

„In einer Welt, in der längst die Bücher nicht mehr aussehen wie Bücher, sind es nur noch solche, die keine mehr sind.“<sup>1</sup>

## Read only? Read only!

Andreas Herberger

ROM steht (bekanntermaßen) für „Read Only Memory“ und man übersetzt manchmal mit „Nur-Lese-Speicher“. Man könnte über diese Bezeichnung philosophisch ins Grübeln geraten, wenn man bedenkt, daß seit Erfindung des Schreibens in Schriftkulturen nur das Lesen kulturstiftend das Gedächtnis der Menschheit gestützt hat. Kann es elektronisch etwas Vergleichbares geben? Oder sind die CD-ROMs der Tod des Buches, wie manche Auguren oder solche, die dafür gehalten werden (wollen), orakeln?

CD-ROM und Buch:  
Nur formal Konkurrenten

Formal betrachtet konkurrieren CD-ROMs sicherlich mit den Büchern, dienen sie doch wie Bücher der Verbreitung von Textinformationen. Doch bereits hier erkennt man den ersten maßgeblichen Unterschied: Das Buch ist primär als Informationsspeicher mit sequentielltem Zugriff („Lesen“) angelegt. Ein wahlfreier Zugriff ist nur teilweise durch buchgeschichtlich erst im Laufe der Entwicklung eingeführte Techniken wie Inhaltsverzeichnisse und Register möglich. Demgegenüber ist der Text auf der CD-ROM vollständig „delinearisiert“, ermöglicht doch die Technik prinzipiell den wahlfreien Zugriff auf jedes einzelne Wort des gespeicherten Bestandes. Dennoch ist es naheliegend, daß in den meisten Fällen der „Lese“-Zugriff auf den Text in der vom Buch gewohnten Reihenfolge möglich ist. Daß der sequentielle Zugriff allein aber nicht das vom Buch gewohnte „Lesegefühl“ erzeugen kann, wird jeder bestätigen, der einmal versucht hat, einen längeren Text am heute noch typographisch ungestalteten 80-Zeichen-Bildschirm eines PCs aufzunehmen. Auch die Verwendung von Farbbildschirmen vermag nicht über die Nachteile gegenüber typographisch gut gestalteten Büchern hinwegzutäuschen. Werbungen wie „Die tragbare Staatsbibliothek ist eröffnet“ erweisen sich bei näherem Hinsehen als unrealistisch, insbesondere dann, wenn für ein Gerät erworben wird, das sich aufgrund der geringen Bildschirmabmessungen eigentlich nur für den nichtlinearen Textzugriff, kaum aber zum Lesen eignet. So kann allenfalls der Katalog einer Bibliothek sinnvoll präsentiert werden.

„Kategorial neue Form der  
Textdarbietung“

Auch für Wörterbücher oder Lexika mag ein primär nichtlinearer Zugriff ohne komfortable begleitende Lektüremöglichkeit vielleicht noch akzeptiert werden. Wenn jedoch (im Extremfall) kaum ein vollständiger Satz eines Gesetzes im Zusammenhang sichtbar ist, stellt sich die Frage, ob eine solche Reduktion des gleichzeitig sichtbaren Informationsumfangs durch die Möglichkeit wahlfreien Zugriffs gerechtfertigt wird.

Konzeption elektronischer  
Produkte

So betrachtet konkurrieren Buch und CD-ROM offensichtlich gar nicht realiter – zu hoffen ist, daß das allgemeine Bewußtsein das auch so wahrnimmt. Vielmehr ist die CD-ROM eine kategorial neue Form der Textdarbietung, die allenfalls in Ergänzung zum Buch ihren Stellenwert hat und das sicherlich auch nur für die Bücher, die im Sinne des einleitenden Adorno-Zitats noch als solche zu betrachten sind.

