

Pilotmigration UB des BIZ der Europa-Universität Viadrina – Frankfurt (Oder) in den BVB

Henning Klaufß

Ende 2007 wurde die „Vereinbarung zur Strategischen Allianz zwischen dem Bibliotheksverbund Bayern (BVB) und dem Kooperativen Bibliotheksverbund Berlin-Brandenburg (KOBV)“ beschlossen.¹ Infolgedessen werden die Kataloge von vielen Bibliotheken aus Berlin und Brandenburg zusammengeführt. Die beteiligten Bibliotheken² sind teils SISIS-, teils Aleph-Bibliotheken. Die UB des Bibliotheks- und Informationszentrum (BIZ) Frankfurt (Oder) hat für die SISIS-Bibliotheken die Pilotierung übernommen.

Es soll im Folgenden nicht das Verfahren der Beschlussfassung in dieser Sache thematisiert werden, sondern es sollen einige Anmerkungen gemacht werden zur Berlin-Brandenburgischen Bibliotheksgeschichte bzgl. Verbundkatalogisierung und lokal autonomer Katalogisierung im Hinblick auf die Allianz sowie gewisse Rahmenbedingungen und einiger Details der Migration bei uns.

Im Jahr 1994 trat das Land Brandenburg dem „Bibliotheksverbund Berlin-Brandenburg“ (BVBB), davor „Berliner Bibliotheksverbund“, BBV), (ehem. „Berliner Monographienverbund“ bei. Dieser Anschluss lief für die UB der EUV so: Es wurden keine bestehenden Titelaufnahmen, d.h. keine alten Bestände, in den BVBB-Katalog gemeldet, sondern nur neue. Um Doppelarbeit zu vermeiden und die Titelaufnahmen sofort im Katalog zu haben, haben wir in Frankfurt (Oder) Titelaufnahmen mit der sog. Softcopy in den Lokalkatalog eingearbeitet. Es ging bei dieser Migration deshalb nicht um die vollständige Übernahme der Titelaufnahmen der UB der EUV, sondern „nur“ um die Übernahme der neu erstellten Titelaufnahmen. Es gab daher in unserem Katalog damals zugleich BVBB- und nicht-BVBB-gestützte Titelaufnahmen.

Schon vor Abwicklung des BVBB 1999 arbeitete die UB der EUV lokal autonom: Der für uns abschließende Filetransfer, durch den uns die kompletten Titel- und Normdaten geliefert wurden, fand Anfang 1998 statt. Als im April 1997 der KOBV gegründet wurde, gehörte es zu dessen Prämissen, keinen durch einheitliche Verbundregeln einheitlich gestalteten Verbundkatalog zu generieren, sondern die

1 Siehe Notiz in *Bibliotheksdienst* 42 (2008), Heft 3, S. 300–301.

2 Am KOBV nehmen u.a. teil in Berlin: Bibliothekssystem der Freien Universität Berlin, UB der Humboldt-Universität zu Berlin, UB der Technischen Universität Berlin, UB der Universität der Künste Berlin sowie weitere Hochschulbibliotheken; in Brandenburg: UB der Europa-Universität Viadrina Frankfurt (Oder), UB der Brandenburgischen Technischen Universität Cottbus, HSB der Hochschule für Film und Fernsehen „Konrad Wolf“ Potsdam sowie weitere Hochschulbibliotheken.

lokal autonom erstellten Titelaufnahmen nachträglich in einem Verbundkatalog zu integrieren. Für die 1999 in Betrieb genommene Suchmaschine wurde den KOBV-Partnern in Bezug auf die Erschließung daher ein größerer Handlungsspielraum gegeben; zu beachten war der „Gemeinsame KOBV-Standard“ (Mitte 1999). Dieser sah die Erfassung bestimmter obligatorischer und optionaler Elemente vor. Obligatorische Elemente waren z.B. Personennamen, Erscheinungsjahr und Hauptsachtitel (HST) mit Zusatz. Hierzu eine kleine Anekdote: Der Zusatz zum HST sollte gemäß KOBV-Standard „ohne Abkürzungen“ angegeben werden. Diese bei uns umfangreich getätigte Arbeit für eine Verbesserung des Retrievals durch Auflösung der Abkürzungen wird durch diese Migration hinfällig, weil bei u.a. den Titelaufnahmen, die wir aus dem BVB genutzt haben, an dieser Stelle umfangreich abgekürzt wurde und wir nach der Migration erneut die ehem. lokal aufgelösten Abkürzungen wieder im Katalog haben, also ein Stück Arbeit kaputt gemacht wurde!

