

Das Lexikon des Mittelalters auf CD-ROM: verlegerische und technische Aspekte der Umsetzung von Buchinhalten in elektronische Form

Lutz, Bernd; Meier, Thomas; Weerth, Thomas

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Lutz, B., Meier, T., & Weerth, T. (2002). Das Lexikon des Mittelalters auf CD-ROM: verlegerische und technische Aspekte der Umsetzung von Buchinhalten in elektronische Form. *Historical Social Research*, 27(2/3), 274-284. <https://doi.org/10.12759/hsr.27.2002.2/3.274-284>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:

<https://creativecommons.org/licenses/by/4.0/deed.de>

Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see:

<https://creativecommons.org/licenses/by/4.0>

Das Lexikon des Mittelalters auf CD-ROM: Verlegerische und technische Aspekte der Umsetzung von Buchinhalten in elektronische Form^{*}

Bernd Lutz, Thomas Meier, Thomas Weerth

I. Das Lexikon des Mittelalters: Das Buch, die Datenbank und die CD-ROM

1.

Eine gängige Form, Informationen rasch auffindbar und in kompakter Form zur Verfügung zu stellen, ist seit jeher das Lexikon. Insofern ist in Gestalt des Lexikons ein gutes Stück dessen schon lange realisiert, wovon uns die Propagandisten der „Datenautobahn“ vorschwärmen. In diesem Zusammenhang wird ja ohnehin meist übergangen, dass Information nicht einfach am Wegrand liegt, sondern im Gegenteil erst mühsam erarbeitet und aufbereitet werden muss.

Davon kann nun auch ein Lied singen, wer am Lexikon des Mittelalters, einem in mehrerer Hinsicht einmaligen Lexikonprojekt, mitgearbeitet hat. Die Geschichte dieses historischen Nachschlagewerks, die weit ins letzte, also ins 20. Jahrhundert zurückreicht, ist zweifelsohne ein Stück Wissenschafts- wie Verlagsgeschichte, sie war aber auch schlicht ein Stück Kärnerarbeit.

Die Geschichte des LexMA – so heißt das Werk abgekürzt – reicht zurück bis in die 1950er Jahre. Es waren v.a. einzelne Forscher und Gelehrte, die sich ein umfassendes Nachschlagewerk zu jener Epoche wünschten, die gemeinhin als Mittelalter bezeichnet wird. Allein, es fand sich damals kein Verlag, der ein solch ehrgeiziges Vorhaben gewagt hätte.

Die Situation sollte sich erst 1968 ändern, als der Stuttgarter Druckemüller Verlag vom Artemis Verlag übernommen wurde. Mit dieser Akquisition hatte der Zürcher Verlag auch eine Perle ganz besonderer Art erworben. Die Rede ist

^{*} Aus dem Protokoll des 81. Kolloquiums über die Anwendung der Elektronischen Datenverarbeitung in den Geisteswissenschaften an der Universität Tübingen vom 10. Februar 2001.

vom monumentalen, über 80 Teilbände umfassenden, seit 1893 erscheinenden und damals kurz vor dem Abschluss stehenden Standardnachschlagewerk zur antiken Welt: Pauly's Realencyclopädie der Classischen Altertumswissenschaft. Als Verlag für große Nachschlagewerke hatte sich Artemis bereits drei Jahre zuvor einen Namen gemacht. Mit eigenen Mitteln war das erfolgreiche und heute immer noch sehr beliebte Lexikon der Alten Welt produziert worden.

In diesem neuen verlegerischen Umfeld fiel die alte Idee eines Lexikons, das den breiten Bereich des Mittelalters abdecken sollte, plötzlich auf einen fruchtbaren Boden, und bereits 1969 wurde offiziell eine Redaktion am Stuttgarter Verlagssitz eingerichtet. Das Werk sollte nicht 80, wie der Pauly, sondern lediglich 3, allenfalls 4 Bände umfassen, und der Verlag rechnete mit einer Produktionszeit von 10 Jahren für das geplante Lexikon des Mittelalters.

