

Das eVerlage-System: Verwaltung und Bereitstellung kostenpflichtiger hypermedialer Dokumente im Internet

Dietrich Boles, Cornelia Haber, Frank Oldenettel
Informatik-Institut OFFIS
Escherweg 2
D-26121 Oldenburg
email: {boles, haber, oldenettel}@offis.de

Zusammenfassung

eVerlage ist eine Sonderfördermaßnahme im Rahmen des BMBF-Förderkonzeptes Global Info. Elf deutsche Verlage aus Wissenschaft und Technik haben sich zu einer Arbeitsgemeinschaft zusammengeschlossen mit dem Ziel, einen relevanten Teil ihrer Bücher Wissenschaftlern und Studierenden als hypermediale Dokumente über das Internet anzubieten. Partner dieser Arbeitsgemeinschaft sind sieben Forschungseinrichtungen, Hochschulen und Bibliotheken, die die Modellierung, die technische Realisierung und den Betrieb des Online-Angebots übernehmen.

Das eVerlage-System dient den Verlagen als Testumgebung für die Erprobung verschiedener elektronischer Angebotsformen, Abrechnungsmodelle und Zahlungsverfahren für den Zugriff auf die gespeicherten Dokumente. Es ist als verteiltes System realisiert und über Standard-WWW-Browser nutzbar. In diesem Artikel werden neben der Architektur und der Funktionalität des Systems insbesondere die Datenverwaltungsaspekte beschrieben.

1. Einleitung.....	1
2. Anforderungen an das eVerlage-System	2
2.1 Anforderungen aus Sicht der Nutzer	2
2.2 Anforderungen aus der Sicht der Informationsanbieter.....	3
2.3 Anforderungen aus Sicht der Betreiber/Administratoren	3
3. E-Commerce Aspekte.....	3
3.1 Lizenzmodelle.....	3
3.2 Zahlung und Abrechnung	4
4. Architektur.....	4
4.1 Zentralagent	5
4.2 Anbieteragenten.....	5
4.3 Nutzeragenten	5
4.4 Weitere Komponenten	5
5. Dokumentenverwaltung.....	5
5.1 Dokumentenmodell.....	6
5.2 Verwaltung der Volltexte	7
5.3 Anbieteragent.....	7
5.4 Zugriffsschutz auf Dokumente	8
5.5 Administrationswerkzeug	9
6. Andere digitale Bibliotheken.....	10
7. Zusammenfassung und Ausblick.....	10
8. Literatur	11

1. Einleitung

Das vom BMBF aufgelegte Förderkonzept Global Info hat das Ziel, Wissenschaftlern den effizienten Zugang zu den weltweit vorhandenen elektronischen und multimedialen Informationen, die in verteilten Informationssystemen digital gespeichert sind, von ihrem Arbeitsplatzrechner zu eröffnen. Ein besonderes Merkmal von Global Info ist, dass alle am Publikationsprozess Beteiligten, d.h.

Autoren, Verlage, Bibliotheken, Fachinformationszentren und Nutzer bei der Gestaltung der neuen Strukturen zusammenwirken sollen.

Auf Seiten der Anbieter gibt es immer noch große Unsicherheit über die Akzeptanz von kostenpflichtigen elektronischen Verlagsprodukten. Erschwerend kommt hinzu, dass der Wandel vom gedruckten zum elektronischen Fachinformationsprodukt einher gehen muss mit sich ändernden Nutzungsgewohnheiten und -bedürfnissen auf Seiten der Leser.

Aus diesem Grund wurde das Global Info Testbett eVerlage initiiert, ein kommerzieller Service für das Angebot elektronischer, multimedialer Verlagsprodukte im Internet. Elf deutsche Wissenschaftsverlage haben sich zu einer Arbeitsgemeinschaft zusammengeschlossen mit dem Ziel, einen relevanten Teil ihres Bücherbestands in elektronischer Form gemeinsam anzubieten. Partner dieser Verlags-ARGE sind sieben Forschungseinrichtungen, Hochschulen, und Bibliotheken, die die Modellierung, die technische Realisierung und den Betrieb des Testbetts übernehmen. Neben diesen technischen Komponenten spielen kaufmännische Aspekte eine Rolle, wie Marketing, Lizenzierung, Preisbildung, Nutzerverwaltung, Inkasso, Abrechnung und Statistik.

