

Kooperative Leistungen bei der Bestandserschließung im virtuellen Verbund

Günter H. Todt

Ein Grundgedanke der Bibliotheksverbände ist es, die Leistungen einer Bibliothek für andere verfügbar zu machen. Dies gilt insbesondere für den ersten und wichtigsten Schritt bei der Erschließung, der Erstellung einer Titelaufnahme für ein bestimmtes Medium.

Er zeigt sich aber auch im Verlauf der weiteren Pflege dieser Titeldaten. Diese besteht im wesentlichen aus der Bereinigung von Fehlern, die bei der Katalogisierung unterlaufen können (Tipfehler etc.) sowie im Nachvollzug neuer Entwicklungen (etwa dem Wechsel von der Einteiligkeit zur Mehrteiligkeit, wie sie z.B. durch das Erscheinen eines Nachtragsbandes erforderlich wird). Auch die eventuell fälligen Änderungen in den mit dem Titel verknüpften Normdaten sind hierunter zu verstehen.

Unbedingt ist an dieser Stelle auch auf die Sacherschließung (sowohl die verbale wie die klassifikatorische) zu verweisen, zu der nicht in allen Bibliotheken ausreichend Kräfte vorhanden sind, weder in qualitativer noch in quantitativer Hinsicht. Gerade auf diesem Gebiet sind zahlreiche Bibliotheken, besonders die kleineren, auf die Leistungen anderer angewiesen.

Im traditionellen Verbund erfolgen solche kooperativen Leistungen durch verteilt vorgenommene Korrekturen und Ergänzungen im zentral vorgehaltenen Datenbestand, dem sogenannten „Union catalog“.

Ein virtueller Verbund, der keinen Union catalog besitzt, scheint daher auf den ersten Blick in einer ungünstigen Situation zu sein, da für ihn solche kooperativen Leistungen ausgeschlossen zu sein scheinen. Man sollte meinen, dass alle oben angesprochenen Maßnahmen (Fehlerbeseitigung, Nachvollzug von neueren Entwicklungen in der Titelbeschreibung und bei den verknüpften Normdaten) entweder gar nicht stattfinden oder nur insoweit, wie sie in den jeweiligen Lokalkatalogen im Einzelfall erkannt werden.

Unstrittig ist, dass jede datenpflegerische Maßnahme im virtuellen Verbund in den jeweiligen Lokalkatalogen geschehen muss, also statt in einer zentralen Datenbank in n verteilten Datenbanken vonstatten gehen muss.

Es liegt daher nahe über Maßnahmen nachzudenken, wie solche datenpflegerischen Maßnahmen in bestimmter Weise automatisiert werden können, um die Zufälligkeit zu minimieren, mit der mal hier, mal da eine Korrektur durchgeführt wird oder eben auch nicht.

Die folgenden Ausführungen wollen Fingerzeige geben und auch von bereits getätigten Erfahrungen berichten, auf welche Weise dieses strukturelle Defizit ausgeglichen werden kann. Es sollen zu den angesprochenen drei Problemkreisen (Tipp-)Fehlerbeseitigung, Aktualisierung von Titelbeschreibungen und Nachtrag verbaler und klassifikatorischer Sacherschließung, Lösungswege vorgestellt werden.

1. Änderungen in verknüpften Normdaten

Der Nachvollzug solcher Entwicklungen erscheint, was den derzeitigen Stand sowohl der Technik wie der Usancen unserer bestimmenden Datenlieferanten angeht, relativ einfach zu sein, auch wenn er noch nicht überall implementiert ist.

Aus den wöchentlichen Änderungsmitteilungen der drei Normdateien gehen sowohl alle Änderungen von Normdatensätzen hervor als auch Substitutionen von Normdaten. Unter Substitution ist hier zu verstehen, dass ein altes Normdatum nicht mehr vorhanden ist und durch ein oder mehrere andere Normdaten ersetzt wird.

Liegt hierbei eine 1:1 Konstellation vor (also eine Umlenkung, wie ganz überwiegend in der PND vorkommend) so kann diese maschinell nachvollzogen werden. Lediglich im Falle einer 1:n Beziehung bleibt intellektuelle Nacharbeit unvermeidbar. Letzteres ist etwa bei der Auflösung eines SWD-Schlagworts die Regel und erfordert normalerweise das Neudurchdenken der Verschlagwortung der betroffenen Titel.

