

Heidrun Wiesenmüller: **Die Konversion der Landesbibliographie-Bände 1973 bis 1985**

Parallel zur Bereitstellung der Landesbibliographie-Datenbank im Internet Ende des Jahres 1999 war die Entscheidung gefallen, die gedruckte Ausgabe mit dem für 2004 geplanten Band 21 (Berichtsjahr 2000) auslaufen zu lassen. Dabei dachte die Kommission für geschichtliche Landeskunde in Baden-Württemberg bereits einen Schritt weiter: Die Gelder, die man ab 2005 bei den Druckkosten einsparen würde, sollten weiterhin der Landesbibliographie zugute kommen. Sie sollten in die Konversion der ersten sechs Bände der gemeinsamen Landesbibliographie für Baden-Württemberg (Berichtszeitraum 1973 bis 1985) fließen, welche bislang nur gedruckt konsultiert werden konnten. So stand also schon seit mehreren Jahren fest, dass vom Jahr 2005 an eine neue, anspruchsvolle Aufgabe auf die Dienststelle Landesbibliographie zukommen würde.

Strategische Planungen

Um einen nahtlosen Übergang von der Druck- in die Konversionsära zu ermöglichen, musste im Laufe des Jahres 2004 parallel zu den Redaktionsarbeiten für den letzten Druckband – die wegen des zusätzlichen Aufsatzteils besonders arbeitsintensiv waren – auch das Konversionsprojekt konzipiert

werden. Die Vorgabe dabei war klar: Die bisher nur über die gedruckten Bände zugänglichen Titel sollten möglichst nahtlos in den bestehenden Datenpool (ab 1986) eingebracht werden. Die Benutzer sollten künftig auch das ältere Material in gewohnter Weise recherchieren können und dabei im Idealfall gar nicht merken, dass zwei unterschiedliche 'Generationen' von Bibliographiedaten vorliegen. Ein Angebot 'neben' der bestehenden Datenbank in Form einer Image-Digitalisierung – also die Erstellung eines virtuellen Abbildes der Druckbände – schied daher von vorneherein aus.

Relativ rasch wurde auch klar, dass eine maschinelle Umsetzung und Aufbereitung für die Datenbank – wie man sie sich im EDV-Zeitalter unweigerlich erhofft hatte – ein unerfüllbarer Traum bleiben würde. Die eigens zu programmierende Software hätte eine Vielzahl komplexer Funktionen erfüllen müssen. Manches wäre noch einigermaßen leicht zu realisieren gewesen, z. B. jedem Titel seine primäre Systemstelle (die nicht immer unmittelbar vor dem Eintrag steht), etwaige Verweisungen (über den ganzen Band verstreut) sowie Nebenschlagwörter (aus dem Sachregister) zuzuordnen. Aber schon bei der Kategorisierung der Titelauf-

nahmen hätte man große Abstriche machen müssen – und selbst bei einer nur ganz groben Untergliederung der bibliographischen Angaben wären erhebliche manuelle Nacharbeiten zu leisten gewesen. Noch schwieriger wäre es gewesen, die Altdaten in das komplexe Verknüpfungsgefüge der Datenbank einzubeziehen, also beispielsweise Zeitschriftenaufsätze mit der zugehörigen Grundaufnahme oder Personentitel mit dem passenden Eintrag in der Personendatenbank zu verknüpfen – dies war aber natürlich eine ganz zentrale Anforderung. Schließlich müssen die Altdaten in verschiedener Hinsicht an heutige Bibliographie-Standards angepasst werden, um eine konsistente übergreifende Recherche über den gesamten Datenpool zu ermöglichen. Auch dies wäre nur zum geringsten Teil maschinell zu bewältigen gewesen.

Entsprechend ernüchternd gestalteten sich die Vorgespräche mit in Frage kommenden Dienstleistern: Hohen Kosten für die Programmierung wäre ein voraussichtlich nur sehr mäßiges Ergebnis gegenüber gestanden, welches manuelle Nacharbeiten in großem Umfang nötig gemacht hätte. In Absprache mit der Kommission für geschichtliche Landeskunde fiel die Entscheidung deshalb gegen eine solche maschinelle Umsetzung und für eine Konversion der klassischen Art, mit der die Qualität und Benutzbarkeit der entstehenden Daten am besten gewähr-

leistet werden kann. Nichtsdestoweniger erfolgt die Konversion EDV-unterstützt: Die Titel werden auf der Grundlage eines Digitalisats und mit Hilfe verschiedener EDV-Tools eingearbeitet – man könnte das Verfahren deshalb auch als teil-automatisiert bezeichnen.