Es sollte anders kommen: Das Erscheinen des ersten Faszikels verzögerte sich, und erst im Oktober 1977, also 9 Jahre nach dem Startschuss und nur 1 Jahr vor dem ursprünglich geplanten Abschluss des Gesamtwerkes, lag die erste Teillieferung vor. Sie umfasste auf 224 Spalten die ersten Artikel von „Aachen“ bis „Ägypten“.

Schuld an der Verzögerung war, dass sich die Planungsarbeiten als viel aufwändiger als vorgesehen gestalteten, denn zum Mittelalter gab es überhaupt keine Vorarbeiten lexikografischer Art. Mittlerweile war in München, wohin der Verlag sein Domizil verlegt hatte, eine eigentliche Fachredaktion mit 4 Redaktorenstellen und einer Sekretärin aufgebaut worden. Ebenso war ein internationaler Berater- bzw. Herausgeberstab verpflichtet worden, mit dem es die Gesamtnomenklatur zu komplettieren und bereinigen galt. Dies alles bewirkte fast naturwüchsig eine Verbreiterung der Nomenklatur, oder besser: Nomenklaturen, und führte eben auch zum erwähnten Rückstand auf den ursprünglichen Terminplan.

Zu den Bereichen, die im LexMA abgedeckt werden sollten und auch wurden, gehören einmal die allgemeine politische Geschichte aller Länder und Regionen Europas, einschließlich des Byzantinischen Reichs, die Rechts-, sodann die Verfassungs-, Sozial- und Wirtschaftsgeschichte, die Stadt- und Agrargeschichte, die Archäologie und Siedlungsgeschichte, die historischen Grundwissenschaften, die Kirchengeschichte, die Philosophie und Theologie, die Geschichte der mittelalterlichen Literaturen, die Kunstgeschichte mit ihren Zweigen, die Kulturgeschichte, einschließlich der Darstellung des Alltagslebens, sowie die Geschichte von Medizin, Naturwissenschaften, Mathematik und Technik. Auch die Bedeutung des Islam und des Judentums für die europäische Geschichte wird eingehend gewürdigt.

Gleich vom ersten Faszikel weg wurde das LexMA von der einschlägigen Fachwelt insgesamt sehr gut aufgenommen, und dem positiven Echo entsprachen vergleichsweise hohe Absatzzahlen.

2.

Soweit die Erfolgsmeldungen. Probleme ergaben sich – wie schon erwähnt – gleichwohl!

Nachdem 1980 der erste und 1983 der zweite Band erschienen waren, aber noch nicht einmal der Buchstabe „C“ hatte abgeschlossen werden können, musste allen Beteiligten klar sein, dass ein Gesamtumfang auch von 5 Bänden illusorisch sein würde. Mit dem 1986 erschienenen dritten Band schließlich konnte nicht einmal der fünfte Buchstabe „E“ des Alphabets beendet werden. Die Umfänge drohten immer mehr auszufern und damit das ganze Projekt aus dem Ruder zu laufen genau in jene Richtung, die man eigentlich immer hatte vermeiden wollen. Das Werk war für den Verlag längst zu einem finanziellen Desaster geworden und konnte nur dank massiven Zuschüssen aus der Privatschatulle des Verlagsbesitzers über Wasser gehalten werden.

In dieser Situation versuchte der Verlag Gegensteuer zu geben und beauftragte zwei Außenstehende mit einem Gutachten – der eine ein Informatiker mit langjähriger Verlagserfahrung, der andere der Verfasser. Die Analyse war rasch gemacht: Ganz offensichtlich mangelte es an den geeigneten Planungs- und Kontrollinstrumenten. Als Folge davon fehlten etwa klare Umfangkontrollen, und ebenso unvermeidlich waren Pannen, die sich auf die Umfänge einzelner Stichwörter wie auf den Erscheinungsrhythmus auswirkten. Hinzu kam, dass damals niemand auch nur annäherungsweise sagen konnte, wie viele Stichwörter mit wie vielen Zeilen noch zu erwarten waren. In der Redaktion war der Überblick über das in Zettelkästen verwaltete Stichwortgut längst verlorengegangen.

