Digitale Bibliotheken in Deutschland

Auswirkungen auf die IT-Infrastruktur

Die digitale Bibliothek ist keine Fiktion. Schon heute sind allerorten die ersten Ansätze zu besichtigen. Es geht heute nicht mehr darum, ob, sondern wie und wann digitale Bibliotheken Realität werden. Bevor man sich damit beschäftigt, wie eine digitale Bibliothek aussehen wird, mit welchen Werkzeugen und welcher Infrastruktur sie zu realisieren ist, sollte zunächst einmal die Motivation hinterfragt werden.

Warum soll eine Bibliothek digitalisiert werden?

Bücher und Zeitschriften in ihrer klassischen Form sind durch moderne Medien nicht verdrängt worden. Im Gegenteil: Noch nie sind so viele Bücher veröffentlicht worden , noch nie so viele Zeitschriften erschienen wie heute, im Zeitalter der Multimedia-Technologie. Nicht einmal das Internet wird die klassischen Printmedien ersetzen. Das zeigt pikanterweise schon die Zahl der Bücher, die zum Thema Internet existieren.

Auf dem Weg zur digitalen Bibliothek geht es vielmehr darum, neue Möglichkeiten zu erschließen, auf die vielfältigen, weltweit vorhandenen Informationen gezielt zugreifen zu können, um sie effektiv nutzbar zu machen. Offensichtlich ist der Bedarf an der Nutzung dieser Ressourcen insbesondere im wissenschaftlichen Umfeld.

Eine Gruppe von IT-Experten[1], befragt im Auftrag der Universität Dortmund, prognostiziert mehrheitlich für einen Zeitraum kurz nach der Jahrtausendwende, daß

- Fachzeitschriften, Kongreßberichte, Habilitationen, Dissertationen, Dokumentationen vorrangig elektronisch verbreitet werden,
- statt „Exemplar“ als Maßstab für die Verbreitung Begriffe wie „print on demand“ und „pay per read“ eine zunehmende Bedeutung haben werden.

Elektronische Verbreitungsformen und damit digitale Bibliotheken eröffnen komfortable Möglichkeiten, auf weltweit vorhandene Informationen praktisch vom Arbeitsplatz aus über das Internet zugreifen zu können und die gefundenen Erkenntnisse praktisch ohne Zeitverzug für die eigene Arbeit nutzbar zu machen.

In einer Zeit, wo „time to market“, also die Geschwindigkeit von Entwicklungen eine entscheidende Rolle spielt, liegt es auf der Hand, daß diese Nutzung für kommerzielle Entwicklungen von Hochtechnologie ein kritischer Erfolgsfaktor mit zunehmender Bedeutung ist.

Die Bundesregierung[2] und auch einzelne Landesregierungen haben erkannt, daß es für den Forschungsstandort Deutschland von eminenter Bedeutung ist, durch die Bereitstellung der notwendigen Infrastruktur eine führende Rolle in der Nutzung von digital verfügbaren Daten sicherzustellen.

Aufgabenverteilung in der digitalen Bibliothek

Die Kompetenz der Bibliothekare liegt traditionsgemäß in der Erschließung, Bereitstellung und Archivierung von Dokumenten.

In einer digitalen Bibliothek sind für praktisch alle Aufgaben elektronische Hilfsmittel verfügbar und geradezu unerläßlich. Schon heute werden Dokumente ausschließlich mit elektronischen Hilfsmitteln produziert. Für die Bereitstellung und Archivierung der gedruckten Exemplare werden EDV-Systeme zur Verwaltung eingesetzt.

