

**LUTHER DIGITAL**  
**Die Digitale Bibliothek**

Leonhard Rau  
IBM Deutschland Informationssysteme GmbH  
Überseering 24  
22297 Hamburg  
Tel. (040) 6389 3945

*A: Das Projekt Luther Digital*

*1. Historischer Rückblick und Einleitung*

Alle Innovationen im Medienbereich basierten auf zwei Voraussetzungen: einer neuen Technologie und einer rapide steigenden Nachfrage.

Jahrhundertlang wurden die Bücher in Klöstern mit der Hand geschrieben. Renaissance und Humanismus förderten im 15. Jahrhundert die Nachfrage nach Büchern. Erste Ansätze für eine Vervielfältigung brachte der Holzschnitt. So entstand in der ersten Hälfte des 15. Jh. das Blockbuch. Bild und Schrift wurden für die Vervielfältigung in Holztafeln geschnitten. Jedoch erst der Typendruck, dessen Vorstufen es schon seit Jahrhunderten in Form von Münzpressen gab, brachte den Durchbruch, Bücher und Schriften in großer Anzahl zu erstellen. Die große Leistung Gutenbergs war die Herstellung von einzelnen Buchstaben aus Metall und die Erfindung der Druckerpresse. Aber erst die Reformation mit ihren auch den ‚gemeinen Mann‘ interessierenden Fragestellungen sicherte dem Buchdruck mit beweglichen Lettern ein großes Publikum. Die von Gutenberg entwickelte Basistechnik ermöglichte die rationelle Erstellung und Verbreitung von Büchern bis in unsere Zeit.

Alle bisherigen Techniken teilen jedoch das Schicksal, daß das Gedruckte mit dem Datenträger Papier eine untrennbare Verbindung eingeht.

Die Computer-Technologie löst nun mehr und mehr die von Gutenberg entwickelte Drucktechnik ab. Dies hat größere Auswirkung als es auf den ersten Blick zu sein scheint. Während die Erstellung von neuen Büchern und Schriften heute schon weitgehend mit Computer-Programmen geschieht, eröffnen sich insbesondere für historische Werke neue Perspektiven, auf die noch eingegangen werden wird.

*2. Herausforderungen für Archive, Bibliotheken und Museen*

Historische Dokumente sind durch eine Vielzahl von Faktoren gefährdet, z.B. durch den Alterungsprozeß des Papiers, durch Schädlinge und durch klimatisch ungünstige Bedingungen. Die Benutzer tragen, wenn sie das Material benutzen, gleichermaßen zum Zerfallsprozeß bei. Daher sind sie häufig vom Zugriff grundsätzlich ausgeschlossen. Wertvolle Informationen bleiben somit der Wissenschaft und dem interessierten Publikum vorenthalten. Vor diesen Problemen stehen alle Bibliotheken, Archive und Museen der Welt, die historische Dokumente zu ihrem Fundes zählen. Ansätze, historische Bestände durch Mikroverfilmung zu sichern und zugänglich zu machen, sind keine dauerhafte Lösung, da auch der beste Mikrofilm altert und irgendwann erneut die Erfassung der dann auch weiter gealterten Originalbestände erfordert. Der Schwerpunkt muß sich von der Dauerhaftigkeit des Mediums auf die Dauerhaftigkeit der Information in einer digitalen Umgebung verlagern.

### 3. Lösungsmöglichkeiten durch Image-Verarbeitung

Image-Verarbeitung (Abbild-Verarbeitung) im öffentlichen und kommerziellen Bereich zielt im wesentlichen auf die Zählung der Papierflut, den schnellen Zugriff zu den erfaßten Dokumenten und die Möglichkeit, die Prozesse effizienter zu gestalten. Anders die wichtige Aufgabe von Bibliotheken und Archiven, historisch seltene oder einmalige Exponate zu erhalten und trotzdem zugänglich zu machen. Hier bietet die Image-Verarbeitung ideale Lösungsmöglichkeiten, die beide Ziele Schutz des Originals und Zugriff auf die Information in optimaler Weise miteinander verbindet.