Um ein zuverlässiges Planungsinstrument zu bekommen, schlugen wir vor, die Gesamtnomenklatur in einer Datenbank zu erfassen. Zweitens sollte stärker als bisher auf die Einhaltung der Umfangvorgaben geachtet sowie der Erscheinungsrhythmus erhöht werden, damit das Werk innerhalb einer vertretbaren Zeitspanne erscheinen konnte.

Der Verlag schloss sich diesen Überlegungen an, und explizit zu deren Umsetzung wurde im Herbst 1987 in Zürich eine Arbeitsstelle LexMA eingerichtet, die von Beginn weg von Frau Dr. Charlotte Bretscher und dem Verfasser betreut wurde.

Zwar befassten wir uns in der Zürcher Arbeitsstelle stärker als ursprünglich beabsichtigt auch mit allgemeinen Redaktionsarbeiten, die vordringlichste Aufgabe aber blieb zunächst die Erstellung verschiedener Datenbanken, um ein Planungs- und Kontrollinstrument zu erhalten.

Nach der Adressdatenbank (mit den 92 Herausgebern und mittlerweile gegen 3000 Autorinnen und Autoren mit oft zwei Anschriften) wurde ein Literatur- und Abkürzungsverzeichnis erstellt. Sodann wurde der größte Brocken in Angriff genommen, nämlich die sich in einem eher schlechten Zustand befindende Gesamtnomenklatur. Sie wurde ab dem Buchstaben „G“ – so weit war

das Werk damals gediehen – samt allen verfügbaren Informationen ebenfalls in eine Datenbank überführt.

Die betreffenden Arbeiten waren im Spätsommer 1988 abgeschlossen, wobei allerdings anzumerken ist, dass die Stichwörterlisten permanent überarbeitet wurden. Trotzdem: Wir wussten damals nun wenigstens, dass noch mit rund 20.000 Stichwörtern zu rechnen war, und es konnte eine einigermaßen zuverlässige Hochrechnung und Umfangplanung gemacht werden. Sie fiel - nicht ganz unerwartet - ernüchternd aus. Demnach war auch nach der vorgängigen Abspeckung von zeilenmäßig überdotierten Stichwörtern zwar nicht mit den von vielen befürchteten 10-12, aber doch mit deutlich mehr als 8 Bänden zu rechnen, und dabei waren die vorgesehenen verschiedenen Anhänge noch nicht inbegriffen. Wenn es im bisherigen Tempo weitergegangen wäre, würden wir heute immer noch auf den Abschluss des Werks warten.

Die Hauptstoßrichtung war demnach klar: Am Stichwörterkorpus selber wollten und konnten wir z.T. gar keine Abstriche machen. Zum einen galt es, eine inhaltliche Ausdünnung im Interesse des Gesamtprojekts zu verhindern, zum andern mussten viele Stichwörter allein deshalb beibehalten werden, weil auf sie bereits ein Verweis vorhanden war. Die einzigen Auswege bestanden also in einer Beschleunigung des Erscheinungsrhythmus einerseits, in der strikten Einhaltung und – wo möglich – Reduzierung der Artikelumfänge andererseits.

Dies sollte erreicht werden mit einer Rationalisierung der Arbeitsabläufe ebenfalls vor allem mittels EDV:

Nachdem in der Zürcher Arbeitsstelle schon im Herbst 1987 zwei Arbeitsplatzstationen des Typs Apple Macintosh installiert worden waren, wurde die Münchner Redaktion ab 1988 ebenfalls mit Geräten ausgerüstet. Die Selbstverständlichkeit, mit der wir heute Computer am Arbeitsplatz oder zu Hause einsetzen, lässt meist vergessen, dass dies noch Ende der 1980er Jahre eher die Ausnahme war.

Endlich war es möglich, einen Überblick zuerst zu gewinnen und – was noch viel wichtiger war! – danach auch zu bewahren. Aber auch die anderen Ziele – Einhaltung der Umfänge und Beschleunigung des Erscheinungsrhythmus – wurden erreicht: Die letzten 5 Bände – jeder Band umfasst 2.200 Spalten mit weit über 8 Mio. Zeichen – erschienen im Zweijahrestakt, und rechtzeitig zur Buchmesse 1998 lag der letzte Halbband vor, dem noch ein Registerband folgte.

