

HANS PETER SCHRAMM

(Hannover)

EDV und das alte Buch

Bei meinen Studenten treffe ich immer mal wieder auf die Argumentation: ich wähle als Studienschwerpunkt Geschichte (des Buchwesens), weil ich Schwierigkeiten mit EDV habe. Wer so argumentiert irrt sich.

Zwar ist richtig: Geschichte befaßt sich mit Vergangenenem, das ja nicht mehr zu ändern ist, ist also retrospektiv, versucht Geschehenes zu verstehen und zu erklären. Elektronische Datenverarbeitung dagegen reduziert Sachverhalte auf Minimaldaten, um sie neu zu organisieren und so neu verfügbar zu machen, sie ist also auf Veränderung aus, d. h. zukunftsorientiert. Aber diese Entgegensetzung ist konstruiert; denn sie ignoriert die Verschiedenheit der Ebenen, von denen gesprochen wird: eine Sachebene und eine Organisationsebene.

Richtig ist, um es knapp zu sagen: EDV kann bei Problemlösungen helfen, und das alte Buch macht Probleme.

Problem Nr. 1: der Steckbrief

Wie weiß ich im Einzelfall, von welchem Buch die Rede ist?

Bibliotheken machen ihre Bestände durch Kataloge identifizierbar. Moderne Bücher können heute in EDV-Kataloge eingegeben werden, können von der lokalen Zerstreung regional in Verbänden zusammengefaßt werden, und durch weitere Zusammenführung (oder auch ohne Zwischenstufe) in einem nationalen Katalog vereinigt werden. Die Nutzung ist dann auch über die nationalen Grenzen hinaus möglich, natürlich mit erheblichen Einschränkungen und Schwierigkeiten, da in den verschiedenen Ländern verschiedene Katalogregeln gelten, die sich vor allen bei der Ansetzung und Ordnung auswirken. Aber dennoch können sich die meisten modernen Bücher nicht in nationalen Verkleidungen verstecken; wenn der Paß auch verschieden aussieht, trägt er doch diesselbe Identitätsnummer: die ISBN (International Standard Book Number).

Wie viel komplizierter ist die Situation für das alte Buch, wieviel mehr Möglichkeiten, sich zu verstecken! Keine ISBN, Titel in ungewohnter Diktion, orthographisch willkürlich, mit fingierten Druckorten, verstecken Beiträgern,

schwer auflösbaren Pseudonymen, mit Aussagen über längst vergessene Sachverhalte. Hinzu kommen exemplarbezogene Informationen über Vorbesitzer, Marginalien, Ausstattung, Einband u.a. Könnten neu katalogisiert werden, wäre das Verfahren noch relativ einfach, wenn man sich international auf ein einheitliches Regelwerk einigen könnte. Aber bei der ungeheuren Bestandsmenge von vielen Millionen Bänden ist eine generelle Neukatalogisierung utopisch. Will man dennoch für die Altbestände einen EDV-gestützten Katalog erstellen, muß von den vorhandenen Katalogen ausgegangen werden, die z.T. noch aus Zeiten stammen, in denen es überregionale Regelwerke noch gar nicht gab. Eine überregionale oder gar internationale Zusammenarbeit scheint da kaum realisierbar. Und doch wird bereits im Verbund gearbeitet, und die Zusammenfassung der einzelnen Kataloge in einer europäischen Datenbank ist in Planung.

Wie strukturiert sich diese Zusammenarbeit?