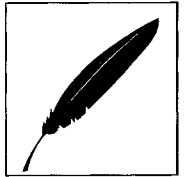
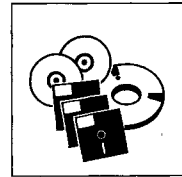
Das Verständnis der Parallelität von Buch und CD-ROM (und möglicherweise je nach Anwendungsfall anderer Medien) hat entscheidenden Einfluß auf die Konzeption eines elektronischen Produktes.

Verschiedene Ausgangssituationen sind denkbar:

Andreas Herberger ist als geschäftsführender Gesellschafter der Makrolog Gesellschaft für Logik- und Computeranwendungen sowie der MediConsult Beratungsgesellschaft für Medientechnologie für die jur-pc CD-ROM-Edition verantwortlich. Makrolog ist Hersteller der OCR-Software OPTOPUS und entwickelte gemeinsam mit CDI Ltd, Israel, die deutsche Version der Retrievalsoftware CDIS.

1. Es liegt ein als Buch gedrucktes „Vorbild“ für die elektronische Ausgabe auf CD-ROM vor.
2. Der als CD-ROM zu publizierende Text liegt nur in elektronischer Form als Datenbank vor, die möglicherweise online zugänglich ist.
3. Das Produkt wird vollkommen neu konzipiert und muß sich weder an einem gedruckten Vorbild noch an der vorgegebenen Struktur eines vorhandenen Datenbestandes orientieren.

1 T. W. Adorno, Minima Moralia, S. 83, Suhrkamp Verlag, 1951



## CD-ROM-Produktion ohne „Altlasten“

Offensichtlich ist die gestellte Aufgabe für den dritten Fall ohne Probleme zu lösen:

Die Konzeption des Produktes kann sich ganz und gar an den spezifischen Vorteilen der CD-ROM orientieren und sollte dies auch tun. Neben den üblichen Fragen bei der Datenbankkonzeption (Volltext, Felder, Synonyma, Metrik, Stopworte etc.) sollte größtmöglicher Wert auf eine mediengerechte Aufbereitung gelegt werden.

„Mediengerecht“ meint in diesem Zusammenhang nicht allein die technischen Voraussetzungen, die für das Pressen der Daten auf die CD-ROM gegeben sein müssen. Es geht vielmehr auch um Fragen der Präsentation des Ergebnisses an dem für die Abfragen genutzten Computer, dessen bestimmende Komponente das Bildschirmformat ist. Heute wird man hier auf noch überwiegend verbreiteter Technik zunächst mit maximal 80 Zeichen je Zeile und 25 Zeilen operieren können – ein Erbe aus der Zeit der Lochkarte übrigens. Keineswegs selbstverständlich scheint zu sein, daß eine mediengerechte Aufbereitung unter den genannten Voraussetzungen bedeutet, daß auf Worttrennungen oder einen Pseudoblocksatz durch Einfügen von Leerstellen zu verzichten ist, da beides bei den genannten limitierten Darstellungsmöglichkeiten keineswegs zur Verbesserung der Lesbarkeit führt.

Mit zunehmender Verbreitung der graphischen Oberflächen und, damit zumeist einhergehend, dem Einsatz größerer und höher auflösender Bildschirme werden zusätzliche Kriterien die mediengerechte Präsentation beeinflussen. Zu fordern ist dann z. B., daß in Abhängigkeit der gewählten Ausschnittgröße der Text automatisch reformatiert und damit optisch so weit wie möglich an den Eindruck der heute üblichen Typographie angepaßt wird, die – richtig eingesetzt – sicher kaum noch Verbesserungsmöglichkeiten erwarten läßt. Selbstverständlich sollte sein, daß nicht allein die gewählte Software die Aufbereitung des Datenbestandes bestimmen darf. Die Anpassung der Strukturierung des Datenbestandes an die Möglichkeiten der aus anderen Gründen ausgewählten Software wäre grundfalsch, da dadurch die Einschränkungen der Software gewissermaßen dem Datenbestand 'übergestülpt' würden und von vorlagenangemessener Aufbereitung keine Rede mehr sein könnte.<sup>2</sup> Der Markt ist heute so vielfältig, daß diesem die Parallelität mehrerer Retrievalprogramme für einen Bestand von Daten aus der Hand eines Verlegers zumutbar sein sollte. Wer demgegenüber um „Marktverwirrung“ besorgt ist, unterschätzt den Markt und das Prinzip demokratischer Konkurrenz. Würde man sich anders entscheiden, so müßte man in Kauf nehmen, daß der Benutzer bei der Arbeit mit einem größere Sinnzusammenhänge enthaltenden Text ein „Telefonbuchgefühl“ nicht los würde.

