusätzlich immer eine Nebeneintragung unter dem Sachtitel und mitunter Nebeneintragungen unter sonstigen beteiligten Personen; Sachtitelwerke haben gelegentlich zusätzlich Nebeneintragungen unter beteiligten Personen. Außerdem enthält der Katalog Karten mit Verweisungen von alternativen Namensformen auf die jeweilige Ansetzungsform.

Auf der Grundlage der Kataloganalyse und des ermittelten Mengengerüsts galt es, eine Reihe von grundsätzlichen Entscheidungen über die Art des Konversionsverfahrens zu treffen:

#### Erfassung durch einen externen Dienstleister

Die sprachlichen Besonderheiten der Kataloge ließen den Einsatz einer automatischen Texterkennung nicht zu. Die Retrokonversion der Kataloge mit insgesamt rund 166.000 zu bearbeitenden Katalogkarten war mit vorhandenem Personal in einem überschaubaren Zeitraum nicht zu leisten, zumal sie den Einsatz von Muttersprachlern erforderte, um die handschriftlichen Angaben in chinesischer und japanischer Schrift auf den Katalogkarten lesen und bibliographisch korrekt interpretieren zu können. Diese Anforderung einerseits und Kostengründe andererseits gaben den Ausschlag dafür, das Erfassen der Daten von einem externen Dienstleister durchführen zu lassen.

#### Erfassung im Online-Offline-Verfahren

Auf die Entscheidung für einen Dienstleister folgte die Frage, ob eine Online-Erfassung der Titelaufnahmen über eine Netzverbindung direkt in der Verbunddatenbank erfolgen oder ob besser eine Offline-Erfassung zur Anwendung kommen sollte. Da eine Online-Erfassung im B3Kat einen hohen Schulungsaufwand für den Dienstleister erfordert hätte und auch in technischer Hinsicht (z. B. garantierte Verfügbarkeitszeiten) problematisch gewesen wäre, wurde von diesem Verfahren Abstand genommen.

Eine reine Offline-Erfassung in einer separaten Datenbank hätte zwar die Vorteile einer einheitlichen Feldbelegung und Datenqualität sowie einer unkomplizierten Qualitätsprüfung und Fehlerkorrektur mit sich gebracht. Allerdings wären die Kosten bei diesem Verfahren deutlich höher gewesen als im Falle einer Erfassung in einem Katalogsystem, das eine Nutzung vorhandener Titelaufnahmen bzw. eine Fremddatenübernahme zulässt.

Als Verfahren wurde deshalb ein dritter Weg, nämlich ein gemischtes Online-Offline-Verfahren, gewählt. Die Katalogisierung erfolgte im ersten Schritt online in der WorldCat-Datenbank. Stichproben hatten ergeben, dass etwa zwei Drittel der zu erfassenden Titel in WorldCat enthalten sind. An die vorhandenen Titelaufnahmen, sogenannte Master Records, wurden im Rahmen der Konversion sogenannte Institution Records, in denen die Lokaldaten (insbesondere die lokalen Signaturen) ergänzt wurden, angefügt (»copy cataloging«). Ein positiver Nebeneffekt davon war, dass die eigenen Titelaufnahmen mit in den WorldCat-Aufnahmen vorhandenen Sacherschließungsdaten wie Subject Headings oder DDC-Notationen angereichert werden konnten. Durch die eindeutigen OCLC-ID-Nummern sind die Datensätze bei zukünftigen Änderungen in der B3Kat-Datenbank problemlos in die regulären Update-Lieferungen des BVB an WorldCat eingebunden.

gemischtes  
Online-Offline-Verfahren

Um Kosten für eine Datenausspeicherung durch OCLC aus WorldCat zu vermeiden, wurden die in WorldCat erfassten Titelaufnahmen nicht nur dort, sondern parallel dazu auch durch den Auftragnehmer lokal gespeichert (Erstellen einer »copy on file« mit dem Connexion Client, dem Katalogisierungsclient von WorldCat). Nach Beendigung der Erfassung lieferte er die Daten zur Offline-Einspeicherung an die Bayerische Staatsbibliothek.