In den einzelnen Bibliotheken werden maschinenlesbare Katalogaufnahmen erzeugt, entweder in Neuaufnahmen oder durch Umwandlung vorliegender Katalogisate. Neuaufnahmen können nur für eine relativ kleine Menge realisiert werden, auch wächst die Titelmenge nur langsam an. Allerdings haben diese Aufnahmen einen hohen Anspruch und bieten alle für Altbestände relevanten Informationen (Beispiel; Herzog August Bibliothek Wolfenbüttel). Der weitaus grössere Teil der EDV-Aufnahmen entsteht durch die Umsetzung vorliegender Katalogisate in maschinenlesbare Form: Katalogkonversion. Da hierzu nur in Ausnahmefällen die Bücher selbst für die Aufnahme herangezogen werden, können diese neuen Katalogisate nicht mehr bieten als die vorliegende Katalogaufnahme. Und da die Kataloge z.T. noch aus den Zeiten stammen, als jede Bibliothek nach ihren eigenen Regeln erfaßte, ist mit Uneinheitlichkeiten zu rechnen, müssen Abstriche gemacht werden. Trotz vieler Bedenken wird das Verfahren praktiziert, und es werden so relativ schnell große Mengen von Katalogaufnahmen über EDV verfügbar. Nicht nur in Göttingen und München wird konvertiert, auch in Paris an der Bibliothèque Nationale und in London an der British Library.

Der Nutzen der nationalen Katalogarbeit mit EDV kommt aber erst dann zum Tragen, wenn die Katalogisate international verfügbar sind. Dann besteht auch eine größere Möglichkeit, minderwertige Katalogaufnahmen an höherwertige anzuhängen — sofern nur die Identität sichergestellt ist — und sie so aufzubessern. Für eine solche Internationale Zusammenarbeit gibt es inzwischen ein Konzept: Im Entstehen ist eine „Eurodatenbank“ für Altbestandskatalogisierung (Claudie Fabian, *Die Gründung der „Eurodatenbank“ für Altbestandskatalogisierung in Europa*. In: Zeitschrift für Bibliothekswesen und

Bibliographie, 39, 1992, S. 207 ff.) Standort und Organisation sind noch unklar; ein Präkonsortium führt die Planungen, der Teilnehmerkreis ist beachtlich: Neben den Nationalbibliotheken von Frankreich, England, Schottland, Irland, Belgien sind die Staatsbibliotheken in München und Göttingen und mehrere bedeutende Universitätsbibliotheken vertreten. Natürlich wird an ein gemeinsames Regelwerk gedacht; wahrscheinlich das ISBD (A) = International standard bibliographic description for older monographic publications (Antiquarian), das dann aber nicht verbindlich vorgeschrieben ist, sondern den neu mit Katalogisierung beginnenden Bibliotheken empfohlen wird. Die meisten Altbestandsdaten liegen heute in einem MARC-Format vor bzw. in einem nach UNIMARC konvertierbarem Format; für die deutschen Bibliotheken ist wichtig, daß auch das MAB-Format in UNIMARC konvertierbar ist. Es könnte auf diesem Wege sogar einmal Wirklichkeit werden, daß für die alten Titelerückwirkend eine ISBN-alt vergeben wird. Das setzt allerdings voraus, daß die Identitätsfrage geklärt ist. Alte Bücher sind in dieser Hinsicht eigensinniger als neue.

Problem Nr. 2: Der Doppelgänger

Wenn zwei Bücher dieselbe Titelaufnahme erhalten, ist damit noch nicht gesagt, daß sie Erzeugnis desselben Setz- und Druckvorgangs sind. Die Titelblätter von zwei Exemplaren des Kirchen- und Ketzeralmanachs aufs Jahr 1781 sehen auf den ersten Blick identisch aus. Der zweite Blick, oder doch ein späterer erkennt dann, daß die Titelblätter von verschiedenen Sätzen stammen. Was aber hilft diese Erkenntnis, wenn die Informationen des Titelblattes in beiden Exemplaren identisch sind, ebenso wie die Kollation? Bibliothekarisch sind beide Exemplare identisch, also zwei Exemplare desselben Titels. Daß sie von zwei verschiedenen Setz- und Druckvorgängen stammen, kommt in der Titelaufnahme nicht zum Ausdruck, ist normalerweise auch gar nicht erkennbar, da in der Regel zur Erfassung nur ein Exemplar vorliegt, also gar nicht verglichen werden kann. Der Doppelgänger ist von dem Original nicht zu unterscheidenwürden Páße ausgestellt, erhielten beide Identitätsnummer.