Rahmenbedingungen des Projekts

Die Konversion der älteren Landesbibliographiebände ist durchaus keine Kleinigkeit, sondern eher ein 'Mammutprojekt': Zu konvertieren sind sechs Bände mit insgesamt 72.884 Einträgen, so dass sich ein Durchschnittswert von 12.147 Einträgen pro Band ergibt. Allerdings schwankt die Titelzahl pro Band sehr stark: Sie liegt zwischen 9.768 im zweiten Band (Berichtsjahre 1975/76) und 18.070 – also fast dem Doppelten! – im sechsten Band (Berichtsjahre 1983/84/85). Da die Konversion sozusagen 'von hinten nach vorne' geht, ist ausgerechnet dieser besonders umfangreiche Band als erstes zu bearbeiten. Mit Fortschreiten der Konversion wird sich die Grenze des von der Datenbank abgedeckten Zeitraums (bisher 1986) schrittweise nach vorne verschieben.¹⁾ Alle sechs Bände wurden im Dezember 2004 von einem Dienstleister eingescannt und mit einer OCR-Software bearbeitet, so dass sie nicht nur als Bilddateien, sondern auch im Textformat vorliegen. Die Zahl der Titeldatensätze, die für das Konversionsmaterial in der Daten-

bank anzulegen sind, lässt sich nicht genau ermitteln, denn sie stimmt nicht exakt mit der Zahl der Einträge in den Druckbänden überein: Einerseits müssen in der Datenbank zusätzliche Grundaufnahmen für Sammelbände und teils auch Zeitschriften angelegt werden, aus denen Aufsätze aufgenommen wurden. Andererseits wiederholen sich Einträge für bestimmte Zeitschriften oder Reihen in mehreren Bänden – in der Datenbank genügt in solchen Fällen eine einzige Aufnahme. Weitere Abweichungen können sich aufgrund unterschiedlicher Auswahlkriterien ergeben.

Mit den vorhandenen Mitteln kann die Kommission für geschichtliche Landeskunde eine 37,5 %-Personalstelle für das Projekt finanzieren. Diese verteilt sich derzeit auf zwei Teilzeitkräfte: In Stuttgart ist seit Februar 2005 Frau Romeis – eine Mitarbeiterin der Kommission – mit 25 % für die Konversion tätig. Sie hat ihren Arbeitsplatz im Zimmer der Landesbibliographie; manchen Kolleginnen und Kollegen wird daher das neue Gesicht bereits begegnet sein. Am Standort Karlsruhe arbeitet seit März 2005 Frau Brawanski – eine der beiden dortigen Stammkräfte – mit 12,5 % an der Konversion. So stehen knapp 14,5 Wochenstunden speziell für die Konversion zur Verfügung. Aber natürlich wird auch vom Stammpersonal während der normalen Arbeitszeit Arbeit für das Projekt geleistet.

Bis einschließlich August 2005 wurden bereits ca. 4.500 Titel konvertiert. Die Gesamtlaufzeit des Projekts lässt sich noch nicht genau kalkulieren²⁾, jedoch ist klar, dass sich die Konversionsarbeit über einige Jahre erstrecken wird.

Das Verfahren

Wie läuft die Konversion nun im Einzelnen ab? Als erstes wurden Arbeitsanweisungen formuliert und Frau Romeis musste in ihre neue Aufgabe eingearbeitet werden. Neben inhaltlichen Grundlagen ging es dabei vor allem darum, die (sehr komplexe) Bedienung des Dienstprogramms und die verschiedenen Typen der anfallenden Titelaufnahmen zu erlernen. Nicht nur diese Einarbeitung, sondern auch die laufende Betreuung sowie die Koordination mit der Karlsruher Bearbeiterin liegen in der Verantwortung von Frau Blaser. Frau Romeis bearbeitet derzeit den (von der Sacherschließung her relativ einfachen) Personenteil: Sie legt die Titelaufnahmen in der Datenbank an und verknüpft diese mit dem zugehörigen Eintrag in der Personendatenbank. Parallel dazu hat Frau Brawanski mit der Bearbeitung des Ortsteils begonnen, wobei sie auch eine Sacherschließung gemäß heutigem Standard vergibt.