Selbstverständlich wurden die Computer auch in der täglichen Redaktionsarbeit an den Texten eingesetzt, und die Verfügbarkeit der Texte in elektronischer Form verringerte auch die Satz- und Korrekturkosten, die gerade bei einem solchen Werk mit seiner spezifischen Fachterminologie und komplexen Typografie sehr erheblich sind.

Die Kommunikation zwischen den beiden Redaktionsorten war – wenigstens technisch – kein Problem, konnten doch die Artikel via E-Mail bequem

zwischen Zürich und München hin- und hergeschickt werden. Bezüglich E-Mail waren wir übrigens seit 1988, also schon in der diesbezüglichen Pionierzeit, mit von der Partie.

3.

Auch wenn seit der Umstellung auf EDV 1987 stets das Ziel vor Augen stand, das LexMA in elektronischer Form verfügbar zu machen, so blieb unsere Datenbank primär ein redaktionelles Arbeitsinstrument. Nebenbei produzierten wir mit Hilfe der Datenbank auch noch die Honorarabrechnungen, wofür im Verlag vorher zwei Personen über mehrere Monate beschäftigt gewesen waren.

Der Entscheid, 1987 auf Apple Macintosh und nicht auf PCs zu setzen, war ein nicht ganz zufälliger. Damals gab es noch sehr wenige Datenbank-Produkte für Kleinrechner auf dem Markt, und alle hätten unsere Bedürfnisse kaum abgedeckt bzw. unsere Datenmengen schwerlich bewältigen können. Die Applikation, die wir schließlich wählten, trägt die futuristische Bezeichnung „4th Dimension“.

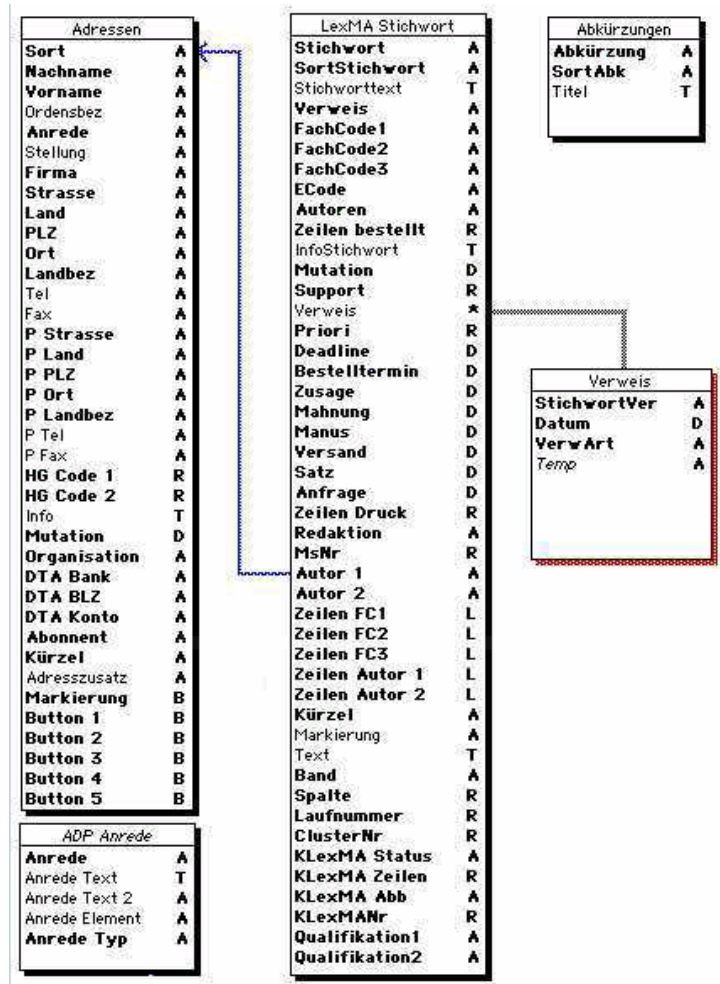
Dabei handelt es sich um eine relationale Datenbank, die längst auch auf PCs läuft. Sie weist Funktionalitäten auf, wie wir sie in der täglichen Arbeit benötigten, darunter etwa komplexe Suchmöglichkeiten, Verlinkung von Feldern unterschiedlicher Datenbanken, Einrichtung von Subdatenbanken, beliebige Bildschirm- und Druck-Layouts, Datenexport usw. Die Applikation ist handlich und kann jederzeit fast im Handumdrehen auch von Nicht-Fachleuten an neue Bedürfnisse angepasst werden. Das wichtigste Plus ist aber darin zu sehen, dass sie auch auf den neuesten Geräten immer noch klaglos läuft, wir also keine zeit- und kostenintensiven Datenmigrationen vornehmen mussten. Ein einziges Mal – 1994 – implementierten wir eine verbesserte Version der Software, die dann auch Serverbetrieb erlaubte, Farbe usw. unterstützte und prinzipiell bereits fit fürs Internet war. Figur 1 zeigt die Struktur der LexMA-Datenbank mit den Einzeldatenbanken sowie einer Subdatenbank.