Im Rahmen des Projektes eVerlage wird eine Anwendungsumgebung – das eVerlage-System – geschaffen, die es den Global-Info-Beteiligten erlaubt, neue digitale Angebotsformen, Abrechnungsmodelle und Zahlungsverfahren für elektronische Dokumente zu erproben. Die Aspekte der Akzeptanz bei Nutzern und Anbietern, der Wirtschaftlichkeit und der Durchführbarkeit stehen dabei im Mittelpunkt.

Grundvoraussetzung für eine wirtschaftlich tragfähige elektronische Bibliothek ist der Nachweis hochwertiger Fachliteratur. Dies wird durch die Einbindung namhafter Verlage in das Projekt sicher gestellt. Die Dokumente werden über das eVerlage-System im Volltext angeboten und sind über Standard-WWW-Browser zugreifbar. Besonderer Wert wurde auf flexible Abrechnungs- und Zahlungsverfahren gelegt, um ein Gefühl für akzeptable Lizenzmodelle und realistische Preise zu bekommen. Das eVerlage-System ist offen für alle Nutzer. Durch eine Anbindung des Systems an Bibliothekssysteme wird sicher gestellt, dass auch Nutzer herkömmlicher Bibliotheken an das System herangeführt werden.

Das eVerlage-System basiert auf den Ergebnissen des ebenfalls vom BMBF geförderten Projektes MeDoc [Barth et al. 1998], in dem von 1995 bis 1997 ein erster Prototyp einer kostenpflichtigen digitalen Bibliothek entwickelt wurde.

2. Anforderungen an das eVerlage-System

Eine digitale Bibliothek wie das eVerlage-System hat verschiedene Nutzergruppen, die unterschiedliche Anforderungen stellen. Eine typische Nutzergruppe umfasst die klassischen Nutzer, die über das System Literatur suchen und lesen. Eine zweite Nutzergruppe bilden die Informationsanbieter, die Informationen (Dokumente) in das System einstellen und eine dritte Nutzergruppe stellen die Betreiber des Systems dar, die sich um die technische Wartung kümmern. Im folgenden werden die Anforderungen an das eVerlage-System aus der Sicht der verschiedenen Nutzergruppen beschrieben, die die generelle Funktionalität des inzwischen entwickelten Systems widerspiegeln.

2.1 Anforderungen aus Sicht der Nutzer

Die klassischen Nutzer verwenden das System in erster Linie zur Literaturrecherche. Sie suchen dabei entweder gezielt nach einem bestimmten Dokument oder aber ganz allgemein nach Literatur zu einem gewissen Thema. Eine benutzungsfreundliche Bibliothek muss also sowohl eine Suche auf den bibliographischen Angaben der Dokumente als auch eine Suche in den Volltexten ermöglichen. Dabei sollte die Anfragesprache intuitiv verständlich sein (auch für Erst-Nutzer). Andererseits muss die Anfragesprache mächtig genug sein, um auch komplexe Anfragen formulieren zu können.

Für Fälle, in denen der Nutzer keine genaue Vorstellung davon hat, welches Dokument er sucht bzw. wie eine Suchanfrage am besten zu formulieren ist, ist eine Navigationsmöglichkeit wünschenswert. Mit Hilfe der Navigation kann sich der Nutzer den Dokumentenbestand – nach verschiedenen Kriterien gefiltert – auflisten lassen.

Nachdem der Nutzer ein ihn interessierendes Dokument gefunden hat, möchte er dieses entweder sofort oder zu einem späteren Zeitpunkt lesen können. Der Kauf eines Dokumentes kann dabei durch das Angebot elektronischer Zahlungsverfahren, die Angabe einer Kreditkartennummer oder eine Abbuchung von einem systeminternen Kundenkonto ermöglicht werden. Von zentraler Bedeutung ist, dass Nutzer auf Wunsch an dieser Stelle nicht auf ihre Anonymität verzichten müssen.

Zielgruppe des Systems sind nicht nur Einzelnutzer sondern auch institutionelle Nutzer, wie z.B. der Fachbereich einer Universität oder ein kleines oder mittelständisches Unternehmen. Diese Institutionen müssen Lizenzen für bestimmte Dokumente für ihre Nutzer erwerben können.

2.2 Anforderungen aus der Sicht der Informationsanbieter

Aus der Sicht der Informationsanbieter ist es besonders wichtig, dass möglichst wenig Aufwand beim Verwalten des Dokumentenbestands entsteht, dass also neue Dokumente eingefügt oder veraltete Dokumente entfernt werden können. Ein weiterer wichtiger Aspekt für die Informationsanbieter ist der Schutz der Dokumente. Dazu zählt insbesondere der Schutz vor illegalem Zugriff und der Schutz vor Raubkopien. Weitere Sicherheitsrisiken, die beachtet werden müssen, sind die Übertragung über Netzwerke und die ungenügende Authentifizierung der Nutzer.