Ende 2007 wurde das Modell der nachträglichen Aggregation von autonomen Lokalkatalogen zugunsten eines traditionellen Verbundkataloges aufgegeben und die Allianz mit dem BVB beschlossen. Die UB der EUV hat hierfür die Rolle der Pilotbibliothek übernommen, da deren Katalog besonders gute Startbedingungen bietet (intensive Datenpflege/Vorhandensein von vielen, nicht überall vorfindlichen Informationen, z.B. Erscheinungsform, überregionale IDs ..., also allerlei zweckdienliche Ansatzpunkte für eine Migration).

Für die Migration wurde ein Werkzeugkasten eingesetzt, mit dem Dubletten erkannt, ggf. eine Anreicherung der Titelaufnahmen des BVB mit denen aus Berlin und Brandenburg (Merge) sowie die Erstellung neuer Datensätze geregelt werden sollte. Bezüglich des vom KOBV für die Migration eingesetzten Werkzeugkastens gab es keine Alternative und es wurde vorab festgelegt, dass nichts programmiert wird, nur Voreinstellungen geändert werden. Dieser Umstand hat zu gewissen Konsequenzen geführt, die der Absicht der „Nordlichter“, bestimmte Kataloginhalte zu bewahren, und der vertraglichen Artikulation davon, manchmal abhold waren.

Für diese Migration wurde ein WIKI eingerichtet, in dem etliche Details der Migration protokolliert sind. Die an der EUV genutzten, für die Migration modifizierten Feldstrukturtabellen finden sich unter der URL <http://www.addon.ub.euv-frankfurt-o.de/sisisaf/saf-bb.htm>.

Die UB Frankfurt (Oder) hat bereits eine ganze Reihe von unterschiedlichen Migrationen hinter sich. So wurde der gesamte Datenbestand (abgesehen von den vielen „kleinen Updates“ der Software) in SISIS von zuerst SE auf ONL, dann von ONL auf SunRise migriert – alle diese Migrationen liefen hausintern ab. Als es in den BVBB ging, wurde dagegen nicht der komplette Datenbestand migriert, sondern, wie bereits geschildert, nur die neuen Titelaufnahmen im BVBB erstellt bzw.

eine bloße Ansigelung durchgeführt. Die in der Zeit der Zusammenarbeit von uns mit dem BVBB erstellten Titelaufnahmen sind hochgradig normiert. Es hat zu fast jeder Kategorie umfangreiche Anleitungen gegeben.

Es ist jetzt das erste Mal, dass der *komplette* Katalog der UB der EUV in einen anderen integriert wurde. Nach dem Ende des BVBB war die UB der EUV wie andere Bibliotheken in Berlin und Brandenburg lokal autonom – abgesehen von der Staatsbibliothek und der UB Potsdam, die in den GBV gegangen sind. In dieser Zeit haben sich in den einzelnen lokalen Katalogen eine Reihe von Usancen/Besonderheiten entwickelt (Beachtung von Codes/Abgleich Schlagwörter mit der SWD/überregionale IDs/Datenpflege...). Nun sollen Bibliotheken mit einer Reihe von lokalen Besonderheiten in einen Verbundkatalog implementiert werden.

Es wird hierbei so sein, dass es in *jedem* Katalog unmittelbar nach der Migration zum einen gewisse Verluste, also Verschlechterungen und Reduktionen von Informationen, zum anderen gewisse Verbesserungen und Bereicherungen geben wird. Es wird lokal quantitativ und qualitativ unterschiedlich sein, in welchem Katalog zum einen das Plus und zum anderen das Minus ausfällt. Wenn ich von „gewissen“ Verlusten und gewissen Verbesserungen spreche, so kann das „gewisse“ manchmal eher Marginales, manchmal aber auch Gravierendes bedeuten.

Ich hoffe natürlich, dass im Allgemeinen mittel- und langfristig das Plus überwiegt, aber ich warne davor, zu glauben, dass es in den Lokalkatalogen selbst kurzfristig an keiner Stelle ein Minus geben wird. Es ist bei jeder Angleichung so, dass aufsteigt, wer unter dem Durchschnitt war und sinkt, wer über dem Durchschnitt war. Diese Homogenisierung passt nicht zu dem ansonsten von Hochschulen geforderten Wettbewerb, der ja auch an das Herausputzen von Besonderheiten gebunden ist!