### **Erfassung in MARC21 und Daten-Mapping nach MAB2 bei der Einspeicherung**

Die Entscheidung für eine primäre Erfassung in WorldCat bedeutete, dass die Titeldaten in MARC21 erfasst und geliefert würden. Dies stellte jedoch kein Problem dar, da eine für die Einspeicherung in die Aleph-Verbunddatenbank notwendige Mapping-Routine von MARC21 nach MAB2 bereits seit 2007 existierte und für die Dateneinspeicherung nach Abschluss der Datenerfassung nur leicht modifiziert werden musste: Im Jahr 2007 waren im Zuge der Migration von MARC21-Titeldaten aus RLIN zunächst die seit 1990 in dieser Datenbank erfassten CJK-Titeldaten ausgespeichert worden.<sup>15</sup> Auf der Grundlage einer von der Verbundzentrale vorgenommenen Datenanalyse waren schrittweise die Vorgaben für das Mapping und die Anpassung der Daten entwickelt worden. Diese Entwicklung einer speziellen Mapping-Routine von MARC21 nach MAB2, die eine fehlerfreie Übertragung auch der originalschriftlichen Felder beinhalten musste, war sehr aufwändig. Nach mehreren Testeinspielungen in eine Aleph-Testdatenbank wurden im Sommer 2008 die Datensätze über eine Batch-Schnittstelle erfolgreich in die Verbunddatenbank eingespeichert. Es handelte sich dabei um die erste größere MARC-Einspeicherung in die Verbunddatenbank des Bibliotheksverbunds Bayern überhaupt.

### **Erfassungsstruktur mehrbändiger Werke:**

#### **Erfassung von Einzelaufnahmen**

Bei mehrbändigen Werken, die ohne Autopsie rein aufgrund der Katalogkarte als solche zu identifizieren waren, erhielt jeder Einzelband eine separate und vollständige Titelaufnahme inkl. aller Angaben, die sonst nur in der Gesamtaufnahme zu finden sind. Eine eigene Gesamtaufnahme wurde nicht angelegt und somit auch keine Verknüpfungen zwischen den Einzelaufnahmen vorgenommen. Dies ist in WorldCat technisch gar nicht möglich. Die Folge, nach der Einspeicherung der Datensätze auch in den eigenen Katalogsystemen Einzelaufnahmen mehrbändiger Werke unverknüpft nachzuweisen, wurde dabei in Kauf genommen.

The screenshot shows a list of bibliographic records in a cataloging client. The records are organized into groups, with some records having a '100' field and others having a '240' field. The records include titles in Chinese and English, along with publication information and mapping codes. For example, one record is 'Xie, Guozhen, Ad 1901-1982.' and another is 'Huang Lihou (Zongqi) nian pu'. The interface includes a search bar at the top and a list of records below.

Abb. 3: Chinesische Titelaufnahme im Katalogisierungsclient des WorldCat (Ausschnitt)

Die Zahl der elektronisch zu erfassenden Katalogkarten betrug knapp 83.000. Dazu kamen mehrere Tausend zu erfassende Aufnahmen für Bände mehrbändiger Werke, so dass letztendlich insgesamt rund 90.000 Datensätze zu erstellen waren.

## **4. VORBEREITUNG UND DURCHFÜHRUNG DER KONVERSION**

Zwar entsprechen die Japonica-Kataloge in ihren allgemeinen Ordnungsprinzipien weitgehend den Sinica-Katalogen, doch weisen sie eine ganze Reihe von Besonderheiten auf, die eine spezifische Erfassungsanleitung erforderlich machten. Aus diesem Grund und aus Gründen der Qualitätskontrolle wurden die chinesischen und japanischen Kataloge nicht gleichzeitig, sondern in aufeinander folgenden Projekten bearbeitet, beginnend mit den Sinica-Katalogen. Die vorbereitenden Arbeiten und die Phasen der Durchführung ähnelten sich bei beiden Projekten jedoch so stark, dass sie, um Wiederholungen zu vermeiden, nicht separat dargestellt werden.

### **Anforderungsprofil und Marktsondierung**

Eine zentrale Rolle bei jeder aus der Bibliothek ausgelagerten Retrokonversion spielen die Vorgaben für die

chinesische und japanische Kataloge getrennt bearbeitet

zu erbringenden Leistungen des Auftragnehmers und eine detaillierte Beschreibung der Arbeitsschritte. Die hierzu zu erstellenden Leistungsbeschreibungen und Erfassungsanleitungen sind bereits für die Angebots-einholung und Preiskalkulation der Anbieter wesentlich, darüber hinaus aber auch für die Projektdurchführung inklusive Qualitätskontrolle unverzichtbar.