Aufgabe	IT- Umsetzung
Produktion	Textverarbeitung, DTP
Erschließung	GUI-Interface für OPAC, Internet-Recherche, Search Engines
Veränderung und Verarbeitung	DTP, Annonationen und URL-Links
Bereitstellung	elektronischer Katalog, WWW-basierte Services, CD-ROM-Server
Archivierung	RDBM-Systeme, CD-ROM-Archive

Die Notwendigkeit einer EDV-Infrastruktur steigt aber mit dem Grad der Digitalisierung. Es entstehen sehr große virtuelle Dokumentenspeicher in TeraByte-Größenordnung, die in entsprechenden RDBM-Systemen verwaltet werden müssen, um die Recherchierbarkeit über Jahre hinweg sicherzustellen. Die Bereitstellung der Informationen kann durch die Nutzung von Internet-Technologie - entkoppelt von der physischen Präsenz des Benutzers in der Bibliothek - gewährleistet werden. So kann ein Wissenschaftler an einem Terminal innerhalb der Universität den OPAC der Bibliothek genauso effektiv nutzen wie ein Student an seinem Heimarbeitsplatz (Intranet/Internet).

Für die wissenschaftliche Nutzung sind Internet-Techniken jedoch in einigen Fällen mit Bedacht einzusetzen. Um elektronisch publizierte Arbeiten auf Dauer zitierfähig zu archivieren, muß u.a. sichergestellt sein, daß die Referenz, z.B. die URL, erhalten bleibt.

Viele Zeitschriftenverlage (Springer, Silverplatter-ERL Archive) bieten bereits elektronische Publikationen an, die nicht nur über Internet-Technologie recherchiert werden können, sondern auch direkt zur Nutzung bereitstehen (IBIS, JASON, SUBITO)[3].

HighWire Press[4], ein elektronischer Publikationsdienst, der an der Universitätsbibliothek der Stanford-Universität beheimatet ist, zeigt diese Möglichkeiten und den Mehrwert der bibliographischen Aufbereitung.

Schnittstelle EDV und Rechenzentrum: In der digitalen Bibliothek ist eine Kooperation unerläßlich

Die Schnittstelle zwischen der EDV-Infrastruktur einer Hochschule und den inhaltlichen Aufgaben der Bibliothek werden schon in diesen Beispielen offensichtlich.

Die Aufgaben der Bibliotheken werden darin liegen, den virtuellen Dokumentenspeicher zu pflegen, indem

sie Authentizität, Verfügbarkeit und Recherchierbarkeit sicherstellen.

Um diese Ziele zu erreichen, ist eine Reihe von technischen Fragen zu klären, für die die Kompetenz in aller Regel bei den Rechenzentren bereits vorliegt. Hier einige weitere Beispiele:

Durch Firewalls muß sichergestellt werden, daß die lizenzrechtlichen Rahmenbedingungen von kommerziellen Informationsanbietern erfüllt werden. Existiert z.B. eine Campus-Lizenz für den elektronischen Zugriff auf ein CD-ROM-Zeitschriftenarchiv, müssen unberechtigte Benutzer am Zugriff gehindert werden. Auch die Abrechnung von Leistungen ist ein durchaus ernstzunehmendes Problem, für das geeignete technologische Unterstützung gewährleistet sein muß, z.B. durch SSL-Transaktionen. Letztendlich müssen auch Wege gefunden werden, die Kompetenz der Hochschulen für die Industrie nutzbar zu machen.

Auf der anderen Seite gilt es aber auch, für die technische Basis der digitalen Bibliothek einen unterbrechungsfreien Betrieb mit permanenter Erreichbarkeit der Dienste zu gewährleisten. Dies hat Konsequenzen für den Betrieb der Netzwerkinfrastruktur, der Netzwerk- und Datenbankserver, aber auch in der Bereitstellung von entsprechend qualifiziertem Personal.

Für den neutralen Beobachter aus der Industrie erscheint daraus eine Kooperation sinnvoll: Auf der einen Seite Bibliotheken mit dem Schwerpunkt auf inhaltlichen Aspekten und auf der anderen Seite Rechenzentren mit ihrer Kompetenz im Betrieb.

Die IT-Infrastruktur der digitalen Bibliothek

Während früher, und teilweise auch noch heute, die Suche in der Präsenzbibliothek nur über begrenzt zugängliche Großrechnerterminals oder gar über Mikrofilmviewer möglich war, existieren heute Client-Server-Anwendungen, die mit benutzerfreundlichen Oberflächen von jedem Arbeitsplatz im Netz aus die Recherche erlauben. In einem heterogenen Umfeld, wie dies üblicherweise heutzutage an deutschen Hochschulen anzutreffen ist, ist die Nutzung eines Recherche-Clients am wissenschaftlichen Arbeitsplatz ein organisatorisches Problem. Für viele verschiedene Windows-Derivate und Releasestände, Macintosh-PCs und Unix-Workstations müßten unterschiedliche Versionen erstellt, verteilt und gepflegt werden. Jeder Software-Administrator in einem RZ weiß, was dies in der Praxis bedeutet.