In einem zweiten Schritt werden die Konversionsdaten – genau wie die laufenden Bibliographiedaten auch – einer formalen und sachlichen Redaktion durch Frau Blaser bzw. Frau Wiesen-

müller unterzogen. Zusätzlich zu den normalen Tätigkeiten, die bei der Redaktionsarbeit anfallen, müssen bei der sachlichen Redaktion der Personenliteratur z. T. fehlende Personendatenätze neu angelegt und Nebenstellen ergänzt werden. Anders als bei den laufenden Bibliographiedaten, deren Endredaktion erst mit einem gewissen Verzug und in Jahresringen erfolgt, wird die Redaktion des Konversionsmaterials möglichst zeitnah in kleinen Portionen durchgeführt.

Die Projektsteuerung und die Redaktion der Konversionsdaten, die auf Jahre hinaus parallel zur laufenden Bibliographiebearbeitung erfolgen wird, stellt eine beträchtliche Zusatzbelastung für das Stammpersonal an der WLB dar, die nur zu einem kleinen Teil durch Einsparungen aufgrund des Wegfalls der Druckausgabe kompensiert wird.³⁾

Werkzeuge für die Erfassung

Die eigentliche Erfassungsarbeit wird – wie bereits erwähnt – durch verschiedene Hilfsmittel erleichtert. So wurden für die am häufigsten vorkommenden Formalgruppen (Aufsätze aus regelmäßig ausgewerteten Zeitschriften, Aufsätze aus Sammelbänden sowie einbändige Monographien) maßgeschneiderte Scripte für die Terminalemulation geschrieben. Diese füllen bestimmte Felder der Erfassungsmaske automatisch aus und geben die zu besetzenden Kategorien vor. Um regelmäßig benötigte Informationen nicht jedes Mal

wieder von Neuem im System recherchieren zu müssen, werden außerdem verschiedene Word-Tabellen geführt. Beispielsweise wurde eine Liste der regelmäßig ausgewerteten Zeitschriften angelegt, in die Identnummern und Signaturen (von WLB und BLB) eingetragen werden. Bei der Erfassung von Aufsätzen daraus können die Angaben komfortabel per Copy-and-Paste in die Erfassungsmaske übertragen werden.

Einen Rationalisierungseffekt erhofften wir uns natürlich auch vom Digitalisat des Druckbandes: Um sich das Abtippen zu ersparen, sollten daraus der/die Verfasser und der Sachtitelbereich kopiert werden. Diesem Wunsch standen jedoch zunächst die Konventionen der Erfassungsmaske entgegen: Nicht nur müssen dort die Kategorien getrennt eingegeben, sondern auch längere Textinhalte in mehrere Zeilen mit jeweils maximal 38 Zeichen zerlegt werden. Am Beginn jeder Zeile ist das entsprechende Kategorienkürzel, ein Kategorien- sowie ein Zeilenzähler anzugeben. Direktes Copy-and-Paste vom Digitalisat in die Erfassungsmaske ist daher nicht möglich. Doch auch für dieses Problem wurde eine technische Lösung gefunden: Der Autor-Titel-String wird in einem ersten Schritt in ein eigens dafür programmiertes Excel-Tool⁴⁾ kopiert. Dieses bereitet den Text so auf, dass ein direktes Einfügen in die Erfassungsmaske möglich ist.

Z. B. wird der folgende, aus der Volltextdatei herauskopierte Textstring:

Gubler, Hans Martin: Johann Caspar Bagnato (1696-1757) und das Bauwesen des Deutschen Ordens in der Ballei Elsaß-Burgund im 18. Jahrhundert. Ein Barockarchitekt im Spannungsfeld v. Auftraggeber, Bauorganisation u. künstlerischem Anspruch.

im Excel-Tool automatisiert umgewan-

delt in die folgende, nach Kategorien strukturierte Darstellung ('v0' steht für den Verfasser, 'st' für den Hauptsachtitel und 'ut' für den Zusatz zum Sachtitel), wobei die Zeilen zugleich die richtige Länge erhalten und die nötigen Angaben am Zeilenanfang ergänzt werden:

v0	01	01	Gubler, Hans Martin
st	01	01	Johann Caspar Bagnato (1696-1757) und
st	01	02	das Bauwesen des Deutschen Ordens in
st	01	03	der Ballei Elsaß-Burgund im 18.
st	01	04	Jahrhundert
ut	01	01	Ein Barockarchitekt im Spannungsfeld
ut	01	02	v. Auftraggeber, Bauorganisation u.
ut	01	03	künstlerischem Anspruch

Dies kann nun en bloc in die Zwischenablage kopiert und daraus wiederum in die Erfassungsmaske der Datenbank eingefügt werden (Abb. 1).