Nicht alle Doppelgänger sind gleich raffiniert; wird ein Text ein zweites Mal gesetzt, weicht er in der Regel im Umfang, zumindest aber im Umbruch von dem Vorgängerdruck ab. Eine Ausnahme bilden die Texte, die zeilengetreu neugesetzt werden, „Männchen auf Männchen setzen“ heißt das in der Druckersprache. Derartige Doppelgänger bleiben oft unerkannt und sind ohne Hilfsmittel nur schwer zu entlarven. Diese Detektivarbeit hat sich ein besonderer Forschungszweig der Buchgeschichte vorgenommen: die analytische Druckforschung. In Deutschland arbeitet v.a. Martin Boghardt in dieser in den angelsächsischen Ländern entwickelten Disziplin (Martin Boghardt, *Analytische*

Druckforschung. Ein methodischer Beitrag zu Buchkunde und Textkritik. Hamburg, 1977.) Lange Zeit hat Boghardt hierzu den Hinman Collator verwendet, ein Gerät, das die entsprechenden Seiten zweier verschiedener Exemplare zeitlich abwechselnd auf dieselbe Fläche projiziert; dabei erscheint dann alles identische unbewegt, die Unterschiede aber scheinen zu springen. Inzwischen setzt Boghardt ein besonderes, für seine Zwecke entwickeltes EDV-Programm ein, das weit mehr Komfort bietet und die mühsame Arbeit an dem Objektiv des Hinman Collators ablöst. Zur Überprüfung zweier Exemplare auf Identität oder Differenz werden die entsprechenden Seiten der verschiedenen Exemplare eingescannt, in der Ausgabe erscheint das identische schwarz, die Abweichungen werden rot für das eine, grün für das andere Exemplar aufgezeigt.

Problem Nr. 3: Fahndung mittels Fingerabdruck

Ein solches Verfahren lässt allerdings nicht generell für die Katalogisierung einsetzen, da zum einen der Zeitaufwand viel zu groß wäre und zum andern, wie schon erwähnt, in der Regel gar nicht mehrere Exemplare desselben Titels für die Titelaufnahme einsehbar sind. Und ein weiteres Problem: Wie kann ein überführter Doppelgänger derart in seinem Paß gekennzeichnet werden, daß klar ist, um wen es sich handelt? Besteht nicht die Gefahr, daß die Fahndungsmühe und der Fahndungserfolg umsonst war, weil es nicht machbar ist, im Paß des überführten Doppelgängers einen entsprechenden eindeutigen Vermerk einzutragen, eine Art Fingerabdruck zum Beispiel? Möglich ist eine Annotation in der Titelaufnahme, aber in welcher Form? Mit welchen Daten? Im angelsächsischen Bereich ist ein derartiges Fingerabdruck-Verfahren entwickelt worden, das jetzt auch bei der Planung des Verzeichnisses der deutschen Drucke des 17. Jahrhunderts diskutiert wird. „Der ‘Fingerprint’ ist eine Folge von Zeichen, die definierten Seiten und Zellen eines gedruckten Buches entnommen werden und die in Verbindung mit dem Erscheinungsdatum dazu bestimmt und geeignet sind, mit Hilfe des Computers Exemplare der gleichen Ausgabe zusammenzuführen und Drucke unterschiedlicher Ausgaben zu unterscheiden, und zwar ohne physischen Vergleich der Objekte“ (John P. Feather: *Tests on the use of the „Fingerprint“ in library catalogues.* Oxford, 1977, S. 1). Dieser Fingerabdruck besteht aus 16 Zeichen, die der letzten und vorletzten Zeile auf vier definierten Seiten eines Druckes entnommen werden, ergänzt um ein Erläuterungszeichen und das Erscheinungsjahr. Mit dieser Zeichenkombination lassen sich erstaunlich genau Drucke knapp definieren — es sei denn, es sei sorgfältigst — „Männchen auf Männchen“ gesetzt worden. Obwohl laut Umfrage vor knapp zehn Jahren schon damals 38 Bibliotheken mit Fingerprints katalogisierten, ist bis heute kein einziger Fall bekannt geworden, daß derselbe Fingerabdruck für zwei verschiedene Drucke