Herausgreifen möchte ich nur kurz die Stichwortdatenbank. Auf eine Erklärung der einzelnen Datenfelder kann verzichtet werden, die Bezeichnungen sind meist sprechend. Hinweisen möchte ich einmal auf die drei mit „FachCode“ bezeichneten Felder mit Auswahllisten. Die vorgehaltenen rund 100 sogenannten FachCodes bezeichneten ursprünglich nicht systematische Fachgebiete, sondern vielmehr die für das betreffende Stichwort zuständigen Fachberater und Herausgeber, d.h. diese Codes deckten ganz unterschiedlich breite, teils auch mehrere Gebiete ab (z.B. Iberische Halbinsel, Zoologie). Auch wenn sie für die inhaltliche Zuordnung der einzelnen Lemmata auf der CD-ROM überarbeitet werden mussten, war die damit gegebene inhaltlich-thematische Zuordnung des Stichwortguts natürlich dennoch von unschätzbarem Wert.

Die Felder „Band“, „Spalte“ und „Laufnummer“ wurden erst im Hinblick auf die CD-ROM kreiert und nachgetragen, ebenso die beiden Felder „Qualifi-

kation“, in denen im nachhinein verzeichnet wurde, ob es sich um einen Personenartikel oder aber um eine Stadt oder ein Kloster handelt.

Fig. 1 : Struktur der LexMA-Datenbank

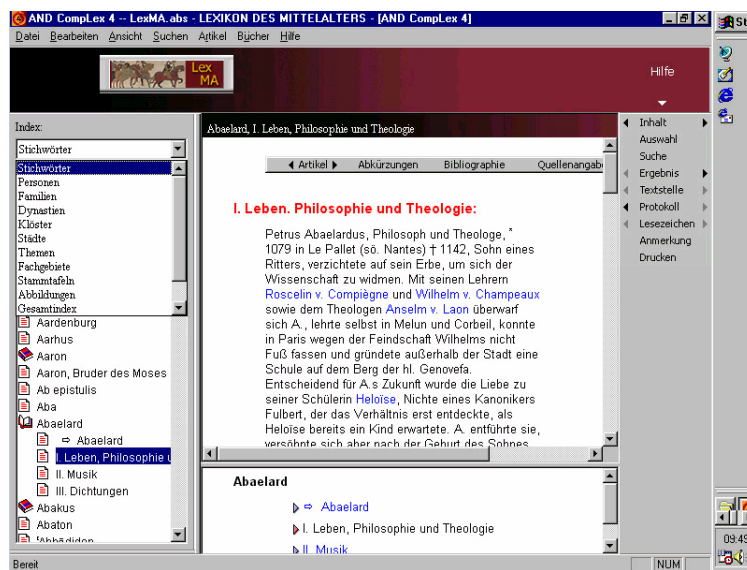


4.

Die CD-ROM umfasst einmal sämtliche Stichwörter und Texte wie in der Buchausgabe und übernimmt dabei auch die vorhandene Gliederung. Im Vergleich zur Buchausgabe weist sie nun aber ganz wesentliche zusätzliche Funktionalitäten auf, die gerade im wissenschaftlichen Alltag äußerst nützlich sind.