#### Technische Beschreibung

Informationen können durch Scanner (Bildabtaster) elektronisch als Image (Abbild) erteilt werden. Danach kann das Original konsequent archiviert und somit geschont werden, da die Informationen von nun an digital (Bild- und Farbpunkte) vorliegen. Sind die Bestände einmal erfaßt, so sind diese genau wie neue Werke leicht und wirtschaftlich reproduzierbar und einem breiten Publikum zugänglich. So ergeben sich z.B. für die Geschichtsforschung völlig neue Möglichkeiten, da bisher verborgene Quellen zugänglich sind.

Der Begriff ‚digital‘ stammt aus dem lateinischen ‚digitus‘, dem Finger, als Hilfe zum Zählen. Unter Digital wird die Darstellung von Informationen (hier Bildpunkte) in einer Folge von Nullen und Einsen verstanden. Die Bildinformationen sind von nun an abstrakt, medien-unabhängig und keinem Zerfallsprozeß mehr unterworfen. Mit dem fortschreitenden Technologiewandel können diese Informationen verlustfrei auf andere Datenträger übertragen werden. Die Erfassungstechnik ist heute so ausgereift, daß die Reproduzierbarkeit den allerhöchsten Anforderungen der Drucktechnik entspricht und weitere Steigerungen für das menschliche Auge nicht mehr wahrnehmbar sind. Die digitalisierten Informationen sind genau wie kodierte Informationen (z.B. Buchstaben, Zahlen, Formeln) Programmen zugänglich, was völlig neue Perspektiven eröffnet (z.B. elektronisches Restaurieren, komplexe Suchvorgänge, MultiMedia).

Obwohl die Basiskomponenten eines Image-Systems bei allen Image-Anwendungen die gleichen sind, werden an die einzelnen Komponenten der Image-Verarbeitung im Bibliotheks- und Archiv-Bereich sehr spezielle Anforderungen gestellt. Während die Erfassung (durch Scanner) im kommerziell/öffentlichen Bereich mit größtmöglicher Geschwindigkeit bei minimaler manueller Einflußnahme erfolgen soll, sind hier die maximale Schonung des Originals bei der Erfassung und größtmögliche Auflösung gefordert. Dies führt dazu, daß der Erfassungsprozeß zeit- und speicheraufwendig ist. Spezial-Scanner wurden entwickelt, da die am Markt allgemein bekannten Geräte diesen Anforderungen nicht genügen.

### 4. Nutzen

Der Nutzen der digitalen Informations-Speicherung ist u.a.:

- *leichter und schneller Zugriff zu den Informationen*

Die Informationen können kostengünstig einer breiten Öffentlichkeit zugänglich gemacht werden.

- *Zugriff mit Suchkriterien, die bisher kaum möglich waren*

Durch die Verknüpfung von Suchkriterien ergeben sich für die wissenschaftliche Bearbeitung neue Möglichkeiten, die bisher schon aus zeitlichen Gründen nicht durchführbar waren.

- *Nutzung anderer bzw. zukünftiger Technologien*

Es ist z.B. möglich, Dokumente elektronisch zu restaurieren mit der Folge, daß die Kopie in besserer Qualität als das Original vorliegt.

- *Reproduzierbar in Büchern*

- *Verknüpfung mit Bibliotheks-Verwaltungs-Programmen*

- *Der bedeutsamste Nutzen ist,*

daß die Informationen, wenn sie einmal digital vorliegen, der Nachwelt als Kopie in unveränderlicher Qualität erhalten bleiben und das Original die größtmögliche Schonung erfährt.

## 5. Bekannte Lösungen

Die IBM arbeitet u.a. an einigen international bekannten Projekten in diesem Bereich. Diese sind:

- ‚Archivo General de Indias‘, Sevilla, Spanien Dokumente aus der Zeit von 1492 bis 1898
- ‚Library of Congress‘, Bibliothek des amerikanischen Kongresses, Washington, USA
- Bibliothek des Vatikans in Zusammenarbeit mit der ‚Pontifical Catholic University of Rio de Janeiro‘. In zahlreichen Presseveröffentlichungen wurde in jüngster Zeit darüber berichtet.
- Gemeinsam realisieren die Lutherhalle Wittenberg und IBM ein weiteres Vorhaben von internationalem Rang, das Projekt ‚Luther Digital‘.