ke zuträfe. Da ein Fingerprint von einem eingearbeiteten Katalogisierer in durchschnittlich 3 Minuten bestimmt werden kann, läßt sich vorstellen, wieviel Zeit erspart wird, wenn man einen Altbestand mit Hilfe dieses ISBN-Ersatzes an einem bereits vorhandenen Katalog, z.B. einem Verbundkatalog, abgleicht.

Da die Kombinationsmöglichkeiten der 16 Zeichen sehr groß sind, reichen häufig schon vier bis acht der 16 Zeichen zur Identifizierung des Druckes aus. Entsprechend bietet sich aus demselben Grund auch oft die faszinierende Möglichkeit, unvollständige Drucke, insbesondere Drucke ohne Titelblatt zu identifizieren — freilich nur in bestimmten Grenzen. Ist nichts von dem Anfang eines Druckes erhalten, so hilft der Fingerprint nicht weiter. Was dann?

Problem Nr. 4: Torso-Fahndung

Fragmente der Frühdruckzeit, natürlich besonders Inkunabel- und Postinkunabelfragmente, werden bei der Katalogisierung mit berücksichtigt — und sind gefürchtet, da die Identifizierung ganz besondere Detektivarbeit erfordert, wie im Kriminalroman die Leiche, der man keine Fingerabdrücke mehr abnehmen kann. Trotz Konrad Haeblers Typenrepertorium kann diese Arbeit viel Zeit kosten. Wie EDV-Vorleistungen in einem solchen Fall helfen können, ist vor kurzem von Peter Zahn am Beispiel eines Handschriftenfragments nachvollzogen worden (Peter Zahn, *Kirchenväter im CD-ROM-Speicher*. Alte und neue Wege zur Ermittlung von Handschriften-Fragmenten. In: Bibliothek, 16, 1992, S. 419 ff).

In einem Archiv auf Elba wurde im Deckel eines Aktenbandes ein Fragment einer mittelalterlichen Handschrift entdeckt. Für die Identifizierung des darauf geschriebenen Psalmenkommentars wurden fünf volle Arbeitstage benötigt — es hätten auch gut mehr sein können. Der Kommentar stammt von Augustinus. Nach Abschluss dieser Identifizierungsarbeit testete Peter Zahn, wie lange er für dieselbe Arbeit gebraucht hätte, wenn er die CETEDOC Electronic Data Library, CETEDOC Data Library of Christian Latin Tests on CD-ROM (Turnhout 1991) zur Verfügung gehabt hätte. Auf dieser einen CD-ROM-Scheibe sind die Texte der bisher erschienenen (mehr als 250!) Bände des „Corpus Christianorum, Series Latina“ zuzüglich „Continuatio Mediaevalis“ gespeichert, insgesamt etwa 21,6 Millionen Worte. Da in dieser Menge frei gesucht werden kann, hatte der Autor durch die Wahl geeigneter Wortkombinationen innerhalb von weniger als fünf Minuten Augustin als Autor und den fraglichen Text ausfindig gemacht — fünf Minuten im Vergleich zu fünf Arbeitstagen!