Auch mehrere Verfasser, die in den Druckbänden durch Semikolon ge-

trennt sind, werden vom Konversionswerkzeug erkannt und richtig umgesetzt. Hauptsachtitel und Zusatz zum Sachtitel werden in den Druckbänden nicht durch einen Doppelpunkt, sondern nur durch einen Punkt voneinan-

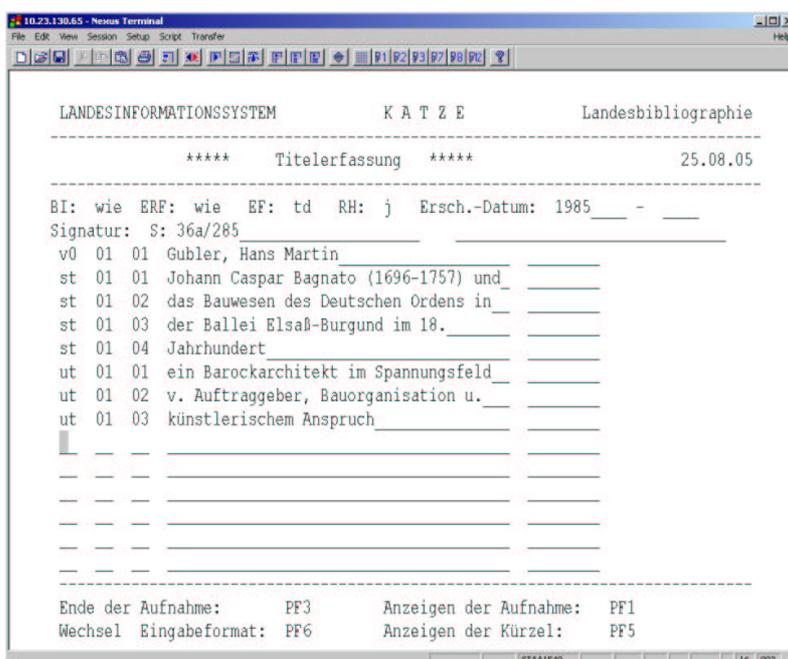


Abb. 1: In die Erfassungsmaske eingefügter Block aus dem Excel-Tool

der abgesetzt. Das Programm sucht daher standardmäßig nach dem ersten Punkt innerhalb des Textstrings. Den davor stehenden Teil betrachtet es als Hauptsachtitel, den Rest als Zusatz. Punkte können aber natürlich auch in anderer Funktion innerhalb des Titelbereichs vorkommen, wie etwa im obigen Beispiel („18. Jahrhundert“). In diesem Fall ist es also nicht der erste, sondern erst der zweite Punkt, der Hauptsachtitel und Zusatz voneinander trennt. Dies kann freilich das Excel-Tool nicht selbsttätig erkennen. Die Erfasserin muss daher in diesem Fall zunächst den Defaultwert ‘1’ im Feld ‘Trennpunkt’ des Excel-Tools auf ‘2’ umstellen, damit die beiden Kategorien richtig umgesetzt werden. Zugleich ist dies ein gutes Beispiel für die Probleme, die bei einer rein maschinellen Kategorisierung aufgetreten wären. Die Interpunktionszeichen, die im Druckband am Ende der Kategorien standen, werden vom Konversionswerkzeug eliminiert. Stattdessen werden für die Anzeige in der Online-Datenbank die gemäß ISBD üblichen Interpunktionszeichen maschinell vom Dienstprogramm ergänzt.⁵⁾

Das Digitalisat ist mitunter auch bei der Sacherschließung nützlich, weil man mit einer Suche nach der laufenden Nummer des Eintrags zugehörige Verweisungen finden kann. Die Suche nach Schlagwörtern aus dem Register ist ebenfalls möglich, macht in der Praxis aber wenig Sinn, da zum einen in Band 6 nur wenige Nebenschlagwörter

vergeben wurden (das Sachregister für über 18.000 Titel ist nur 30 Seiten stark!) und zum anderen ohnehin das heutige Normvokabular verwendet werden muss. Deshalb geht es im Normalfall schneller, die Schlagwörter gemäß aktuellem Standard neu zu vergeben.