Dazu gehören einmal verschiedene Indizes, die aus dem riesigen Stichwörterkorpus vielfältige Auswahlen nach fachlichen, inhaltlichen und thematischen Kriterien (z.B. Personen, Städte, Fachgebiete, Themen usw.) treffen (vgl. Fig. 2). Dies war nur möglich, weil auf die entsprechenden Informationen aus der LexMA-Datenbank zurückgegriffen werden konnte.

Fig. 2 : Maske der LexMA-CD-ROM



Volltextsuchen sind natürlich ebenso möglich wie andere einfache und komplexe Suchläufe. Über die als Verweise markierten Wörter innerhalb eines Artikeltexts kann direkt auf das betreffende Stichwort gesprungen werden, dies ebenfalls ein wesentlicher Vorteil gegenüber der Buch-Version. Die einzelnen Arbeitsgänge werden zudem protokolliert, so dass der Überblick über die Sitzung jederzeit gewahrt bleibt. Es besteht ferner die Möglichkeit, Lesezeichen zu setzen und eigene Bemerkungen an Artikel anzuhängen sowie Texte oder Listen auszudrucken oder – und das macht die CD-ROM erst recht zum attrak-

tiven Arbeitsinstrument von Mediävistinnen und Mediävisten – in eine Textverarbeitungs-Umgebung zu exportieren.

Thomas Meier (Zürich)

II. Über die technischen Möglichkeiten bei der Digitalisierung von Verlagsinhalten

Die Erkenntnis, dass die Anlegung bzw. die Umsetzung von Verlagsinhalten in SGML- oder XML-basierte Datenbestände zwar oft mit etlichen Anfangsschwierigkeiten zu kämpfen hat, mittelfristig aber eine signifikante Senkung der Kosten in nahezu allen Produktionsbereichen bewirkt, beginnt sich auch in den Verlagshäusern des klassischen Wissenschafts- und Belletristikbereiches durchzusetzen. Die Vorteile eines neutralen SGML/XML-Datenbestandes liegen auf der Hand: Neben einem systemneutralen Datenzugriff besteht die Möglichkeit redaktioneller Arbeit an den Daten, die nicht mehr in die Ratio anderer Produktionsprozesse eingreift; und nicht zuletzt bieten SGML- oder XML-Daten den Vorteil, jederzeit für die Publikation verfügbar zu sein, ohne dass dafür die Inhalte jeweils neu erschlossen werden müssten.

Beispielhaft für die Primäerschließung von Verlagsinhalten, ihre Umsetzung in XML und anschließende Publikation ist das Projekt „Das Lexikon des Mittelalters (LexMA) auf CD-ROM“. TUSTEP gewährleistete dabei, dass die Daten von Anfang an in SGML-basierten Systemen vorgehalten wurden; so kamen die oben beschriebenen Vorteile schon für die allerersten Arbeitsschritte des Projektes zum Tragen. Und mit der Software AND CompLex wurde eine Präsentationsform des LexMA auf CD-ROM gefunden, die neben den Vorteilen der Anwendung – komplexe Suchanfragen für die inhaltliche Recherche, zuverlässige Darstellung aller Sonderzeichen, vielfältige Navigation im verzweigten Artikelsystem des LexMA – vor allem den Vorteil bietet, XML-basiert zu arbeiten.

Mit diesen Voraussetzungen konnte der Prozess der elektronischen Erschließung von neun voluminösen Bänden LexMA äußerst rationell gehalten werden und erlaubte dennoch eine Anreicherung des Textbestandes durch zahlreiche redaktionelle Meta-Informationen. Trotzdem mussten die Bände erst vollständig neu doppelt erfasst und durch einen anschließenden Abgleich der beiden Versionen auf Erfassungsfehler hin analysiert werden, bevor das Textcorpus in elektronischer Form (ca. 90 MB) vorlag. Dieses war der Ausgangspunkt für den anschließenden, auf 6 Monate terminierten Umsetzungsprozess in das CompLex-Endformat, wobei die Möglichkeit einer umfangreichen Kontrolle der Datenqualität in den dann erreichten ca. 170 MB Gesamtdaten nicht mehr zur Verfügung stand.