Auch für die Bibliotheken, für die die Migration zunächst ein Verlust ist, und dazu rechne ich in vielen Punkten die der Europa-Universität Viadrina, entsteht hoffentlich mittel- und erst recht langfristig ein Gewinn. Die Realisierung dieser Hoffnung ist daran gebunden, dass das, was der Verbund für uns leistet, über die Bereitstellung von Fremddaten usw. hinausgeht (die standen uns nämlich bereits bisher in erheblichem Umfang zur Verfügung), sondern dass die Verbundzentrale Aufgaben erfüllt, die unseren Benutzern dienen und zu denen wir lokal entweder finanziell und/oder aufgrund der arg dünnen Personaldecke kaum in der Lage sind.

Die EUV hat im Mai 2008 beim BVB einen Antrag gestellt, ihren Katalog bei der Migration dergestalt zu behandeln, dass die Titelaufnahmen der EUV 1:1 erhalten bleiben, nicht aus Eitelkeit, sondern weil die EUV einen hochgradig homogenen und nach aktuellem Regelwerk erstellten, ohne Retrokonversion zustande gekommenen und in hohem Maße auf den Import von guten Fremddaten gestützten Katalog hat. Im Fall einer Dublettenerkennung wollte die EUV also Testsieger

sein. Die Antwort des BVB war (abgesehen von bestimmten Anreicherungen) weitgehend ein „Nein“, weil es angeblich die „Logik des Verfahrens“ so vorsähe. Eine vollständige Ersetzung von bereits im BVB bestehenden Titelaufnahmen durch die der EUV wurde abgelehnt. Der BVB hat sich also als Testsieger definiert:

- Ein nicht als dublett erkannter Datensatz führt zur Neuanlage eines Datensatzes, der so erstellt wird, wie er angeliefert wurde.
- Wenn eine Titelaufnahme im BVB vorhanden ist, erfolgt eine Ansigelung, ggf. mit Anreicherungen

Eine Anreicherung wird in folgenden Fällen vorgenommen:

- Sacherschließungsinformationen: Schlagwörter werden ergänzt, wenn im BVB keine vorhanden sind; sind dort welche vorhanden und werden neue geliefert, werden die neuen, und d.h. hier die aus Berlin-Brandenburg, weggeworfen. Das liegt maßgeblich an dem Umstand, dass es mit dem genutzten Werkzeugkasten keinen sog. „Kreuzvergleich“ gibt, der es ermöglicht, eine Identität der in der einen Datenbank z.B. im 1., in der anderen Datenbank im 3. Schlagwortfeld vorhandenen Schlagwortkette festzustellen. Neue Notationen werden stets addiert.
- Codierungen in begrenztem Umfang (Sprachencode)
- Überregionale ID-Nummern
- Fußnoten: Im BVB wird überwiegend das Sammelfeld für Fußnoten aller Art (MAB 501) genutzt; dieses wird um die spezifischen MAB-Felder für unterschiedliche Fußnotenarten angereichert. Die damit u.U. verbundene Redundanz wird hingenommen.
- URLs für Elemente der Kataloganreicherung, also Links zu ToC etc.

Inwieweit diese Regelung, z.B. bzgl. der verbalen Sacherschließung, der „Vereinbarung zur Strategischen Allianz zwischen BVB und KOBV“, insbesondere Anlage 3: „Aufbau des gemeinsamen Verbundkataloges“ entspricht, ist fraglich.³

Es gibt also drei Fälle, wenn eine Titelaufnahme eines „KOBV-Nordlichtes“ im Katalog des BVB implementiert wird:

- 1) keine Dublette gefunden → Neue Titelaufnahme im BVB
- 2) Dublette gefunden → Titelaufnahme des BVB ist „Sieger“ *ohne* Änderungen durch Anreicherungen
- 3) Dublette gefunden → Titelaufnahme des BVB ist „Sieger“ *mit* Änderungen durch Anreicherungen.

3 S. Vereinbarung zur Strategischen Allianz zwischen BVB und KOBV, Anlage 3 Aufbau des gemeinsamen Verbundkataloges: „Auf der Basis von Datenanalysen werden zwischen den Vertragspartnern Verfahren abgesprochen, um wertvolle Informationen aus dem Datenbestand des KOBV (Sacherschließungsdaten, Regelwerkskorrekturen etc.) in den BVB-Katalog einzubringen.“

Die Ermittlung von Dubletten (Match) verläuft nach einem differenzierten Verfahren, dessen Ergebnisse zwar gut, aber auch verbesserungsfähig sind: Das eingesetzte Programm zur Match-Bestimmung weist Defizite im Bereich der materialspezifischen Codes und der Erscheinungsform auf. Erstere werden gar nicht berücksichtigt, letztere teilweise in ähnliche (aber nicht identische!) Gruppen zusammengefasst. Dies führte, auch im Zusammenhang mit anderen Festlegungen, zu prinzipiell vermeidbaren Kollateralschäden. So stellte sich z.B. die Unmöglichkeit, irgendeinen Satz mit einer übergeordneten Titelaufnahme (ÜG), die die Erscheinungsform p oder z hat, zu verknüpfen, erst im Nachhinein heraus. Im Vorfeld hatte es lediglich geheißen, dass nur U-Sätze auszunehmen seien, so dass nur diese von uns vorab ausgefiltert wurden.