Bei allen Vorteilen, die möglicherweise ein elektronisches Produkt für bestimmte Bestände aufweist, sollte nie außer acht gelassen werden zu prüfen, ob nicht gegebenenfalls eine parallele Veröffentlichung in gedruckter Form sinnvoll ist. Der Aufwand hierfür dürfte bei korrekter Datenaufbereitung normalerweise gering sein, könnte aber die „Nachbereitung“ von Suchergebnissen im Papier unter Umständen komfortabler gestalten und vielleicht sogar einen breiteren Kundenkreis ansprechen, der über die Papierfassung zur CD-ROM kommt.

## Aus der Datenbank zur CD-ROM

Zum zweiten Fall: Umsetzung eines elektronisch vorliegenden Bestandes. Hier gilt im Grundsatz alles vorher für den Fall des vollständig neu konzipierten Produkts Angeführte. Allerdings verhindert eine wesentliche Einschränkung normalerweise die optimale Umsetzung: Der vorliegende Datenbestand ist nicht so aufbereitet, wie dies nach dem heutigen Stand der Technik geschehen würde. Ein einfaches Beispiel: Die metrische Suche (also die Berücksichtigung von sprachlichen Einheiten wie Satz und Absatz) ist für CD-ROMs eigentlich Stand der Technik. In den seltensten Fällen wird aber in einem über Jahre (oder Jahrzehnte) gepflegten Datenbestand eine (eindeutige) Satzdemarkierung vorhanden sein: Die einfache Verwendung von Punkten als Satzdemarkierung ist kaum als präzise anzusehen, da auch Abkürzungen durch einen Punkt abgeschlossen werden können. Allerdings kann man an diesem Beispiel erkennen, daß auch ein nicht optimal aufbereiteter Da-

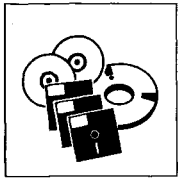
*Mediengerechte Aufbereitung*

*„Pluralität von Retrievalprogrammen“ gegen „Telefonbuchgefühl“*

*Synergie-Effekte von CD-ROM- und Buch-Ausgabe*

*Strukturverbesserungen in bereits elektronisch vorhandenem Bestand.*

<sup>2</sup> Vergleiche zum Stand der Technik bei Volltextdatenbanken: Yoram Hacohen, Entwicklung eines Informationsretrievalsystems für große Datenbestände auf CD-ROM, MAKROLOG newsletter 3/90, S. 10 ff.



tenbestand keineswegs 1 : 1 übernommen werden muß. Viele qualitätsverbessernde Maßnahmen (auch die korrekte Satzendemarkierung oder die Auflösung von Abkürzungen, um deren Suchfähigkeit zu ermöglichen) lassen sich in deutschen Texten heute mit ausreichenden Trefferquoten maschinell durchführen, so daß der Arbeitsaufwand zur Qualitätsverbesserung gering ist.<sup>3</sup> Ausscheiden werden in den meisten Fällen aufwendige manuelle Strukturänderungen oder -verbesserungen, da dies kostenmäßig selten zu rechtfertigen sein wird.