Voraussetzung für die Einhaltung eines derart kurzen Bearbeitungszeitraums unter Gewährleistung optimaler Datenqualität war die durch TUSTEP gegebene Möglichkeit, aus verschiedenen SGML- bzw. XML-Komponenten Daten-systeme zu interpolieren, die sich gegenseitig strukturell und inhaltlich kontrollieren und zum endgültigen Datenbestand hin ergänzen. Im Folgenden werden drei wesentliche dieser Komponenten des LexMA-Projektes vorgestellt, auf die sich die Architektur der SGML-basierten, interpolierenden Datensysteme stützte.

Komponente 1: Redaktionseigene Lemma-Datenbank

Nach der Überführung der redaktionseigenen Lemma-Datenbank in SGML-Strukturen konnten ihre Meta-Informationen zur Strukturierung des erfassten Textcorpus genutzt, ihre Informationen selbst wiederum durch das Textcorpus korrigiert und ergänzt und dann in den Gesamt-Datenbestand eingespielt werden.

SGML-basierte Strukturierung des Textcorpus:

- Abgleich der aus den erfassten Daten gewonnenen Lemmaliste mit den in der Datenbank verzeichneten Lemmata (ca. 36.000);
- Abgleich der Lemmaschreibung;
- Abgleich der internen, hierarchischen Strukturierung von Lemmata. Gemeint sind hier Lemmata, die in sich selbst in weitere Unterabschnitte untergliedert sind. Im erfassten Buchtext sind diese durch Zwischenüberschriften erkennbar, in der Datenbank gelten sie als eigene Lemmata;
- SGML-Auszeichnung der nur per Verweispfeil gekennzeichneten Querverweise im erfassten Buchtext durch Abgleich mit den in der Datenbank verzeichneten Lemmata (ca. 250.000; Trefferquote der realisierten Hyperlinks auf der CD-ROM: ca. 92 %).

Korrektur/Ergänzung durch das Textcorpus:

- Gegenkorrektur fehlerhafter Lemmaschreibungen, nicht erkennbarer interner Hierarchisierung der Lemmata in der Datenbank;
- Einspielung nicht vorhandener Meta-Informationen, wie die Band-Spaltenreferenz eines jeden Lemmas oder die explizite Angabe des hierarchischen Ortes jedes Datenbank-Lemmas (Zugehörigkeit zu einem vorgeordneten Lemma als Zwischenüberschrift).

Einspielung der Meta-Informationen in den Gesamt-Datenbestand:

- Verschlagwortung der Lemmata nach in der Datenbank vorliegenden Kategorien, wie „Person“, „Fachgebiet“ etc. Diese Meta-Informationen sind auf der CD-ROM als Indices realisiert.

Komponente 2: XML-Daten des Textcorpus

Die Strukturierung des erfassten Textcorpus hin zu einem validen XML-Datenbestand musste ohne Hilfe einer vorliegenden DTD erfolgen, da sich deren Element-Deklarationen vollständig erst nach Abschluss der Strukturierung des Gesamt-Datenbestandes und Analyse seiner Elemente ergeben hätten. Die Erarbeitung der Struktur des Textcorpus vollzog sich deshalb in iterierenden Schritten:

- Strukturierung der Daten nach einer ersten, primitiven DTD mit TU-STEP;
- Parsen der Daten gemäß dieser DTD;
- Analyse der Fehler, Rückschluss auf Datenfehler/Unvollständigkeit der DTD;
- Erstellen einer neuen DTD;
- Wiederholung des Prozesses.

Komponente 3: Parallelsatz zur Buchausgabe

Die Erstellung eines Band-Spalten-gleichen Parallelprints zur Originalausgabe sowie umfangreicher Register indizierter Stellen aus den XML-Daten des Textcorpus gewährleistete neben der Dokumentation des Arbeitsprozesses die Kontrolle der Umsetzung von Zeichenkodierungen. Durch die Indizierung konnten fehlerhafte Kodierungen im SGML-basierten Register selbst korrigiert und dann in den Gesamt-Datenbestand zurückgespielt werden. Dies ermöglichte eine Abkoppelung der Einzelkorrekturen vom übrigen workflow.