Weitere Schäden traten insbesondere auf durch die große Anzahl der in der Zieldatenbank erzeugten Dubletten. Diese können dadurch entstehen, dass der migrierende Satz und der BVB-Zielsatz unterschiedlich sind (andere Erschließungstiefe), oder dass das Matchprogramm durch bestimmte Konstellationen ausgetrickst wird. Dies gilt besonders für fortlaufende Sammelwerke.

Im Falle von Verknüpfungen entstand v.a. ein Durcheinander bei mehrbändig begrenzten Werken in gezählter Serie, da die strukturelle Lösung dieses Sachverhalts im BVB signifikant von der in der DNB und dem Lokalkatalog abweicht.

Bei mehrbändig begrenzten Werken bleibt der bis zur Migration lokal existente Zusammenhang ÜG und Unterordnungen erhalten. Bei Schriftenreihen und Stücktiteln aber wird dieser Zusammenhang evtl. zerstört, weil ansonsten zig Dubletten der Titelaufnahmen von Schriftenreihen erzeugt worden wären: Die vor der Migration an einer ÜG hängenden Stücktitel werden nach der Migration evtl. an mehreren ÜGs hängen.

Von dem Sachverhalt „X Datensätze haben in der EUV y Tage gebraucht“ darf nicht voreilig 1:1 auf sonstige Lokalsysteme umgerechnet werden, denn es gibt neben der Anzahl der Datensätze allerlei sonstige lokale Besonderheiten, die zu gravierenden Differenzen führen können: Welche Felder sind wie gefüllt (der Flaschenhals des Verfahrens ist die Indexierung!), ist Datenpflege betrieben worden, welche Hardware ist lokal vorhanden usw.

Die Migration aller Kataloge von SIS-Bibliotheken soll nach dem bei der EUV entwickelten Schema verlaufen. Die Migration der Kataloge von Aleph-Bibliotheken werden auf diesen Erfahrungen aufbauen. Wir empfehlen den Lokalsystemen, bei dem Verfahren nicht auf eigene Kontrolle zu verzichten, um sicherzustellen, dass nicht irgendwelche Erfahrungen, die mit der Migration des EUV-Kataloges gemacht wurden, voreilig verallgemeinert werden! Die lokalen Titelaufnahmen sollten auf jeden Fall in der Exportumgebung angesehen werden, d.h. bevor die teils neuen, teils „nur“ gematchten, teils gematchten und gemergten (angereicherten) Datensätze in den jeweiligen Lokalkatalog gespielt werden. Der Fall „neuer Da-

tensatz“ dürfte gerade bei kleinen Bibliotheken relativ selten und ohnehin weitestgehend unproblematisch sein; der Fall „gemergter Datensatz“ ist der Standard und insbesondere der kann Probleme mit sich bringen.

Der Zeitplan des KOBV sieht vor, dass im 1. Quartal 2009 die Migration sämtlicher SISIS-Bibliotheken abgewickelt sein wird. Pläne soll man (erst recht, wenn man sie selbst mitgestaltet hat) ernst, aber nicht als Dogma nehmen: *Absolute Priorität für eine jede Bibliothek hat das Interesse von Benutzern am Katalog, nicht der Zeitplan des KOBV!* Die die EUV betreffenden Zeitpläne wurden zwischen EUV und KOBV abgestimmt und haben sich als richtig bzw. machbar erwiesen; notwendige Verschiebungen waren lange vorher bekannt!

Es war einmal vorgesehen, die Daten der Bibliotheken paketweise zu migrieren. Das wurde geändert: Um den Besonderheiten der Daten der einzelnen Bibliotheken besser gerecht zu werden, werden die Daten der Bibliotheken jeweils einzeln migriert, d.h., dass das gesamte bei uns praktizierte Verfahren, also Einfrieren der Katalogisierung, Export der Daten, Dublettencheck, Merge, Rückführung der BVB-IDs in SISIS usw. in jeder Bibliothek einzeln durchgeführt wird.