Bearbeitung der Titel

Zwischen der damaligen und der heutigen Praxis der Landesbibliographie gibt es beträchtliche Abweichungen – bei der Formalerschließung (erst seit dem Wechsel auf die Datenbank kommen die RAK zur Anwendung), bei der Sacherschließung (unterschiedliche Systematiken, Schlagwörter heute nach SWD) und schließlich auch bei den Kriterien für die Auswahl des Materials. Die Einträge der alten Druckbände können daher nicht einfach 1:1 umgesetzt werden, sondern müssen – mit entsprechendem intellektuellen Aufwand – nach den heutigen Grundsätzen überarbeitet werden. Dabei wird natürlich keine 100 %-ige Angleichung angestrebt. Statt Perfektionismus ist vielmehr Pragmatismus angesagt, um Kosten und Nutzen in ein sinnvolles Verhältnis zu bringen und zu gewährleisten, dass die Konversion in gutem Tempo voranschreiten kann.

Das online recherchierbare Konversionsmaterial stellt daher auch kein exaktes Abbild der Einträge aus den Druckbänden dar. Zumeist enthalten die entstehenden Datensätze mehr Informationen, als in den Bänden gebo-

ten wurden (z. B. genauere bibliographische Angaben, mehr Schlagwörter, feinere Systemstellen). Mitunter müssen aber auch Verluste hingenommen werden: Beispielsweise wurden bei Ortsgeschichten früher die enthaltenen Aufsätze zwar nicht speziell erschlossen, aber zumindest im Petitdruck beim Haupteintrag mit angegeben. Dieses Verfahren lässt sich nicht auf die Datenbank übertragen. Will man die Information erhalten, so müsste man entsprechend viele vollwertige Aufsatzaufnahmen an die Hauptaufnahme anhängen. Dies widerspricht jedoch der heutigen Praxis, nach der Ortsgeschichten nur eine Hauptaufnahme erhalten.⁶⁾ Deshalb fallen die Petitdruck-Angaben zu Ortsgeschichten bei der Konversion 'unter den Tisch'.

Im Einzelnen werden bei der Bearbeitung des Materials für die Datenbank folgende Änderungen vorgenommen:

- Bei Monographien sind die bibliographischen Angaben in Band 6 häufig sehr stark verkürzt, z. B. fehlen Reihenangaben und Urheber. Die Titelaufnahmen werden daher mit den Aufnahmen im Südwestverbund (SWB) abgeglichen und entsprechend 'aufgefüttert'. Aufsatzaufnahmen bleiben hingegen praktisch unverändert.
- Abkürzungen im Zusatz zum Sachtitel werden aufgelöst, wenn sie recherchierelevant sind (z. B. „B-W“ wird zu „Baden-

Württemberg“). Nicht recherchierelevante Abkürzungen (z.B. „e.“ für „einer“, „eine“ etc.) können stehen bleiben, was das Copy-and-Paste-Verfahren erleichtert.

- Signaturen der WLB und BLB werden ergänzt, um auch für die Altdaten den direkten Wechsel in die Ausleihsysteme zu ermöglichen. Bei Monographien geschieht dies in der Regel, wenn eine SWB-Aufnahme vorhanden ist. Systematische Recherchen im Digikat der WLB oder dem Mikrofiche-Katalog der BLB würden zuviel Zeit in Anspruch nehmen.
- Stellen der alten Systematik müssen auf die entsprechenden Stellen der jetzigen Systematik abgeändert werden. Im Ortsteil geht dies mit einer Verfeinerung der sachlichen Erschließung einher, da hier in den gedruckten Bänden meist nur die oberste Hierarchiestufe (bzw. bei größeren Orten die erste und zweite Hierarchiestufe) angegeben wurde.
- Haupt- und Nebenschlagwörter nach heutigem Standard werden zugeordnet bzw. ergänzt, wobei zumeist eine Verdichtung der Sachinformation erfolgt.

Zu Schwierigkeiten führen gelegentlich die unterschiedlichen Aufnahmeprinzi-

pien der damaligen und heutigen Zeit. So wurden früher – zumindest in Auswahl – auch Artikel aus den großen Stuttgarter und Karlsruher Zeitungen sowie Aufsätze aus Adressbüchern aufgenommen; beides wird heute nicht mehr praktiziert. Hier ist abzuwägen, ob der Wert dieser Titel so hoch ist, dass man eine Inkonsistenz in der Datenbank – die auch zu falschen Erwartungen bei den Benutzern führen kann – in Kauf nehmen will, oder ob man sie lieber weglässt. Derzeit werden solche Aufnahmen nicht konvertiert, jedoch im Arbeitsexemplar gekennzeichnet.