Aus den Ergebnissen der Kataloganalyse und der Entscheidung für ein bestimmtes Konversionsverfahren resultierten folgende Leistungsanforderungen an den Auftragnehmer:

- Der Auftragnehmer sollte bereits über Erfahrungen mit Konversionsprojekten in wissenschaftlichen Bibliotheken verfügen.
- Das eingesetzte Personal sollte mit dem OCLC Connexion Client, dem Katalogisierungsclient von WorldCat, und der Recherche in WorldCat vertraut sein und vorhandene Titelaufnahmen dort zuverlässig identifizieren können.
- Erfasst werden sollten die bibliographischen Angaben sowohl in chinesischer und japanischer Originalschrift als auch in lateinischer Umschrift.
- Die Konversion erforderte den Einsatz chinesischer und japanischer Muttersprachler, die die handschriftlichen Angaben in chinesischer und japanischer Schrift lesen und korrekt erfassen können.
- Das eingesetzte Personal sollte mit dem Datenformat MARC21 vertraut und in der Lage sein, bibliographische Daten sicher zu interpretieren und Titelaufnahmen in strukturierter Form zu erfassen. So müssen z. B. Haupteintragungen, Nebeneintragungen und Verweisungen sicher identifiziert werden. Besonders die Interpretation und elektronische Erfassung von Katalogkarten zu chinesischen und japanischen Handschriften ist bibliographisch anspruchsvoll.

Anfragen bei potentiellen Dienstleistern im europäischen Raum ergaben, dass vor allem aufgrund der sprachlichen Anforderungen kein Interesse an der Durchführung des geplanten Vorhabens bestand. Die Marktsondierung konzentrierte sich deshalb auf mögliche Dienstleister in Fernost. Sowohl das Vergabeverfahren zur Konversion der Sinica-Kataloge als auch dasjenige zur Konversion der Japonica-Kataloge hatte zum Ergebnis, dass der Auftrag der Firma China International Book Trading Corporation (CIBTC) mit Sitz in Peking erteilt wurde.

#### Scannen der Katalogkarten

Der mit der Retrokonversion beauftragten Firma wurden die Katalogkarten nicht im Original, sondern in digitalisierter Form zur Verfügung gestellt.<sup>16</sup> Dies entspricht der gängigen Praxis bei Konversionspro-

jekten. Einerseits ist die Benutzbarkeit der Bestände auch während der Konversion gewährleistet, zum anderen war die Bereitstellung digitaler Images auf einem Server zum Download weit weniger aufwändig als der Transport von 176 Katalogkästen.<sup>17</sup> Darüber hinaus ließen sich für die Firma mithilfe der Images die Arbeitsprozesse effizienter strukturieren, als dies mit den Originalkarten möglich gewesen wäre. Aus Gründen der besseren Lesbarkeit wurden auch für die Sinica bis 1983 digitale Images der originalen Katalogkarten und nicht nur deren verkleinerte Abbildung im gedruckten Bandkatalog bereitgestellt.

#### Erstellen von detaillierten Erfassungsanleitungen

Für die Sinica- und die Japonica-Kataloge wurden jeweils detaillierte Erfassungsanleitungen erarbeitet, auf deren Grundlage die Erfassung zu erfolgen hatte. Hier wurden u. a. Vorgaben zur Belegung der MARC-Felder und -Unterfelder, zum Zeichensatz – UTF-8 (Unicode) – und zur Behandlung der originalschriftlichen Felder gemacht.

#### Projektphasen

Entsprechend der Leistungsbeschreibung führte der Auftragnehmer folgende Arbeitsschritte aus: Vor der elektronischen Erfassung der Titelaufnahmen mussten in einem ersten Arbeitsschritt zunächst Haupteintragungen, Nebeneintragungen und Verweisungen als solche identifiziert werden. Nebeneintragungen waren nicht zu bearbeiten, Katalogkarten mit Verweisungen waren zunächst zurückzustellen und die Verweisungsformen erst am Ende der Erfassung in den entsprechenden Titelaufnahmen zu ergänzen.

Beim zweiten und wichtigsten Arbeitsschritt, dem Erfassen der Haupteintragungen, lautete die Vorgabe, grundsätzlich alle Informationen, die auf den Katalogkarten enthalten sind, vorlagegemäß zu erfassen. Dies bedeutete, dass bei der Übernahme von WorldCat-Aufnahmen in diesen Aufnahmen Ergänzungen vorgenommen wurden, sofern die Katalogkarten zusätzliche bibliographische Angaben enthielten.

Erfasst wurden bibliographische Angaben sowohl in chinesischer und japanischer Originalschrift als auch in lateinischer Umschrift. Chinesische Schriftzeichen waren wie auf den Katalogkarten angegeben entweder als Kurz- oder als Langzeichen im Code UTF-8 (Unicode) zu erfassen. Die auf den chinesischen Katalogkarten verwendete Umschrift Wade-Giles musste im Zuge der elektronischen Erfassung nach Pinyin umgesetzt werden.<sup>18</sup> Silben waren getrennt zu schreiben. Einzige Ausnahmen (entsprechend der Konvention in WorldCat): Zusammenschreiben waren Ortsnamen (z. B. »Zhongguo«, nicht »Zhong guo«) und

bei Personennamen die Vornamen (z. B. »Gao Minde«, nicht »Gao Min de«).<sup>19</sup>