2.3 Anforderungen aus Sicht der Betreiber/Administratoren

Die Administratoren erwarten ein stabiles System, das wenig Wartungsaufwand erfordert. Die Installation sollte möglichst einfach sein und das System sollte – soweit möglich – auf Standardkomponenten beruhen. Wichtig ist, dass das System im laufenden Betrieb gewartet werden kann.

3. E-Commerce Aspekte

Prinzipiell kann das eVerlage-System von seiner Funktionalität her als ein Online-Shop im WWW angesehen werden, über den digitale Produkte, nämlich die hypermedialen Dokumente ver- und gekauft werden können. Als Geschäftsmodell wird aktuell der Erwerb von Zeitlizenzen unterstützt. Integrierbar sind jedoch auch Subskription oder Zugriffslizenzen (pay-per-view).

3.1 Lizenzmodelle

Informationsanbieter können ihre Dokumente Nutzern über das eVerlage-System anbieten. Aus Gründen eines besseren Dokumentenschutzes ist kein Download von Dokumenten vorgesehen. Vielmehr können Nutzer Zeitlizenzen für Dokumente, Teile von Dokumenten oder auch Dokumentenbundle (mehrere Dokumente zu einem Paket zusammengefasst) erwerben, über die sie eine gewisse Zeitspanne online über WWW-Browser auf die Dokumente zugreifen können. Für Einzelnutzer werden dabei zwei Typen von Lizenzen unterschieden:

- **Kurzzeitlizenzen:** Will ein Nutzer nur mal kurz in ein Dokument „hineinschnuppern“, kann er eine Kurzzeitlizenz erwerben. Zur Zeit umfassen Kurzzeitlizenzen eine Zeitspanne von einer Stunde und kosten pauschal 1,- DM. Sowohl Zeitspanne als auch Preis sind jedoch variierbar.
- **Einzellizenzen:** Einzellizenzen sind längerfristig gültige Lizenzen mit einer Mindestbestelldauer von aktuell einem halben Jahr. Die Preise betragen einen bestimmten Prozentsatz der Ladenpreise der Bücher.

Für institutionelle Nutzergruppen werden zwei weitere Lizenzmodelle angeboten:

- **Campuslizenzen:** Campuslizenzen erlauben allen Gruppenmitgliedern jederzeitigen Zugriff auf das Dokument. Die Preise richten sich hier nach der Größe der Nutzergruppe.
- **Gleitlizenzen:** Gleitlizenzen sind an die Nutzungsmodelle traditioneller Bibliotheken angelehnt. Besitzt eine Gruppe n Gleitlizenzen für ein Dokument, können n Nutzer gleichzeitig auf das

Dokument zugreifen. Eine Gleitlizenz bleibt dabei einem Nutzer eine bestimmte Zeitspanne (zur Zeit eine Woche) zugeordnet.

3.2 Zahlung und Abrechnung

Lizenzen für den Zugriff auf Dokumente können über das eVerlage-System online erworben werden. Für die Bezahlung werden unterschiedliche Verfahren unterstützt. Für jeden registrierten Benutzer existiert ein privates eVerlage-Konto, das per Überweisung aufgeladen werden kann. Alternativ kann auch mittels elektronischem Geld bezahlt werden. Hier wird zur Zeit auf den Payment-Handler des Chablis-Projektes [Brüggemann-Klein et al. 1998] zugegriffen, der mehrere elektronische Zahlungsverfahren (eCash, IBM MP, Millicent und CyberCash) kapselt. Geplant ist die Integration der Geldkarte und des Pay-Box-Verfahrens, das auf der Basis von Handys arbeitet.

Für institutionelle Nutzergruppen existieren sogenannte Gruppenkonten. Der Gruppenadministrator kann dort einen bestimmten Geldbetrag einzahlen und den Gruppenmitgliedern erlauben, bis zu einem pro Person festgelegten Betrag das Geld für den Erwerb von Kurzzeit- oder Einzellizenzen zu nutzen. Auf diese Art und Weise können bspw. Bibliotheken von Hochschulen sehr flexibel auf konkrete Nutzerbedürfnisse reagieren.