## 6. Szenarium ‚Luther Digital‘

Die Lutherhalle Wittenberg ist das größte reformationsgeschichtliche Museum der Welt. Es beherbergt mehr als 50.000 Exponate ab dem Mittelalter. Gemeinsam realisieren die Lutherhalle Wittenberg und IBM das Projekt ‚Luther Digital‘, um diese historischen Exponate digital zu erfassen und einer breiten Öffentlichkeit zugänglich zu machen. Auf der CeBIT94 sowie während des Stadtfestes 1994 anlässlich Luther's Hochzeit in Wittenberg haben wir mit großer Resonanz bei Besuchern aus Politik, Kultur und Wirtschaft die Machbarkeit dieses Vorhabens demonstriert. Am 1. August 1995 hat Sachsen-Anhalts Ministerpräsident Dr. R. Höppner das System offiziell in Betrieb genommen. In den Medien wurde über das Gezeigte und die sich eröffnenden Perspektiven berichtet.

In enger Abstimmung mit bereits laufenden Projekten wird nun das Projekt realisiert. ‚Luther Digital‘ ist das erste deutsche Digitalisierungs-Projekt in dieser Größenordnung wird von namhaften Museen und Bibliotheken aufmerksam verfolgt.

## B: Aspekte für die Planung und Realisierung eines Digital- Library-Projektes

Die hier beschriebenen Aspekte sind von grundsätzlicher Natur und dokumentieren die Überlegungen zum Projekt ‚Luther Digital‘. Genauere Planungsdetails und -erfahrungen werden im Verlaufe des Projektfortschrittes festgehalten und ihren Niederschlag in der Fortschreibung dieses Beitrages finden.

### 1. Digital-Library - ein Innovatives Projekt

Traditionelle Anwendungen im Bibliotheks- und Archivbereich behandeln weitgehend die Beschreibung und Verwaltung der Bestände. Mit der digitalen Erfassung der vorhandenen Bestände und dem Zugang zu diesen Informationen betreten wir Neuland. Datenorganisation (wegen der großen Datenmengen die hochauflösendes Scannen generiert) und Präsentation (wegen des hohen Anspruchs) sind neue Herausforderungen. Anwendungen wie z.B. elektronische Bildmanipulation (z.B. Restauration) oder Suchen nach Bildinhalten stecken noch in den Kinderschuhen. Analogien zu kommerziellen Dokumenten-Management-Systemen sind nur sehr begrenzt möglich.

Die Erfassung (Scannen) der Original-Bestände steht am Beginn des Projektes. Sie erfolgt für alle Zeiten und muß daher mit äußerster Sorgfalt durchgeführt werden, damit eine Reproduktion in der Bildqualität dem Original entspricht und keine Restriktionen für zukünftige Anwendungen bestehen. Nur in seltenen Ausnahmefällen darf eine Neuerfassung erlaubt sein. Das Ziel der lückenlosen Erfassung kann je nach Größe des zu erfassenden Bestandes Jahrzehnte dauern. Dies verdeutlicht, daß eine Wiederholung dieser Aufgabe schon aus Zeit- und Kostengründen kaum möglich ist.

Designüberlegungen am Projekt-Anfang müssen also mehr die strategische Beständigkeit und Qualität der Informationen und weniger die darauf aufbauenden Anwendungen im Blickfeld haben. Dieses Denken findet nicht auf Anhieb die ungeteilte Zustimmung.

Wegen der großen strategischen Bedeutung des ersten Schrittes, der Erfassung und Organisation der Informationen, muß das Design- und Planungsteam interdisziplinär sein, d.h. ihm müssen Experten aus dem Bibliotheks- und Archivwesen sowie der Informationsverarbeitung angehören.

## 2. Strategische Planungs-Aspekte

### 2.1 Primäre Aspekte

Die Codierung der Bildinformation einerseits und die verwendeten technischen Komponenten zur Speicherung und Bearbeitung andererseits müssen derart sein, daß weder eine Abhängigkeit zu einem Hersteller besteht, noch die Verarbeitbarkeit in anderen Systemumgebungen sowie die Migration der Informationen auf neue Speichertechnologien gefährdet sind.