Japanische Schriftzeichen waren, wie auf den Katalogkarten angegeben, entweder in Hiragana, Katakana oder Kanji im Code UTF-8 (Unicode) zu erfassen. Die auf den japanischen Katalogkarten verwendeten Transkriptionssysteme Hepburn und modifiziertes Hepburn mussten einheitlich in das modifizierte Hepburn-System konvertiert werden, wie es von der Library of Congress verwendet und vom japanischen Verbundkatalog NACSIS empfohlen wird, mit zwei Ausnahmen: Anstelle von Bindestrichen waren Leerzeichen zu verwenden (z. B. Yüraku chō, nicht Yüraku-chō), und japanische Wörter waren getrennt zu schreiben (z. B. Tōkyō Daigaku Kyōyō Gakubu).<sup>20</sup>

Da MAB-Felder im Aleph-System des B3Kat nicht ausschließlich mit Originalschrift belegt werden dürfen, ist für die MAB-Standard-Felder eine Feldbelegung in Umschrift obligatorisch. Deshalb mussten alle Angaben, die auf den Katalogkarten nur in Originalschrift verzeichnet waren, transkribiert und in den entsprechenden Feldern ergänzt werden.

Für ca. 70 % der zu erfassenden Titel existierte bereits eine Titelaufnahme in WorldCat. In diesen Fällen mussten nur sog. Institution Records angelegt und dort Lokaldaten ergänzt werden. Für rund 30 % der zu erfassenden Titel war eine Neuaufnahme erforderlich. Dafür musste zunächst ein sog. Master Record angelegt werden, aus dem ein Institution Record erzeugt werden konnte, in dem die Lokaldaten ergänzt wurden. Am Ende mussten also zu allen Titeln sowohl ein Master Record als auch ein Institution Record in WorldCat vorhanden sein.<sup>21</sup>

In einem dritten Schritt wurden die auf den Verweiskärtchen angegebenen alternativen Namensformen in den betreffenden Titelaufnahmen nachgetragen. In einem vierten Schritt wurden die in WorldCat angelegten Institution Records nach einer letzten Qualitätsprüfung als »copy on file« in MARC21 durch den Auftragnehmer lokal gespeichert. Diese Datensätze wurden im fünften und letzten Schritt der Bayerischen Staatsbibliothek zur finalen Qualitätskontrolle geliefert.

### \_\_\_ Qualitätssicherung

Wie bei allen Retrokonversionsprojekten sind erprobte Verfahren zur Qualitätssicherung – sowohl im Hinblick auf Vollständigkeit als auch auf Fehlerfreiheit – unabdingbar. Der Auftragnehmer war dazu aufgefordert, geeignete Verfahren zur Qualitätssicherung darzulegen und anzuwenden. Zur Gewährleistung eines hohen Qualitätsstandards wurden Prozentzahlen festgelegt, deren Nichterreichen eine kostenlose Nachbesserung seitens des Auftragnehmers zur Folge hatte.<sup>22</sup>

Während der gesamten Dauer der Erfassung der Titeldaten wurden diese in Stichproben anhand der OCLC-Nummern vom Auftraggeber auf ihre Qualität hin überprüft. Besonders auf die richtige Erfassung der Signaturen wurde großer Wert gelegt. Als von einiger Bedeutung hat sich die enge und zeitnahe Kommunikation zwischen Auftraggeber und Auftragnehmer herausgestellt. Denn trotz der ausführlichen Erfassungsanleitung hatte die Heterogenität der Kartenkataloge zur Folge, dass sich im Projektverlauf immer wieder Problemfälle und Detailfragen ergaben, für die auf diese Weise eine schnelle Lösung gefunden werden konnte.

## 5. INTEGRATION DER DATEN IN VERBUND- UND LOKALKATALOG

Im November 2010 schloss CIBTC die im Januar desselben Jahres begonnene Retrokonversion der Sinica-Kataloge erfolgreich ab und lieferte insgesamt knapp 54.000 MARC-Datensätze. Nach einer abschließenden Qualitätskontrolle und Datenbereinigung wurden im Januar 2011 diese Datensätze von der Verbundzentrale des BVB in die Aleph-Verbunddatenbank importiert, wobei die Daten von MARC21 nach MAB2 umgesetzt und mit BSB-Besitznachweisen versehen wurden. Auf die anschließende Einspielung der Datensätze ins Lokalsystem der Bayerischen Staatsbibliothek<sup>23</sup> folgten einige manuelle und maschinelle Korrekturen der Katalogisate. Korrigiert bzw. gelöscht wurden z. B. mehr als 250 fehlerhaft erfasste Signaturen, falsche Unterfeldbelegungen (z. B. MARC 852 \$a statt \$j) oder doppelt gelieferte Institution Records. Im Februar 2011 wurden schließlich noch die sogenannten Buchdaten (für die Ausleihe relevante Angaben) für das Ausleihmodul des Lokalsystems generiert. Damit hatte der Import der Sinica-Konversionsdaten seinen Abschluss gefunden.