4. Architektur

Das eVerlage-System ist ein verteiltes System, dessen Architektur sich aus den in Abschnitt 2 skizzierten Nutzergruppen ableitet (siehe auch Abbildung 1). Sogenannte Nutzeragenten ermöglichen Nutzern den Zugang zum System. Über Anbieteragenten können Anbieter ihre Dokumente ins System integrieren. Ein Zentralagent dient als Vermittler zwischen Nutzern und Anbietern und realisiert die E-Commerce-Aspekte. Das gesamte eVerlage-System ist in Java implementiert. Die Agenten sind eigenständige Prozesse, die auf der Basis von Java-RMI kommunizieren. Zum Schutz gegen das illegale „Abhören“ der Kommunikationsverbindungen zwischen den einzelnen Komponenten des eVerlage-Systems kommt das *Secure-Socket-Layer-Protokoll* (SSL) zum Einsatz, welches die Datenströme verschlüsselt und somit einen ausreichenden Schutz während der Datenübertragung bietet.

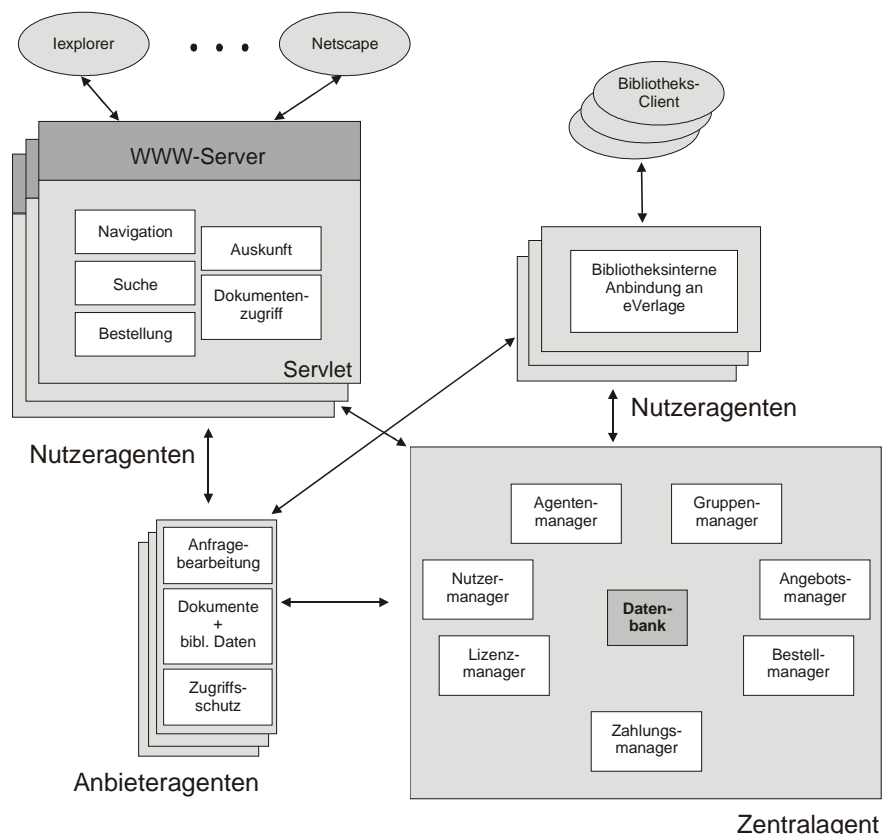


Abbildung 1: Architektur des eVerlage-Systems

4.1 Zentralagent

Der Zentralagent des eVerlage-Systems besteht aus insgesamt sieben verschiedenen Komponenten, die bestimmte Funktionalitäten kapseln:

- Der Agentenmanager verwaltet die angeschlossenen Nutzer- und Anbieteragenten.
- Der Nutzermanager stellt Funktionen zum Speichern und Laden von Nutzerdaten bereit.
- Der Gruppenmanager ist für die Organisation von institutionellen Nutzergruppen zuständig.
- Der Angebotsmanager kapselt die Angebote.
- Der Lizenzmanager verwaltet die erworbenen Lizenzen.
- Der Zahlungsmanager ist für den Ablauf von Zahlungsvorgängen und die Verwaltung der Konten verantwortlich.
- Der Bestellmanager kümmert sich im Zusammenspiel mit den anderen Managern um die Abarbeitung kompletter Bestellvorgänge.

Alle Daten werden in einer relationalen Datenbank gespeichert, auf die mittels JDBC zugegriffen wird.