„*Java Enterprise Computing*“[5,6] wird zunehmend auch von den führenden Softwarehäusern im Bibliothekenumfeld als ein Meilenstein in der Entwicklung zur digitalen Bibliothek angesehen. Zum einen entfällt das Problem der Software-Verteilung. Ein Benutzer greift auf den entsprechenden Web-Server der Bibliothek zu und erhält seine Software-Umgebung automatisch übermittelt. Die Hardware seines Arbeitsplatzes spielt dabei keine Rolle, sobald ein Web-Browser mit einer Java VM installiert ist. Die Vorteile von Client-Server-Computing werden also um die Plattformunabhängigkeit erweitert. Daß Java auch ein gutes Instrumentarium für sichere Datenübermittlung und sogar Zahlungstransaktionen bietet, zeigt der Einsatz eines Java-Clienten für das Home-Banking bei der Bank 24.

NCs wie das Java-Terminal von Sun bilden als „*Zero-Administration-Client*“[5] ebenfalls eine Plattform für diese Java-Clienten. Für die Aufstellung in Lesesälen bieten sich solche Geräte durch den niedrigen Wartungs- und Pflegeaufwand (keine Festplatte, kein File-System) geradezu an.

Selbst komplexe Datenbankabfragen können mit *JDBC* realisiert werden. Führende Datenbankhersteller wie Oracle, Sybase und Informix stellen in ihren nächsten Versionen bereits JDBC-Schnittstellen für JAVA-Clienten zur Verfügung. Full-Text-Optionen sind ebenfalls heute für die gängigen RDBM-Systeme verfügbar.

Die Basistechnologien im Hardware- und Softwarebereich, die die Basis der digitalen Bibliotheken bilden werden, sind heute schon vorhanden. Doch auch im Softwarebereich gibt es massive Fortschritte. Schon auf der Bibliotheka in Dortmund wird die Fa. Sisis z.B. einen Java-Clienten für ihr Bibliothekssystem anbieten. Führende Softwarehäuser, wie Ameritech/Dynix mit Horizon oder Silverplatter mit ERL, sind ebenso mit der Entwicklung von Java-Clienten beschäftigt, wie die Firma Dataware mit ihrem Bieblis, das in dem heute schon verfügbaren Internet Bibliothekssystem für Zeitschriftenrecherche und Versand (*IBIS/JASON/SUBITO*) eingesetzt wird.

Wenn nun noch eine sinnvolle Arbeitsteilung zwischen den Bibliotheken und den Rechenzentren gefunden wird, sind die Weichen für eine Realisierung von digitalen Bibliotheken in Deutschland eigentlich gestellt. In dieser Hinsicht ist der Forschungsstandort Deutschland – objektiv betrachtet – zwar nicht auf Platz 1 in der Welt, aber doch auf einem sehr guten Weg.

Jörg Schwarz
Sun Microsystems
Geschäftsentwicklung Forschung und Lehre Deutschland

Referenzen:

1 Expertenbefragung Uni Dortmund: http://www.uni-dortmund.de/HRZ/docs/KIT200x/bericht.html
2 Bundesregierung bzw. BMBF: http://www.bmbf.de/inhalt.html, hier insbesondere „Information als Rohstoff für Innovation, Programm der Bundesregierung 1996 - 2000“
3 IBIS/JASON: http://www.ub.uni-bielefeld.de/ibispro.html
4 HIGHWIRE-PRESS: http://www-jbc.stanford.edu.l
5 Java Enterprise Computing, Zero-Administration-Client: http://java.sun.com/
6 Java für Bibliotheken: http://www.sun.com/edu/admin/library/sunlibx/index.html