Mit der Retrokonversion der Japonica-Kataloge war im April 2011 begonnen worden; im Dezember 2011 wurden rund 36.000 Datensätze geliefert, deren Einspeicherung in die Verbunddatenbank und in das Lokalsystem Anfang 2012 nach demselben Schema wie die Einspeicherung der chinesischen Titelaufnahmen erfolgt.

## 6. FAZIT

Mit Abschluss der beschriebenen Retrokonversionsprojekte ist das Ziel eines durchgehend maschinenlesbaren Nachweises der konventionellen Ostasienkataloge der Bayerischen Staatsbibliothek erreicht: Seit Anfang 2012 sind sämtliche CJK-Titel im Online-Katalog der Bayerischen Staatsbibliothek, im Verbundkatalog des Bibliotheksverbunds Bayern (Gateway Bayern), in WorldCat sowie in einschlägigen Fachportalen, allen voran der Virtuellen Fachbibliothek Ost- und Südostasien (CrossAsia) elektronisch nachgewiesen.

enge Kommunikation mit Auftragnehmer

30 % der Titel bisher noch nicht in WorldCat

sämtliche Ostasiatica der BSB elektronisch nachgewiesen

**stark gestiegene  
Nutzung der Bestände**

Das Verfahren, die Arbeit im Wesentlichen aus der Bibliothek auszulagern und einem Dienstleister zu übertragen, hat sich – wie schon bei früheren Konversionsprojekten – bewährt. Es erlaubte eine relativ genaue zeitliche und finanzielle Kalkulation. Eine Erfassung mit eigenem Personal in der Bayerischen Staatsbibliothek wäre außerdem bei weitem zeitaufwändiger und teurer gewesen. Die Nutzung der in WorldCat in großem Umfang vorhandenen Fremddaten wirkte sich nicht nur günstig auf die Kosten aus, sondern trug auch deutlich zur Qualität der Titeldaten bei. Entscheidend für die Güte waren aber vor allem die gründliche und detaillierte Projektvorbereitung sowie die laufende Qualitätskontrolle seitens der Bibliothek und die große Zuverlässigkeit des Dienstleisters.

Bei der Einspeicherung der Datensätze in die Verbunddatenbank wurden kaum Dubletten erkannt, der Titeldatenpool des B3Kat verzeichnete somit einen quantitativen wie qualitativen Zuwachs. Da sich eine ganze Reihe gerade der älteren Titel außerhalb Chinas und Japans ausschließlich im Bestand der Bayerischen Staatsbibliothek befindet, dürfte der elektro-

nische Nachweis dieser Titel die Sichtbarkeit der BSB (und ihrer Ostasien-Bestände) in den verschiedenen nationalen und internationalen Katalogen deutlich steigern. Dass sich durch den elektronischen Nachweis die Nutzung dieses Bestandssegments bereits in den ersten Monaten signifikant erhöht hat, belegt die stark gestiegene Zahl der Ausleihen chinesischer Titel seit Anfang 2011.

Mit der Integration in den B3Kat sind die Titelaufnahmen in einen aktiven und lebendigen Erschließungskontext eingebunden. Der zeitgemäße Nachweis der Bestände eröffnet neue Perspektiven der Nutzung und stellt die Basis für eine weitere Optimierung der Katalogisate dar. Dies kann intellektuell geschehen, indem beispielsweise Erschließungsexperten der BSB oder anderer Bibliotheken in der Verbunddatenbank gezielt tätig werden. Ein großes Potential liegt auch in denkbaren maschinellen Verfahren zur Verbesserung der Erschließungsqualität. Durch den vermehrten Datenaustausch zwischen den deutschsprachigen Bibliotheksverbänden können beispielsweise Sacherschließungsdaten und Kataloganreicherungen nachgenutzt werden. Wird eine Titelaufnahme von einem anderen Bibliotheksverbund nachgenutzt, werden die ggf. von diesem vergebenen Schlagwörter und Notationen sowie gescannten Inhaltsverzeichnisse über den zwischen den Bibliotheksverbänden eingerichteten kooperativen Datenaustausch automatisch in der B3Kat-Aufnahme ergänzt. Dieses Szenario ist, trotz der Tatsache, dass es sich um ältere Titel handelt, nicht unrealistisch, da vielerorts Retrokonversions- und Re-katalogisierungsmaßnahmen erfolgen.