Industriestandards bei den Datenformaten und Datenträgern sind ein unbedingtes Muß. Im Projekt Luther Digital werden die Daten im TIFF (Tag Information File Format) auf optischen 3,5“ Platten gespeichert. Wie im kommerziellen Bereich stellt dies sicher, daß die Daten jederzeit problemlos auf andere Datenträger überspielt werden können. Im Bereich der optischen Speichersysteme wird insbesondere die Verwendung des blauen Lasers sowie des Schichtenspeichers noch beachtliche Kapazitätswachse bringen. Das gesamte Technologieumfeld wird einem rasanten Wandel unterworfen sein. Wie im kommerziellen Bereich müssen die verwendeten Hard- und Softwarekomponenten des Systems die Teilhabe an dieser Entwicklung sicherstellen. Sonderentwicklungen, die nicht diesen Kriterien entsprechen, müssen aus den genannten Gründen verbannt bleiben, auch wenn diese bei isolierter und kurzfristiger Betrachtung gut sein mögen.

### 2.2 Sekundäre Aspekte

Während die primären Aspekten im wesentlichen die Beständigkeit der Informationen und deren Organisation zum Inhalt haben, behandeln die sekundären Aspekte die Leistungsfähigkeit des Systems und die Planung der Anwendungsgebiete. Hier sind jederzeit Nachbesserungen und Modifikationen möglich, da die Organisation und Qualität des erfaßten Bestandes davon unberührt bleiben.

Die Planung der Anwendungsgebiete ist wie die Planung der Datenorganisation Aufgabe eines interdisziplinären Teams. Die Akzeptanz der Anwendungen wird wesentlich durch die Qualität der Suchmethoden und der Präsentation der gesuchten Informationen bestimmt. Grundsätzlich lassen sich die Anwendungsgebiete in zwei Kategorien unterteilen, den wissenschaftlich orientierten Anwendungen einerseits und den allgemeinen od. populärwissenschaftlichen Anwendungsgebieten andererseits.

## 3. Auswahl der Themenbereiche

Für den Projekterfolg müssen möglichst bald und kontinuierlich in kurzen Intervallen geschlossene Themenbereiche erschlossen werden. Man könnte bei der Erfassung mit dem ersten Buch des ersten Regals beginnen, sequentiell fortfahren und mit dem letzten Buch des letzten Regals enden. Bei dieser Vorgehensweise läge der Endtermin der Erschließung des Gesamtbestandes je nach Größe viele Jahre bis Jahrzehnte in der Zukunft. Daß diese Vorgehensweise nicht sinnvoll ist, scheint offensichtlich. Um möglichst schnell verwertbare Ergebnisse zu präsentieren, werden Themenbereiche definiert, die entweder von besonderem wissenschaftlichem oder öffentlichem Interesse sind. Von besonderem wissenschaftlichem Interesse sind z.B. Quellen, die aus konservatorischen Gründen bisher nicht zugänglich waren. Von öffentlichem Interesse, welches Wissenschaftlichkeit nicht ausschließt, sind z.B. Ereignisse wie das Luther-Jahr 1996, der internationale Luther-Kongress 1997 in Heidelberg oder die Weltversammlung des Lutherischen Weltbundes im gleichen Jahr in Hongkong. Wissenschaftliche Themenschwerpunkte werden u.a. in Abstimmung mit Hochschulen gesetzt, die sich reformationsgeschichtlicher oder theologischer Forschung widmen. So werden neue Quellen für die Wissenschaft erschlossen.

Die Definition von Teilzielen fördert darüber hinaus die erforderliche Motivation zur Zielerreichung innerhalb gesetzter Fristen.

#### 4. Zeitplanung

Hochauflösendes und buchschonendes Erfassen (Scannen) ist ein zeitaufwendiger Prozeß. Zielplanung und Zeitplanung stehen in einem engen Zusammenhang. Sind die Themenbereiche definiert, so sind Schwierigkeitsgrade und Mengenvolumen des zu scannenden Materials zu ermitteln. Um die zur Objektschonung unumgänglichen Scanner-Rüstzeiten zu minimieren, muß das zu scannende Material sorgfältig nach Gleichartigkeit der Formate und Schwierigkeitsgrade strukturiert werden. Neben der Buchwippe gibt es eine Reihe weiterer Hilfsmittel, den Scanvorgang möglichst objektschonend zu gestalten.