Der Löwenanteil der in den wöchentlichen Änderungsmitteilungen angesprochenen Normdaten weist glücklicherweise keine Substitution in der Form 1:n auf, sondern beschränkt sich entweder auf eine bloße Änderung im Normdatensatz selbst oder verlangt eine Umlenkung. Alle diese Änderungen können maschinell ohne weiteres intellektuelles Zutun nachvollzogen werden.

Ob diese Umsetzung der wöchentlichen Änderungsmitteilungen nun in einer zentralen Datenbank vor sich geht (und dann von dort in die angeschlossenen Lokalkataloge diffundiert) oder gleich in den jeweiligen Lokalkatalogen vor sich geht, dürfte per saldo nahezu auf ein- und dasselbe hinauslaufen. Gängige Bibliothekssysteme (wie z.B. SISIS) verfügen über hinreichend leistungsfähige Tools, um solche Arbeiten ohne größeren Aufwand jeweils vor Ort leisten zu können.

Welch erstaunliche Korrekturen hier im Einzelfall sich ergeben können, sei an einem Beispiel aus der Praxis erläutert:

Aus der Feder der amerikanischen Wirtschaftswissenschaftler Robert S. Pindyck und Daniel L. Rubinfeld gibt es ein Lehrbuch der Mikroökonomie,

dessen 4. Aufl. 1998 auch auf deutsch erschienen ist (ISBN 3-486-22358-5). Dabei hat der Verlag Oldenbourg allerdings den Namen des 2. Autors fehlerhaft gedruckt und ihn in „David L. Rubinfeld“ umgetauft.

Bis zum April 2004 (also 6 Jahre nach Erscheinen des Buches) wurde dieser Fehler in den folgenden Bibliotheken und Verbänden erkannt¹:

BVB, HBZ, SWB, GBV (dort gibt es 3 Titelaufnahmen für das Werk, gemeint ist hier die „Hauptaufnahme“, an der 24 Bibliotheken angesiegelt sind) sowie im Bereich des KOBV von der HU-, der TU- und der FU-Berlin, (wobei, nebenbei gesagt, in insgesamt 2 Fällen sogar in verfälschender Weise die Vorlageform korrigiert wurde).

Nicht erkannt wurde der Fehler hingegen in:

HEBIS, DDB (beide Standorte!); ZB der Wirtschaftswissenschaften Kiel (GBV) sowie aus dem Bereich des KOBV von der Universität Frankfurt/Oder; der FH Wirtschaft Berlin, der FH Lausitz und der TFH Wildau.

Eine „neutrale“ Position nimmt gewissermaßen die Uni Potsdam (früher KOBV, jetzt GBV) ein: Dort hat man auf die Erfassung des 2. Verfassers verzichtet.

Der Befund bestätigt nebenbei in groben Zügen die Vermutung, dass isoliert arbeitende Bibliotheken (DDB) und kleinere Bibliotheken am wenigsten Gelegenheit haben, auf Fehler dieser Art zu stoßen, während in den größeren Einheiten (in den 3 Berliner Universitätsbibliotheken und allen Verbänden außer HEBIS) dies sehr wohl geschah.

Ende April 2004 floss die Aufdeckung dieses Druckfehlers des Verlags endlich in die PND ein, so dass es dort zu einer Umlenkung von „Rubinfeld, David L.“ auf den Satz von „Rubinfeld, Daniel L.“ kam. Alle Bibliotheken, die diese Umlenkung im Rahmen der PND nachvollzogen, kamen also automatisch in den Genuss der Fehlerkorrektur. Die Eleganz einer solchen Lösung zeigt sich besonders im Vergleich des Aufwands: Während in den oben genannten Bibliotheken und Verbänden der Druckfehler des Oldenbourg Verlags an 7 verschiedenen Stellen jeweils autonom entdeckt und ausgebügelt werden musste, erlaubt die PND-Umlenkung theoretisch in 8 Bibliotheken bzw. Verbänden die Korrektur ohne jedwedes weiteres Zutun!