### Vom Buch zur CD-ROM

*Der „Normalfall“:  
Ein Buch soll elektronisch  
aufbereitet werden.*

Sieht alles vielleicht noch recht einfach aus, wenn der Bestand schon elektronisch vorliegt, so wandelt sich das Bild zum Komplizierteren hin, wenn der „Normalfall“ eintritt und ein Buch (oder gar eine Bibliothek) für die elektronische Darbietung aufbereitet werden soll. Prospektiv wird die Anzahl dieser Fälle abnehmen, da die meisten Bücher ohnedies irgendwo bereits elektronisch vorliegen werden (sei es im PC des Autors oder in einem Rechner der Setzerei).

*Dreifach-Handerfassung als  
Garant für Qualität?*

Aber auch wenn nur Papier vorliegt, was bei auf alte Textbestände bezogenen Wissensprojekten noch lange der Fall sein wird, hat man es nicht mit einem ernsthaften Hindernis zu tun. Zum einen kann der Text durch Handerfassung in eine elektronische Form gebracht werden (eine ganze Anzahl von Datenbeständen wird z. B., ausgehend von der Theorie, das bei mehreren Erfassern selten gleiche Fehler vorkommen, gleich dreimal erfaßt).

*OCR-Erfassung – mehr als eine  
Alternative?*

Alternativ bieten sich leistungsfähige OCR-Programme an, die den Text mittels Scanner automatisch übernehmen. Größter Wert ist dabei auf die Umsetzung aller Strukturinformationen des gedruckt vorliegenden Textes (Absätze, Schrift- und Seitenwechsel etc.) zu legen. Auch Probleme wie die korrekte Erfassung von Sonderzeichen und – insbesondere bei juristischen Texten – die fehlerfreie Umsetzung einer Vielzahl von Abkürzungen sollten Minimal Kriterien für die Eignung eines OCR-Programms sein.

*Erfordert den Spezialisten:  
Buchkompatible  
Datenumsetzung*

Ist dann der Text auf die eine oder andere Art und Weise erfaßt worden, ist der nächste Schritt, der gar nicht so einfach ist, die „buchkompatible“ Umsetzung in eine CD-ROM-Fassung. „Buchkompatibel“ heißt hier, daß die CD-ROM-Fassung mindestens den Nutzen des Buches haben muß, sinnvollerweise aber gegenüber dem Buch Zusatznutzen aufweist. In keinem Falle darf es passieren, daß die CD-ROM-Fassung geringere Möglichkeiten als das Buch bietet. Kann man z. B. auf der CD-ROM nicht gezielt eine Seite „aufschlagen“, zwingt das den Benutzer regelmäßig zum Medienbruch z. B. beim Überprüfen von Zitaten. Gleiches gilt für eingeführte Techniken wie Inhaltsverzeichnisse oder strukturierte Register, die allesamt auf der CD-ROM abgebildet sein sollten. Auch sollte jederzeit auf der CD-ROM die korrespondierende Stelle im Buch erkennbar sein (z. B. durch Mitführen einer Seitenzahl in der Fußzeile), so daß der Wechsel oder die Zitierung einfach möglich ist.

Und last but not least: In Büchern kann man Notizen anbringen und ggf. Seiten kopieren. Daß dies auch in einer CD-ROM-Fassung möglich ist, wurde mit den MediConsuh-Produkten unter CDIS bewiesen. Sie bieten Buchkompatibilität in höchstem Maße.

### Das Ziel: CD-ROMs auf dem aktuellen Stand der Technik

Als Fazit läßt sich festhalten, daß die Produktion guter CD-ROMs äußerst vielschichtig ist und strategische Planung erfordert. Besonders wichtig ist die Kenntnis des technisch Machbaren, damit die resultierenden Produkte dem Stand der Technik entsprechen. Werden diese Voraussetzungen berücksichtigt, ist zu hoffen, daß die von Adorno bereits 1951 geäußerte Kritik am Zustand der Bücher den CD-ROMs erspart bleiben möge.

<sup>3</sup> Die Möglichkeiten der automatischen Strukturverbesserung erläutert Christian Michel in seinem Beitrag „Neue Medien braucht das Land“, MAKROLOG newsletter 1/91, S. 3 ff., und 2/91, S. 9 ff.