Legt man einen 8-Stunden Tage zugrunde, beträgt die Seitenleistung bei den bekannten Projekten je Scanner-Arbeitsplatz ca. 50-100 Seiten pro Tag. Dies variiert stark in Abhängigkeit vom Zustand des zu erfassenden Materials und der Gleichartigkeit desselben. Aus der Anzahl der eingesetzten Scanner und deren Leistungsprofil läßt sich der z.T. sehr erhebliche Zeitaufwand für die Erfassung ermitteln.

#### 5. Datenorganisation

Die erfaßten Objekte werden nach Themenbereichen auf optischen Datenträgern abgelegt. Die Beschreibung dieser Objekte (Indexierung) erfolgt in einer relationalen Datenbank (DB/2). Die Beschreibung der Objekte ist wie die Erfassung ein zeitaufwendiger Prozeß. Nach dem Scannen werden zunächst nur die Inventarnummer, der Buchtitel sowie die entsprechenden Seitennummern als Index- und Suchkriterium eingegeben. Dies ist nur ein erster Schritt, um den Zugriff überhaupt zu ermöglichen. Weitere Beschreibungen z.B. zum Inhalt eines Buches oder Dokumentes müssen jederzeit möglich sein. Jede zusätzliche Beschreibung erweitert die Möglichkeit komplexen Suchens in der Datenbank. So wird der Wert des Archivs schrittweise gesteigert.

#### 6. Technische Voraussetzungen

Hier soll nicht auf die Konfiguration des Gesamtsystems eingegangen werden. Wichtig ist, daß die verwendeten Systemkomponenten ausschließlich Standard-Einheiten sind. Wie schon erwähnt, ist dies eine wichtige Voraussetzung für den Bestand und die Dauerhaftigkeit des Projektes und der Informationssicherung.

Ohne hochauflösende und objektschonende Scanner und optische Platten-Archiv-Systeme als Datenspeicher ist ein Projekt dieser Art nicht möglich. Der Einsatz optischer Speichersysteme ist inzwischen in allen Branchen anzutreffen. Der Einsatz von Spezial-Scannern ist nichts besonderes, denn für viele Anwendungsbereiche verschiedenster Unternehmungstypen sind spezielle branchentypische Scanner verfügbar. Im Bibliotheks- und Archivbereich sind Geräte erforderlich, deren Schwerpunkt bei extrem hoher Auflösung und bei maximaler Objekt-Schonung liegen. Spezielle Erfassungsgeräte (Scanner) sind also keine Besonderheit sondern typisch in allen Branchen.

Für die digitale Bilderfassung kommen zwei Scanner-Typen zum Einsatz. Welches Gerät das geeignete ist, richtet sich nach der Beschaffenheit des zu scannenden Materials.

Zwei im Bürobereich vielfach erprobte Flachbettscanner können Formate bis zu DIN A3 in schwarzweiß oder Graustufen erfassen. Diese Scanner werden für lose Blätter ohne Farbinformationen genutzt. Für Gebundenes kann dieser Scanner nur eingesetzt werden, wenn die Öffnung des Buches um 180 Grad keine Verzerrungen beim Scannen verursacht und der Buchrücken durch diese Öffnung nicht beschädigt wird.

Ein Spezial-Scanner mit Zusatzeinrichtungen zur Buchschonung (z.B. Buchwippe) und Bildausleuchtung wird für alles andere genutzt. Dieser hochauflösende Scanner kann in Farbe und Graustufen erfassen. Das maximale Bildformat kann für reflektierendes Material (z.B. Papier) bis zu ca. 60x70 cm und für durchscheinendes Material (z.B. Farbdias) bis zu 25x25 cm betragen. Der an ein Reprogerät erinnernde Scanner mit seinen Hilfsinstrumenten und unterschiedlichen Objektiven hat anstelle eines Films einen Chip, der die durch das Objektiv erfaßten Lichtwellen in digitale Bildinformationen umsetzt.