Bezüglich der Formalerschließung besteht noch ein Manko in der fehlenden Normdatenverknüpfung mit der Personennamendatei PND und der Gemeinsamen Körperschaftsdatei GKD bzw. der zukünftigen Gemeinsamen Normdatei GND. Mit Einführung der GND im Frühjahr 2012 wird es möglich werden, originalschriftliche Namensformen zu erfassen. Es bleibt zu untersuchen, inwiefern nachträglich ein maschineller Abgleich der GND mit den Konversionsaufnahmen im B3Kat mit dem Ziel einer automatischen Verknüpfung mit den Normdatensätzen erfolgversprechend wäre. Da sowohl die deutschen Normdaten als auch die Titelaufnahmen des BVB sowie die internationalen Normdaten und die Titelaufnahmen aus WorldCat (in dem auch viele chinesische und japanische Bibliotheken ihre Bestände nachweisen) für den Virtual Authority File VIAF<sup>24</sup> herangezogen werden, besteht hier durch eine Nachnutzung der VIAF-Algorithmen und -ergebnisse eine vielversprechende Perspektive für eine automatische Verknüpfung.

Der Wert der Retrokonversion besteht nicht allein

Abb. 4: Chinesische Titelaufnahme im Aleph-System des B3Kat (Ausschnitt)

Sprachencode	037	b	a	chi
Materialcodes	050	—	—	a a
Codes Begr.W.	051	—	—	s          r
ID Bearb. Inst.	070	—	—	DLC
Anwendercodes	070	f	a	chin
Anwendercodes	078	g	a	BSBCJKSINB
Stufe	080	—	—	03
1. Person	100	—	—	Xie, Guozhen <1901-1982>
1. Person OS	A00	—	—	謝國楨 <1901-1982>
			f	Hani
EST	304	—	—	Huang Lizhou xue pu
EST OS	C04	—	—	黃黎澍學譜
			f	Hani
HST VF	331	—	—	Huang Lizhou (Zongxi) nian pu
HST VF OS	C31	—	—	黃黎澍 (宗羲) 年譜
			f	Hani
1. PST VF	341	b	a	Huang Lizhou (Zongxi) xue pu
1. PST VF	341	b	a	Wang Chuanshan (Fuzhi) xue pu
Verfasserangabe	359	—	—	Xie Guozhen bian
Verfasserang.OS	C59	—	—	謝國楨編
			f	Hani
WST	370	a	a	Wang Chuanshan (Fuzhi) nian pu
WST	370	a	a	Huang Lizhou xue pu
WST	370	a	a	Huang Lizhou (Zongxi) nian pu
WST OS	C70	a	a	王船山 (夫之) 年譜
			f	Hani
WST OS	C70	a	a	黃黎澍學譜
			f	Hani
WST OS	C70	a	a	黃黎澍 (宗羲) 學譜
			f	Hani
WST OS	C70	a	a	王船山 (夫之) 學譜
			f	Hani
WST OS	C70	a	a	黃黎澍 (宗羲) 年譜
			f	Hani
Ort 1. Verlag	410	—	—	[Taipei Xian Yonghe Zhen]
Ort 1. Verl. OS	D10	—	—	[臺北縣永和鎮]
			f	Hani
Name 1. Verlag	412	—	—	Wen hai chu ban she
Name 1.Verl. OS	D12	—	—	文海出版社
			f	Hani
Ersch.jahr	425	—	—	[Minguo 61 i. e. 1972]
Ersch.jahr	425	a	a	1972
Umfang	433	—	—	142 ; 4, 2, 219 p.
Format	435	—	—	20 cm
1.GT VF	451	—	—	Jin dai Zhongguo shi liao cong kan. Di 79 ji ; 782-783
1.GT VF OS	D51	—	—	近代中國史料叢刊. 第79輯 ; 782-783
			f	Hani
Bandangabe	455	—	—	782-783
Bandangabe OS	D55	—	—	782-783
			f	Hani
Bandang. SF	456	—	—	782-783

in einem verbesserten Nachweis der Bestände in modernen suchmaschinenbasierten Online-Katalogen. Elektronische Metadaten bilden vielmehr auch die fast zwingende Voraussetzung für eine Digitalisierung von Beständen und die Einbindung der Digitalisate in nationale und internationale Portale. Die Bayerische Staatsbibliothek verfolgt das Ziel, in den nächsten Jahren ihren kompletten urheberrechtsfreien Bestand – inklusive ihrer ostasiatischen Titel – digital zur Verfügung zu stellen.<sup>25</sup>

Nicht zuletzt bieten elektronische Katalogisate die Chance, Teil des semantischen Netzes zu werden. Durch die Freistellung der B3Kat-Daten als Linked Open Data in einem für das Internet geeigneten Datenformat werden die bibliographischen Daten aus ihrem »Datensilo« befreit und finden Eingang in das sogenannte »Netz der Daten«. Damit können sie in vielfältiger Weise nachgenutzt, angereichert und neu kombiniert werden. Durch Verlinkungen und Relationen zu den unterschiedlichsten im Netz vorliegenden Ressourcen und Daten sowie die Programmierung neuartiger Tools und sogenannter Mashups werden neue, intelligente Anwendungen entstehen.