Interpolierende Datensysteme - Anforderungsprofile für ihre Architektur

Die Erschließung eines durch umfangreiche inhaltliche Informationen angereicherten XML-Datenbestandes aus proprietären Systemen oder der Primärerfassung wird sich niemals ohne Handeingriffe und redaktionelle Arbeiten realisieren lassen. Die Architektur interpolierender Datensysteme aber kann einen erheblichen Beitrag zur Unterstützung und Rationalisierung solcher Arbeiten leisten, wie an der erfolgreichen Durchführung des LexMA-Projektes deutlich wird. Bedingung dafür ist allerdings die Nutzung aller verfügbaren Daten-Ressourcen; damit lassen sich, unabhängig von den Gegebenheiten eines konkreten Projektes, einige allgemeine Voraussetzungen für die Implementierung solcher Systeme formulieren:

- freie Überführbarkeit bzw. Rückführbarkeit der Daten in beliebige, auch proprietäre Systeme;
- Vorhaltung der Daten in neutralen SGML-Metasytemen zur Generierung „on-the-fly“, d.h. zur Generierung von temporären, partikulären Teil-Systemen;

- Re-Komponierbarkeit dieser Teil-Systeme wie einzelner Datensätze aus allen Systemen nach frei definierbaren Regeln;
- Indizierung aller Elemente der Systeme, bis hin zu Einzelzeichen als Voraussetzung ihrer Re-Komponierbarkeit;
- Überprüfbarkeit der strukturellen wie inhaltlichen Validität der Systeme während des gesamten Arbeitsprozesses durch Vergleich der indizierten Elemente.

Thomas Weerth (Tübingen)

III. Zusammenfassung der verlegerischen Aspekte

Verlegerischer Fluchtpunkt der technischen Ausführungen von Herrn Weerth und der redaktionellen von Herrn Meier ist die Schaffung einer elektronisch ebenwertigen Datenbasis der zahlreichen Großwerke, über die der Verlag J.B. Metzler bereits verfügt oder in Angriff nimmt.

Die Redaktionen der jetzt laufenden Großwerke („Der Neue Pauly“, „Musik in Geschichte und Gegenwart“ (zusammen mit dem Bärenreiter-Verlag), „Lexikon des Mittelalters“ u.a. m.) arbeiten unter höchst unterschiedlichen, der Entstehungsgeschichte dieser Projekte geschuldeten elektronischen Voraussetzungen. Der Verlag ist dabei, hier gleichartige Voraussetzungen zu schaffen, die sich bis in die Datenverwaltung und die Datenpflege im Verlag erstrecken.

Der Verlag versucht, die mit dieser Aufbereitung verbundenen hohen Investitionen teilweise durch den gezielten Abverkauf von hochpreisigen CD-ROMs abzubauen. Gleichzeitig baut er international relevante Internetkontakte zu Netzvertreibern wie Xipolis, Brepols, Brill oder Chadwyck-Healey auf, die u.a. auch mit englisch-sprachigen Versionen des „Neuen Pauly“ und des „Lexikons des Mittelalters“ auftreten werden. Als Kunden kommen überwiegend weltweit hochkarätige Bibliotheksnetze in Frage.

Für die Redaktionen bedeutet dies u.a., dass ihre Tätigkeit mit dem Abschluss des Print-Produkts nicht abgeschlossen ist, sondern das Internet in bestimmten Zeitintervallen mit updates und Supplementen bedient wird.

Um diese Strategie erfolgreich zu beschreiten, ist eine eigene Abteilung „Neue Medien“ notwendig, welche die Koordination der Redaktionen, der hauseigenen Lektorate, der technischen Herstellung der elektronischen Produkte und der outlets über die Verlagsdatenbank übernimmt.

An ein verlagseigenes Vertriebssystem im elektronischen Bereich ist noch nicht gedacht.

Bernd Lutz (Stuttgart)