Eine weitere wesentliche technische Komponente ist das optische Plattenarchiv-System IBM 3995 (nicht zu verwechseln mit einer CD-ROM). Ohne die immense Speicherkapazität dieses Systems wären die großen Informationsmengen, die hochauflösendes Scannen generiert, nicht wirtschaftlich speicherbar. Dennoch stellt die gegenwärtige Kapazität von 1,3 GB je Scheibe einen Engpaß dar. Neue Entwicklungen wie z.B. kürzere Laser-Licht-Wellen und Schichtenspeicher werden in der Zukunft diesen Engpaß beseitigen. Wir rechnen mit wesentlichen Kapazitätssteigerungen des Speichervolumens der optischen Platten alle zwei bis drei Jahre.

## 7. Präsentations-Planung

Der Zugriff bzw. die Präsentation zu den erfaßten Informationen muß dem Interesse und den Fähigkeiten des Informationensuchenden Rechnung tragen. Die Bandbreite schwankt zwischen wissenschaftlichem und nur allgemeinem Interesse. Die Popularität des Projektes Luther Digital wird in starkem Maße von der wissenschaftlich orientierten Öffentlichkeit geprägt. Diese bestimmt wesentlich die Priorität der Themenbereiche und die Art der Präsentation. Wir stehen erst am Beginn dieser interdisziplinären Diskussion, da die primären Planungsaspekte bisher den Vorrang hatten und für den Start des Projektes wesentlich waren.

Art und Umfang der Präsentation hängen auch von den Zugriffswegen und den Datenmengen ab. Zu unterscheiden sind:

- Direkter Zugriff über Kommunikationsschnittstellen (z.B. INTERNET).
- Zugriff zu speziellen geschlossenen Themenbereichen über CD-ROM.
- Zugriff über bestehende Bibliotheks-Informationssysteme. Schnittstellen-Programme sind erforderlich, um diesen Zugriff zu realisieren. Die Lutherhalle Wittenberg verwendet Allegro/C als Bibliotheks-Verwaltungs-Programm. Über dieses Programm werden wir den Zugriff auf die entsprechenden digitalen Abbildern ermöglichen.
- Lokaler Zugriff über Datenstationen in der Lutherhalle.

Immer, wenn die digitalen Abbilder die Lutherhalle verlassen, wird den entsprechenden Seiten ein Wasserzeichen hinzugefügt. Dies dient dem Copyright-Schutz und Nachweis der Herkunft. Den rechtlichen Aspekten (rights management) einer digitalen Bibliothek wird bei der Fortschreibung dieses Beitrages ein besonderer Abschnitt gewidmet werden.

## C: Zusammenfassung

Der Leser, der ein Kochbuch zur Planung und zum Einsatz digitaler Bibliotheken erwartet hat, mag enttäuscht sein, denn dieser Beitrag behandelt im wesentlichen nur unsere Planungsüberlegungen. Konkrete Planungs- und Anwendungserfahrungen werden zu gegebener Zeit dokumentiert und veröffentlicht werden. Obgleich wir mit den bereits laufenden Projekten in engem Kontakt stehen, betreten wir mit dem Thema Digitale Bibliotheken Neuland. In regelmäßigen Fachveranstaltungen planen wir über Fortschritt und Erfahrungen zu berichten.

Die Projektdauer bezüglich der Archivierung läßt sich aufgrund des Mengenvolumens nur mit großen Unsicherheiten berechnen. Die Erschließung neuer Anwendungsgebiete wie z.B. elektronische Restaurierung und Bildmanipulation, Schrifterkennung (ICR) oder Suchen nach Bildinhalten werden neue Berufsbilder schaffen. Das technologische Umfeld und die daraus ableitbaren Möglichkeiten werden einem rasanten Wechsel unterworfen sein. Das Thema ‚Digitale Bibliothek‘ und die sich eröffnenden Möglichkeiten werden der Lutherhalle Wittenberg und anderen Institutionen, die diesem Pilotprojekt folgen, eine dauernde Aufgabe bleiben.