Problem Nr. 5: Fahnden ohne Grenzen

Das letzte Beispiel macht die Leistung der EDV deutlich: die Verwaltung einer unvorstellbaren Datenmenge, die in verhältnismäßig lächerlich kurzer Zeit durchsucht werden kann. Je umfangreicher durchforschte Datenmenge ist, auf desto soliderer Basis steht das Ergebnis. Dies gilt auch für die erwähnte „Eurodatenbank“ für Altbestände: Hier kann sich eine gewaltige Menge von Buchnachweisen versammeln, über die nationalen Grenzen hinweg, ein europäischer Thesaurus librorum, der die kulturelle Gemeinschaft und Vielfalt unseres Kontinents abbildet, ohne die regionalen Eigenständigkeiten zu verwischen. Es ist nicht erheblich, ob die Datenbank in der bisher konzipierten Form entsteht oder in anderer. Entscheidend ist das erkannte gemeinsame Anliegen, durch den Einsatz von EDV einen europäischen Zusammenschluß der alten Buchbestände zu schaffen und damit auch die überregionale Verantwortung für den nationalen Besitz zu verdeutlichen. Und welche Möglichkeiten werden sich daraus ergeben! Ich weiß nicht ob man durch Literaturstudium erfahren kann, daß in der Bibliothek der Akademie der Wissenschaften in Budapest ein Band aus Hartmann Schedels Bibliothek verwahrt wird, oder ob, um dies zu erfahren, der Zufall einer Bibliotheksführung helfen muß. Bei entsprechender Datenorganisation kann sich einmal die Bibliothek Hartmann Schedels — und vieler anderer — in kürzester Zeit nachweisen lassen, sofern sie sich in Bibliotheken befindet, die an die Eurodatenbank gemeldet haben.

Das Europa des alten Buches — etwas, auf das man sich freuen muß.

HANS PETER SCHRAMM

Elektronikus adatfeldolgozás és a régi könyv

Modern elektronikus adatfeldolgozási technika és történeti diszciplínák: hogyan viselik el egymást? Az adat-menedzser a jövőbe tekint, a történész a múltba: hol van a közös találkozási pont? A kérdés feltevése — így — hibás: ugyanis a régi könyv problémát jelent, az elektronikus adatfeldolgozás pedig megoldást, de legalábbis alátámasztja azt, amennyiben az adatok igen/nem modellben ábrázolhatók.

A régi könyvvel kapcsolatos problémák megoldására az elektronikus adatfeldolgozást ma sokféleképpen használják fel. Emellett a programoknak lényegesen többet kell biztosítaniok, mint a modern könyvek esetében. A katalógizálásnál a közreműködők, a nyomdai kiállítás, a proveniencia, a széljegyzetek, a kötés stb. adatait fel kell venni és visszakereshetővé kell tenni. Az ilyen módon elkészült címfelvételeket aztán az egyedi katalógusból az összesítő (Gesamtkatalog), majd a közös (Verbundkatalog) és végül nemzetközi katalógusba lehet áttemelni és ott újból lekérdezhetővé tenni, mint a nemrégén tervbevert európai adatbankban (Europäisches Datenbank für Altstände).

Az adatbankok minősége függ az egyedi felvételek minőségétől, azaz attól, hogy kiváló színvonalú új felvételek készültek-e vagy pedig régi katalógusfelvételeket konzerválták.

A könyvek feldolgozásának egyértelmű viszonyait akkor lehetne elérni, ha megoldható lenne, visszamenőlegesen a régi könyveknek egy ISBN-hez hasonló azonosítóval való ellátása. Emellett azok a könyvek, amelyek könyvtárosi szempontból azonosak ugyan, de a valóságban különböző kiadási folyamatok

produktumai, nehézséget jelentenek. Ezek felülvizsgálata egy sok ráfordítást igénylő eljárással, az analitikai nyomdai kutatással oldható meg, azaz újabban az elektronikus adatfeldolgozás segítségével; egy plauzibilis és viszonylag egyszerű eljárás az ilyen példányok meghatározására a „fingerprint”, amelynek alkalmazása elsősorban számítógépes megoldással ésszerű. Ezzel természetesen nem lehet fragmentumokat azonosítani; ehhez teljes szövegű adattárolás szükséges, amely – mint nemrég egy példa mutatta – lehetővé teszi, hogy az azonosítási feladat a korábbi időráfordítás töredéke alatt elvégezhető legyen.