2. Änderung in der bibliographischen Beschreibung eines Titelsatzes

Titelaufnahmen sind stets in der Gefahr, durch Tippfehler oder andere Fehlleistungen den erwünschten Qualitätsstandard nicht zu erreichen. Zwar darf man davon ausgehen, dass überall mehr oder minder erfolgreich Fehler er-

1 Basierend auf einer KVK-Recherche Ende April 2004

kannt und dann ausgebügelt werden, doch ist dieser Pflegeaufwand notwendigerweise beschränkt. Es fehlt zumeist an der Kapazität, um solchen Dingen nachzugehen oder an der Information darüber, dass hier überhaupt etwas anzupassen ist (etwa beim Wechsel von Einteiligkeit nach Mehrteiligkeit oder gar hin zu fortlaufender Erscheinungsweise).

Hier wäre viel gewonnen, wenn die Daten der Deutschen Bibliothek, die ja einen beträchtlichen Anteil in den Lokalkatalogen des virtuellen Verbunds KOBV ausmachen, in einer ähnlichen Weise maschinell gepflegt werden könnten, wie dies heute schon mit den ZDB-Daten möglich ist.

Während die ZDB-Daten, analog wie die Normdaten, anhand der Nutzung der wöchentlichen Änderungsmittellungen fortlaufend auch in den Lokalkatalogen auf den neuesten Stand gebracht werden können, ist dies bei DNB-Daten leider nicht der Fall. Es ist andererseits bekannt, dass die Deutsche Bibliothek alljährlich eine nicht unbeträchtliche Zahl von Korrekturen in ihrem Titeldatenbestand vornimmt, davon den nachnutzenden Bibliotheken aber nichts weitergibt.

Während ein traditioneller Verbund auf die Information der Titeldatenkorrekturen der DDB wohl weitgehend verzichten kann (dort geht man davon aus, dass die notwendigen Änderungen im wesentlichen von den Verbundmitarbeitern erkannt und selbst getätigt werden), liegt der Fall im virtuellen Verbund etwas anders. Wie das Beispiel in Kap. 1 schon zeigte, kann aber nicht davon ausgegangen werden, dass ein Fehler in DNB-Daten, die von mehreren Einzelbibliotheken jeweils autonom nachgenutzt werden, überall erkannt und ausgebügelt wird. Erfahrungsgemäß ist eher das Gegenteil der Fall. Es wäre daher von großem Vorteil, ließe DDB solche Titeldatenkorrekturen ihren Nachnutzern, mindestens aber solchen in virtuellen Verbänden, zukommen. Dies würde in der Praxis zur Eliminierung einer Vielzahl von Fehlern führen, die durch die Nachnutzung defizitärer DNB-Daten entstanden sind.

Intellektuell nachzuvollziehen wären dabei lediglich Fälle wie Satzartänderung (was in der Regel die Folge ist beim Wechsel von Einteiligkeit nach Mehrteiligkeit oder gar hin zu fortlaufender Erscheinungsweise) oder wenn DNB-Sätze gemäß lokalen Konventionen bereits verändert worden waren. Für diese Fälle wären natürlich vor Ort geeignete Selektionskriterien zu entwickeln, um hier keine unerwünschten Verschlimmbesserungen zu provozieren. In der Masse der Fälle aber, in denen selbst bei kritischster Sicht der DNB-Satz keine Änderung verlangt, könnten die in Frankfurt/M. und Leipzig getätigten Titeldatenkorrekturen maschinell nachvollzogen werden. Auf diese Weise würde sich auch das notwendige Arbeitsvolumen verringern, das aus dem Nachvollzug von Normdatensubstitutionen der Konstellation 1:n resultiert. Anstatt, wie in Kapitel 1 diesbezüglich formuliert, das „Neudurchdenken der Verschlagwortung bestimmter Titel“ intellektuell leisten zu müssen, könnte die veränderte verbale Sacherschließung zumindest für die DNB-gestützten Titeldaten ma-

schinell nachvollzogen werden. Es müsste dann nur noch die Verschlagwortung der Nicht-DNB-gestützten Titel aktualisiert werden. Gleiches gilt natürlich für die im DNB-Datenbestand vollzogenen Aufspaltungen von PND-Sätzen, die gleichfalls für alle DNB-gestützten Titeldaten maschinell nachvollzogen werden könnten.

Leider sind diese Verfahren derzeit nicht zu realisieren, da die Deutsche Bibliothek nicht willens ist, die von ihr getätigten Titeldatenkorrekturen weiterzugeben. Der Hoffnung, dass sich diesbezüglich eine Änderung ihrer Haltung ergibt, darf hier aber Ausdruck gegeben werden.