Zusammenfassend kann festgehalten werden, dass sich mit dem Abschluss der Retrokonversion der Sinica- und Japonica-Kataloge – und damit der Retrokonversion der Alphabetischen Kataloge insgesamt – Nachweissituation, Sichtbarkeit und Zugänglichkeit der reichen Bestände der Bayerischen Staatsbibliothek signifikant verbessert haben.

<sup>1</sup> Vgl. Leskien, H. (Hrsg.): *Erhaltung, Archivierung und Aussonderung von Druckschriften in Bayern*. Berlin, Deutsches Bibliotheksinstitut 1998.

<sup>2</sup> Siehe hierzu Holbach, W.: *Retrokonversion und Retrokatalogisierung in den Bayerischen Staatlichen Bibliotheken*. In: *Bibliotheksforum Bayern* 34,3 (2006), S. 206–228. Die dort genannte Zahl von 6,1 Mio. noch zu konvertierender Titelaufnahmen staatlicher bayerischer Bibliotheken im Jahr 2006 reduzierte sich bis 2011 auf ca. 2,5 Mio. Aufnahmen.

<sup>3</sup> Vgl. Fabian, C.: *Der Katalog 1501 bis 1840 der Bayerischen Staatsbibliothek: Konversion, Erschließung, Überführung nach UNIMARC*, Präsentation auf CD-ROM. In: *Bibliotheksforum Bayern* 24 (1996), S. 341–369.

<sup>4</sup> Vgl. Fabian, C.; Haller, K.: *Der Image-Katalog als alternatives Modell der Konversion: Die Konversion des Alphabetischen Katalogs 1953–1981 der Bayerischen Staatsbibliothek*. In: *ZfBB* 45 (1998), S. 167–188. Sowie: Fabian, C.: *Erfassen von Daten für einen Image-Katalog: Die Konversion des Alphabetischen Katalogs 1953–1981 (IFK) der Bayerischen Staatsbibliothek*. In: *Bibliotheksforum Bayern* 25 (1997), S. 258–274.

<sup>5</sup> Zur Geschichte der Katalogkonversion an der Bayerischen Staatsbibliothek insgesamt bis 2007 siehe Fabian, C.; Holbach, W.: *Die Konversion des Quartkatalogs der Bayerischen Staatsbibliothek – Schlusstein auf dem langen Weg zu einem vollständigen maschinenlesbaren Bestandskatalog!* In: *ZfBB* 54 (2007), 3, S. 103–119.

<sup>6</sup> Die getroffenen Aussagen beziehen sich auf die sogenannten »Alphabetischen Kataloge/Formalkataloge«. Die Sachkataloge konnten bisher noch nicht konvertiert werden. Zu den historischen Sachkatalogen vgl. Rückert, I.: *Der Schrettinger-Katalog und der Alte Realkatalog*. In: Lorenz, B. (Hrsg.): *Bibliothek und Philologie: Festschrift für Hans-Jürgen Schubert zum 65. Geburtstag*. Wiesbaden 2005, S. 107–136.

<sup>7</sup> Bei Redaktionsschluss war die abschließende Einspeicherung der japanischen Titelaufnahmen in Verbund- und Lokalkatalog noch nicht erfolgt.

<sup>8</sup> Vgl. Stephan, R.: *Die Erwerbung chinesischer Literatur für die Münchner Hofbibliothek*. In: Eikelmann, R. (Hrsg.): *Die Wittelsbacher*

und das Reich der Mitte. 400 Jahre China und Bayern. München 2009, S. 449–454.

<sup>9</sup> Vgl. Kraft, E.: *Japanische Handschriften und traditionelle Drucke aus der Zeit vor 1868 in München: Neuerwerbungen der Bayerischen Staatsbibliothek. Verzeichnis der orientalischen Handschriften in Deutschland, Band 27,5*. Stuttgart 1994, S. 7–16.

<sup>10</sup> Bayerische Staatsbibliothek. *Katalog der Ostasiensammlung*. Bde. 1–6. Wiesbaden 1984.

<sup>11</sup> L.sin. bezeichnet die Fachsignatur, unter der die Sinica der Bayerischen Staatsbibliothek katalogisiert werden.

<sup>12</sup> Der Supplementkatalog erschien nie in gedruckter Form.

<sup>13</sup> <http://liblink.wordpress.com/2007/09/27/worldcat-and-rlink-changes/> (Aufruf am 20.10.2011).