Voraussetzung für eine gelungene Migration ist in jedem Fall eine ausführliche Datenanalyse: Sind y-Sätze vorhanden, gibt es unverknüpfte U-Sätze usw.? Dieser Datenanalyse muss sich dann natürlich eine konsequente Datenbereinigung anschließen. Der hierfür erforderliche Aufwand darf nicht unterschätzt werden! (Diese Datenanalysen sind auch dann durchzuführen, wenn ein ASP-Dienst in Anspruch genommen wird.) Diese Datenanalyse kann und sollte MABLE-gestützt („MAB lässig evaluiert“) durchgeführt werden.

Die Testbedingungen für die Generalprobe waren alles andere als optimal: Es gab verschiedene Datenbanken, an denen verschiedene Aspekte untersucht werden konnten; es war der EUV in manchen Fällen nicht klar, was wo und wie untersucht werden sollte; dann fiel der CITRIX-Server aus ... Es empfiehlt sich, dass Lokalsysteme sich auch für die Testphase planmäßig hinreichend Zeit nehmen und diese in die Zeitpläne einbauen, um einen ausreichend große Puffer zu haben.

Die für die Katalogisierung im BVB durchzuführenden Schulungen sollten möglichst unmittelbar vor der Katalogisierung im BVB durchgeführt werden – sonst ist die Gefahr des Vergessens zu groß. In der EUV ist es so eingerichtet worden, dass sich die Schulung durch Herrn Wilhelm von der UB Augsburg über gut zwei Tage erstreckte: Der letzte Tag vor und der erste Tag nach Beginn der Katalogisierung, weil sich viele Fragen erst nach Verlassen des „Trockendocks“ ergeben. Dabei war der 12.11.08 der erste Tag der Katalogisierung im BVB. Für die Zeit der Schulung und den Rest dieser Woche ist eine Urlaubssperre für die Betroffenen ausgesprochen worden. Aufgrund der Erweiterung der an Arbeitsplätzen vorfindlichen Software durch den Aleph-Client wurde vorher der Personalrat in Kenntnis gesetzt.

Die Vorbereitung des SIKOM-Einsatzes lief so: Im 1. Schritt wurden die BVB-IDs in die Titelaufnahmen eingetragen. In diesem Zusammenhang wurde die Indexierung des Feldes für die Herkunft des Datensatzes aufgehoben (Recherchen machen nur Sinn, wenn Differenzen vorliegen und in diesem Feld steht bei uns künftig immer dasselbe: „BVB“); das Feld für Fremd-IDs wurde so umgestaltet, dass es noch recherchierbar, aber nicht mehr beschreibbar ist. Letzteres ist wichtig, weil ansonsten Fehler durch die Anwendung von SIKOM entstehen können!

Als 2. Schritt (SIKOM-Einsatz im Testmodus) war angedacht, alle migrierten Sätze insgesamt einzuspielen. Dies erwies sich als nicht durchführbar (Defizite der Datenbank, Zeitbedarf etc.). Es wird daher vorerst darauf verzichtet, zumal in vielen Fällen aus inhaltlicher Sicht auch kein Handlungsbedarf besteht: Wir würden mit großem Aufwand das in unseren Katalog holen, was derzeit ohnehin dort steht.

Im 3. Schritt wurde stattdessen der SIKOM-Echtbetrieb vorgezogen, um die Neuaufnahmen, die nach Ende der alten Produktionsumgebung erstellt wurden, in den Katalog zu bekommen. (Die Krücke „Schnelle Übernahme“ ist insbesondere im Bereich der Normdaten defizitär.)

Zudem gibt es noch (aber das wird evtl. Monate dauern) Gesprächbedarf bzgl. der Details des Geschäftsgangs. Darüber hinaus gibt es für Benutzer und Mitarbeiter Änderungen der Recherchestrategien, u.a., weil dublette ÜGs existieren und es nun wieder die zuvor aufgelösten Abkürzungen im Zusatz zum HST gibt. Dubletten finden sich leider nicht nur bei ÜGs, sondern in großer Anzahl auch bei sonstigen Titelaufnahmen und Einträgen in die Personennormdatei. Wir fürchten, dass sich dieses Problem mit der Migration weiterer Bibliotheken in den BVB verschärfen wird. Die im BVB angewendete Verbundsoftware (Aleph) bietet gewisse Vorzüge. Aber u.a. aufgrund der Dubletten wird die Phase der Routinisierung von Arbeitsprozessen noch eine Weile in Anspruch nehmen.

Die EUV hofft (trotz der Warnung, dass die Lokalsysteme Obacht geben sollen, dass deren lokale Katalog-Besonderheiten berücksichtigt werden), an der Entwicklung eines Migrationsverfahrens in einer Weise mitgewirkt zu haben, die eine problemlose Nachnutzung für andere Bibliothekskataloge ermöglicht.