3. Nachträgliche Sacherschließung

Soweit bekannt geschieht dies im traditionellen Verbund durch die sukzessive Beschlagwortung der Verbundtiteldaten, an dem alle mit lokalem Bestand angesiegelten Bibliotheken partizipieren.

Im virtuellen Verbund ist dies in der Form natürlich ausgeschlossen, die verbale Sacherschließung (soweit diese nicht lokal bzw. beim Import geeigneter Fremddaten im Zuge der Katalogisierung erfolgen kann) muss durch retrospektive Nutzung von Fremddaten erreicht werden.

Für DNB-gestützte Titeldaten ist dies relativ einfach zu bewerkstelligen. Es wird im Lokalkatalog eine Selektion durchgeführt nach den Kriterien: DNB-Identnummer vorhanden, aber keine Verschlagwortung.

Diese DNB-Identnummern werden der Verbundzentrale übermittelt und von dieser im Datenpool „DNB-Daten“ selektiert und der anfordernden Bibliothek zurückgeschickt, welche wiederum aus den erhaltenen Titeldaten diejenigen Felder, an denen sie in diesem Fall interessiert ist², in den Lokalkatalog importiert. Dabei dient die DNB-ID als Updatekriterium, so dass es zu keinen Problemen kommen kann.

Für Nicht-DNB-gestützte Daten ist der Prozess der nachträglichen Sacherschließung etwas diffiziler: Hier nutzen wir in Frankfurt an der Oder zur Zeit lediglich die ISBN als Updatekriterium. Für die Zukunft wäre es wünschenswert, ließe sich ein Programm entwickeln, das auch für Titeldaten ohne ISBN einen weitestgehend eindeutigen Identitätsvergleich mit anderen Datenbanken gestattet.

Für den ISBN-gestützten Abgleich werden zunächst alle nicht verschlagworteten Titeldatensätze³ selektiert und in derselben Weise wie oben dargestellt zunächst am Datenbestand „DNB-Daten“ der Verbundzentrale überprüft.

2 Die UB der Europa-Universität Frankfurt/Oder übernimmt hierbei neben dem eigentlichen Schlagwortsegment (MAB2 900ff.) auch die DNB-Klassifikation (MAB2 700, Indikator d).

3 Wobei wir uns hier auf Titelsätze, die nach RSWK in formaler Hinsicht überhaupt verschlagwortungsfähig sind, beschränken. Untersätze sind also von dieser Proze-

Nachdem nun bei allen lokalen Datensätzen eine ggf. verfügbare Verschlagwortung der DDB nachgeführt werden konnte, wird dieser Arbeitsschritt ein 2. mal unternommen, nun aber erfolgt der Abgleich nicht an den DDB-Daten, sondern vernünftigerweise am Bestand der BVB-Daten. Der bayerische Verbund verfügt bekanntermaßen im Vergleich aller Verbünde über die umfanglichste Verschlagwortung, so dass die Chance, dort Sacherschließungsdaten zu finden, sehr hoch ist und in einer Vielzahl von Fällen den lokalen Titelsatz entsprechend aufbessern kann.

Ein Problem, das beim Update über ISBN auftaucht, ist das Vorliegen mehrerer ISBNs, die auch noch, je nach Katalog, in unterschiedlicher Anzahl und Reihenfolge erfasst sein können. Um in solchen Fällen das etwaige Überschreiben einer lokal erfassten ISBN durch eine andere auszuschließen, wird für das Update die ISBN in der Titelladedatei durch die Katalognummer des Zielsatzes ersetzt und das Feld 540 (ISBN) aus der Ladedatei völlig gelöscht. Dadurch bleibt das lokale ISBN-Feld bei dieser Aktion gänzlich unangetastet.

Ein gesondertes Problem ergibt sich, wenn Sacherschließungsdaten aus anderen Sprachkreisen übernommen werden sollen. Dies gilt insbesondere dann, wenn dabei die Entitäten der bibliographischen Beschreibung auseinanderfallen.

Diesem Sachverhalt musste Rechnung getragen werden beim Import der Sacherschließungsdaten der LoC⁴, wie er in den ersten Monaten dieses Jahres an der UB der Europa-Universität Frankfurt an der Oder stattfand. Der in diesem Zusammenhang einschlägige Unterschied zwischen einem RAK-geführten deutschen Katalog und dem AACR-geführten Katalog der LoC besteht bekanntermaßen im Verzicht der letzteren auf das Anlegen von Untersätzen. Es musste also bei der Nachführung der LoC-Sacherschließung zwischen ein- und mehrteiligen Werken genau unterschieden werden.