Auch an die baden-württembergischen Persönlichkeiten werden heute strengere Kriterien angelegt. So wurden früher nicht nur diejenigen Personen als Landeskinder gezählt, die in Baden-Württemberg geboren und/oder gestorben sind bzw. mindestens ca. zehn Jahre hier gelebt und gewirkt haben, sondern auch solche, die „durch vorübergehende Tätigkeiten dort bleibende Spuren hinterlassen haben“.⁷⁾ So tauchen etwa im Personenteil von Band 6 zwei Titel zu Goethe auf. In der heutigen Personendatei gibt es für Goethe jedoch keinen Eintrag mehr, mit dem man sie verknüpfen könnte – und nach den heutigen Regeln darf ein solcher auch nicht mehr angelegt werden. Dieses Problem lässt sich durch Anwendung der heutigen Praxis noch relativ leicht lösen: Titel vom Typ 'Goethe in Heidelberg' findet man unter dem Ort

(mit der Systemstelle 0.8: Biographisch-geographische Beziehungen); die Person Goethe wird als Schlagwort erfasst. In einigen Fällen bleibt jedoch nichts anderes übrig, als die zugehörigen Titel bei der Konversion auszulassen.

Noch nicht endgültig entschieden ist das Verfahren bei Zeitschriften wie dem 'Bloch-Almanach', die sich spezifisch mit einer Person – hier dem Philosophen Ernst Bloch – beschäftigen. Diese werden heute aus Gründen des Aufwands nur noch mit einer Grundaufnahme geführt, ohne dass die enthaltenen Aufsätze einzeln aufgenommen werden – dafür kann der Benutzer schließlich entsprechende Personal- oder Spezialbibliographien konsultieren. Allein in Band 6 sind hingegen nicht weniger als 29 Aufsätze aus dem Bloch-Almanach enthalten. Auch in diesen Fällen liegt es nahe, auf eine Konversion zu verzichten.

Im Herbst wollen sich die Landesbibliographie-Mitarbeiterinnen und -Mitarbeiter aus Karlsruhe und Stuttgart zusammensetzen, um ihre Erfahrungen auszutauschen und Festlegungen für den weiteren Verlauf der Konversionsarbeit zu treffen. Schon jetzt lässt sich freilich sagen, dass das Projekt zwar – wie jede Konversion klassischen Typs – nicht unbeträchtliche Mühe macht, aber den Nutzern der Landesbibliographie auch eine gewaltige Verbesserung bringen wird.

-
- 1) Nach abgeschlossener Bearbeitung von Bd. 6 auf 1983, nach Bd. 5 auf 1981, nach Bd. 4 auf 1979, nach Bd. 3 auf 1977, nach Bd. 2 auf 1976 und schließlich auf 1973. Da inzwischen schon eine beträchtliche Zahl von Titeln vor 1986 in die Datenbank eingebracht wurde, wird in Kürze ein entsprechender Hinweis auf der Homepage der Landesbibliographie angebracht.
 - 2) Die vorliegenden Zahlen können nur bedingt hochgerechnet werden. Nicht nur ist der Aufwand für die Konversion der drei Teile (Allgemeiner, Orts- und Personenteil) unterschiedlich hoch, sondern es sind auch Unterschiede zwischen den Bänden zu erwarten, z. B. weil aus den 1970er-Jahren nicht so viele Titel im SWB zu finden sind wie aus den 1980er-Jahren. Auch die Netto-Zahl der insgesamt zu erstellenden Datensätze kann erst auf einer größeren Zahlenbasis kalkuliert werden.
 - 3) Entlastung gibt es hier primär für Frau Wiesenmüller, da die Arbeiten für die Druckvorbereitung (z. B. Fahnenlesen) wegfallen.
 - 4) Die Programmierung erfolgte durch Frau Wiesenmüllers Ehemann, Bernd Raab.
 - 5) Am Anfang des Zusatzes zum Sachtitels muss gemäß RAK u. U. noch manuell von Groß- auf Kleinschreibung geändert werden.
 - 6) Man hat sich auf diese Praxis geeinigt, da zum einen die Aufnahme der einzelnen Aufsätze mit den vorhandenen Ressourcen nicht zu leisten wäre und zum anderen davon ausgegangen werden kann, dass die Benutzer eine ungefähre Vorstellung davon haben, was in derartigen Publikationen zu finden ist.
 - 7) Einleitung zu Bd. 1, S. XXI.