<sup>14</sup> Im Bayerischen Verbundkatalog (Gateway Bayern) können chinesische, japanische und koreanische Schriftzeichen angezeigt und demnächst auch recherchiert werden. Das SISIS-Lokalsystem der BSB ist auf Unicode umgestellt; die Originalschriftlichkeit wurde mit der Version 4.1 Ende 2011 eingerichtet.

<sup>15</sup> Die Ausspeicherung der von 2007 bis 2008 in WorldCat katalogisierten CJK-Titel war im Jahr 2009 erfolgt.

<sup>16</sup> Mit dem Scannen der Katalogkarten war ein Scan-Dienstleister in München beauftragt worden.

<sup>17</sup> Die Sinica-Kataloge umfassen 106, die Japonica-Kataloge 70 Katalogkästen aus Metall.

<sup>18</sup> Seit 1990 wird an der Bayerischen Staatsbibliothek für die Katalogisierung chinesischer Titel dem internationalen Standard entsprechend Pinyin als Umschrift für das Chinesische verwendet.

<sup>19</sup> Damit wurden die Vorgaben der von der AG Kooperative Verbundanwendungen erarbeiteten verbundübergreifenden Praxisregeln zur CJK-Katalogisierung aus dem Jahr 2010, deren Anwendung im B3Kat von der Kommission für Erschließung und Metadaten auf der Sitzung am 15.4.2011 beschlossen wurde, erfüllt.

<sup>20</sup> Auch dies entspricht den Vorgaben der von der AG Kooperative Verbundanwendungen erarbeiteten Praxisregeln zur CJK-Katalogisierung. Vgl. Fußnote 19.

<sup>21</sup> Bei der Sinica-Konversion kam es in einigen Fällen – besonders bei älteren Titeln – zu einer Verwendung vermeintlich passender, in Wahrheit jedoch falscher Titelaufnahmen. Für die Japonica-Konversion galt deshalb die Vorgabe, für alle älteren Titel (Titel mit Erscheinungsjahr vor 1868) immer einen neuen Master-Record anzulegen.

<sup>22</sup> So waren z. B. folgende Leistungszahlen zu erreichen: fehlerfreies Erfassen des Verfassernamens in Originalschrift und Pinyin zu 99,95 %; fehlerfreies Erfassen der Signatur zu 99,95 %; fehlerfreies Erfassen des Hauptsachtitels in Originalschrift und Pinyin zu 99,9 %; fehlerfreies Erfassen der übrigen Texte zu 99 %; Wahl der richtigen MARC-Kategorie zu 99,5 % usw. Als Fehler galt jedes nicht korrekt erfasste Zeichen.

<sup>23</sup> Vgl. Fußnote 14.

<sup>24</sup> <http://viaf.org/> (Aufruf am 20.10.2011).

<sup>25</sup> Für den wertvollen Altbestand erfolgt dies, teilweise gefördert durch die Deutsche Forschungsgemeinschaft, innerhalb der BSB. Siehe hierzu Tabery, T.: *Altsinica Digital. Erschließung und Digitalisierung chinesischer Handschriften und Drucke an der Bayerischen Staatsbibliothek*. In: *Bibliotheksforum Bayern* 1 (2012). Der urheberrechtsfreie Bestand des 17. bis 19. Jahrhunderts wird von der Firma Google im Rahmen einer Public-Private-Partnership digitalisiert. Siehe hierzu u. a. Ceynowa, K.: *Mass Digitization for Research and Study: the digitization strategy of the Bavarian State Library*. In: *IFLA Journal* 35,1 (2009), S. 17–24 und Baumgartner, M.; Hilpert, W.: *Halbzeit – ohne Pause. Stand und Erkenntnisse der industriellen Massendigitalisierung an der Bayerischen Staatsbibliothek*. In: *B.I.T.online* 14,2 (2011), S. 133–138.

## DIE VERFASSER

**Dr. Thomas Tabery** ist stellvertretender Leiter der Orient- und Ostasienabteilung der Bayerischen Staatsbibliothek, Ludwigstr. 16, 80539 München, E-Mail: [tabery@bsb-muenchen.de](mailto:tabery@bsb-muenchen.de)

**Werner Holbach** ist Referatsleiter für Sacherschließung und Portale/Virtuelle Fachbibliotheken in der Abteilung Bestandsaufbau und Erschließung I der Bayerischen Staatsbibliothek, Ludwigstr. 16, 80539 München, E-Mail: [holbach@bsb-muenchen.de](mailto:holbach@bsb-muenchen.de)

**elektronische Metadaten  
Voraussetzung für  
Digitalisierung**

**bibliografische Daten  
als Linked Open Data  
verfügbar**