In der Praxis wurden zunächst alle einbändigen Monographien⁵ (natürlich nur, insoweit mit ISBN versehen) an der LoC-Datenbank des KOBV überprüft. Nach diesem Schritt lieferten wir die ISBNs von Untersätzen zur Abprüfung, importierten aber die rückgelieferten Informationen nicht in den Untersatz, sondern in den direkt übergeordneten Hauptsatz. Ein Problem ergab sich in etlichen Fällen dadurch, dass zu den ISBNs mehrerer zusammengehöriger Untersätze verschiedene Informationen geliefert wurden. Da diese Fälle intellektuell hätten überprüft

dur ebenso ausgenommen wie analytische Titelaufnahmen und die fortlaufender Sammelwerke.

- 4 Wir haben in diesem Zusammenhang neben den Subject headings der LoC auch deren beide Klassifikationen sowie die elektronischen Referenzwerke, die die LoC-Daten so reichlich nachweisen, mitimportiert.
- 5 Selektionskriterium: Erscheinungsform (MAB2 .051 Pos. 0) m bzw. s

werden müssen⁶, woran denn nun diese Unstimmigkeit liege, wurde darauf verzichtet, solche Sätze retrospektiv mit LoC-Daten anzureichern.

Zu guterletzt kamen schließlich solche mehrteiligen Werke an die Reihe, die bis jetzt noch nicht hatten versorgt werden können, aber selbst eine Gesamt-ISBN aufweisen. Hier erfolgte das Update gleichermaßen wie bei den einteiligen Titelaufnahmen.

Abschließend sei noch auf eine weitere, unerwartete Besonderheit hingewiesen: Erwartet worden war, dass in einigen Fällen für eine gelieferte ISBN in der LoC-Datenbank mehr als eine Entsprechung gefunden würde. Dies ist in all jenen Fällen unausweichlich, in denen ein Verlag ein- und dieselbe ISBN für mehr als einen Titel vergeben hat. Für diesen Fall war vorab vereinbart, gleichfalls auf den Import der LoC-Daten zu verzichten. Zwar wäre es weitestgehend unschädlich, die Sacherschließung der LoC dann zu übernehmen, wenn die beiden Bücher mit derselben ISBN nur 2 verschiedene Auflagen desselben Werkes betreffen, doch konnte dieser Fall maschinell nicht unterschieden werden von dem allerdings erheblichen, wenn mit derselben ISBN 2 Bücher völlig verschiedenen Inhalts erschienen sind.

Überraschend war dann jedoch, wie viele solcher uneindeutigen Fälle in der Praxis auftraten. Eine Überprüfung ergab die erstaunliche Tatsache, dass zu der gesuchten ISBN jeweils 2 Sätze im Datenpool LoC des KOBV vorhanden waren, die eindeutig als dublett erkennbar waren. Eine Nachprüfung des Entstehens dieser Dubletten wiederum ergab, dass über einen langen Zeitraum die LoC selbst bei der Korrektur eines Titeldatensatzes dessen Identnummer mit einem diesbezüglichen Suffix versehen hatte. Dieses nicht ganz DV-gerechte Vorgehen der amerikanischen Kollegen hatte natürlich zur Folge, dass nun zwei Datensätze mit unterschiedlichen Identnummern vorliegen, die keine Maschine dieser Welt als identisch erkennen kann.

Da die große Anzahl dieser Fälle eine intellektuelle Überprüfung ausschloss und maschinell nicht erkennbar war, ob nun die Konstellation „zwei verschiedene Bücher mit gleicher ISBN“ oder die der durch Suffigierung der Identnummer entstandenen Dublette vorlag, wurde auch in diesen Fällen auf den Import der LoC-Erschließungselemente verzichtet. Wir konnten dennoch ca. 100.000 Titelsätze anreichern, d.h. knapp 50% des lokal vorhandenen Monographienbestands.

6 Ein denkbarer Fall, der in der Praxis wahrscheinlich gar nicht so selten sein dürfte, liegt dann vor, wenn ein Gesamtwerk nach RAK als mehrbändiges Werk katalogisiert wird, nach AACR aber kein Gesamtwerk angenommen wird, sondern die einzelnen Teile jeweils als einbändige Monographie behandelt werden.