4.2 Anbieteragenten

Anbieteragenten kapseln den Dokumentenbestand sowie die bibliographischen Daten eines oder mehrerer Informationsanbieter. Sie sind für die Bearbeitung von Suchanfragen zuständig und haben in Kooperation mit dem Lizenzmanager des Zentralagenten dafür Sorge zu tragen, dass keine unberechtigten Zugriffe auf die Dokumente erfolgen. Die Anbieteragenten werden in Abschnitt 5 noch detaillierter beschrieben.

4.3 Nutzeragenten

Nutzeragenten, von denen mehrere installiert sein können, ermöglichen Nutzern den Zugang zum System. Ihre Hauptaufgaben bestehen aus der Kapselung von Zugangscharakteristiken und dem Cachen von bestimmten Daten, um einen performanten Zugriff auf das System gewährleisten zu können. Es existieren verschiedene Typen von Nutzeragenten. Der Standard-Nutzeragent ist als Java-Servlet realisiert und über Standard-WWW-Browser zugänglich. Er realisiert insbesondere Such-, Navigations-, Auskunfts-, Bestell- und Dokumentenzugriffsfunktionalitäten. Für die Bibliotheken, die das eVerlage-System an ihre Systeme anbinden wollen, werden eigene Nutzeragenten entwickelt, die die Anbindung realisieren.

4.4 Weitere Komponenten

Für eine einfache Administration des eVerlage-Systems sind weitere Software-Werkzeuge mit graphisch-interaktiven Front-Ends entwickelt worden:

- ein Tool für das Einstellen und Verwalten von Dokumenten (siehe auch Abschnitt 5.5),
- ein Tool für die Erstellung von Angeboten,
- ein Tool für die Erstellung und Auswertung von Nutzungsstatistiken,
- ein Tool für Administratoren institutioneller Nutzergruppen, das die Verwaltung von Gruppenmitgliedern und Gruppenkonten sowie den Erwerb von Gruppenlizenzen ermöglicht und
- ein Tool für eine einfache Pflege der gesamten Datenbank.

5. Dokumentenverwaltung

Bei einer digitalen Bibliothek wie dem eVerlage-System wird der Dokumentenverwaltung eine große Bedeutung beigemessen. Sie muss eine schnelle, flexible und zuverlässige Recherche sowie den unmittelbaren Dokumentenzugriff ermöglichen. Weitere wichtige Aspekte sind die Integration von Mechanismen für den Dokumentenschutz sowie eine einfache Wartbarkeit des Dokumentenbestandes.

Im eVerlage-System ist der Anbieteragent für die Verwaltung der Dokumente verantwortlich. Informationen zu den Dokumenten, wie die bibliographischen Angaben, aber auch die hierarchische Struktur eines Dokumentes werden im sogenannten Dokumentenmodell festgehalten.

5.1 Dokumentenmodell

Das dem eVerlage-System zugrunde liegende Dokumentenmodell wurde größtenteils aus dem eingangs erwähnten MeDoc-Projekt übernommen und entsprechend der Anforderungen an das neue eVerlage-System modifiziert. Ziel bei der Konzeption des Dokumentenmodells war es, sowohl die Struktur innerhalb eines Dokumentes (Buch → Kapitel → Abschnitte) als auch die Beziehung der Dokumente zueinander (z.B. mehrere Ausgaben einer Zeitschrift, die Bücher einer Serie, ...) abzubilden und eine gezielte Suche innerhalb der bibliographischen Dokumentendaten zu ermöglichen.

Abbildung 2 illustriert das Dokumentenmodell des eVerlage-Systems. Auf den obersten Hierarchieebenen befinden sich *Zeitschriften* und *Serien*. Die Zeitschriften gliedern sich in Bände, welche sich wiederum in Ausgaben unterteilen. Jede einzelne Ausgabe einer Zeitschrift besteht aus Artikeln, die sich aus Abschnitten zusammensetzen. Diese Untergliederung ermöglicht es einem Nutzer des Systems, beispielsweise von einem Artikel einer Zeitschrift zum nächsten Artikel derselben Ausgabe zu navigieren, oder aber zu einem Artikel einer vorangegangenen Ausgabe zurückzuspringen.

Serien können aus Artikelsammlungen, Tagungsbänden, Büchern sowie sonstigen Dokumenten bestehen. Die Artikelsammlungen und die Tagungsbände unterteilen sich wie Zeitschriften ebenfalls in Artikel. Bücher hingegen setzen sich aus Kapiteln zusammen, die sich in Abschnitte gliedern. Für sonstige Dokumente, also Dokumente, die nicht einer der gerade beschriebenen Kategorien zugeordnet werden können, wurde dieselbe Strukturierung wie für Bücher gewählt.

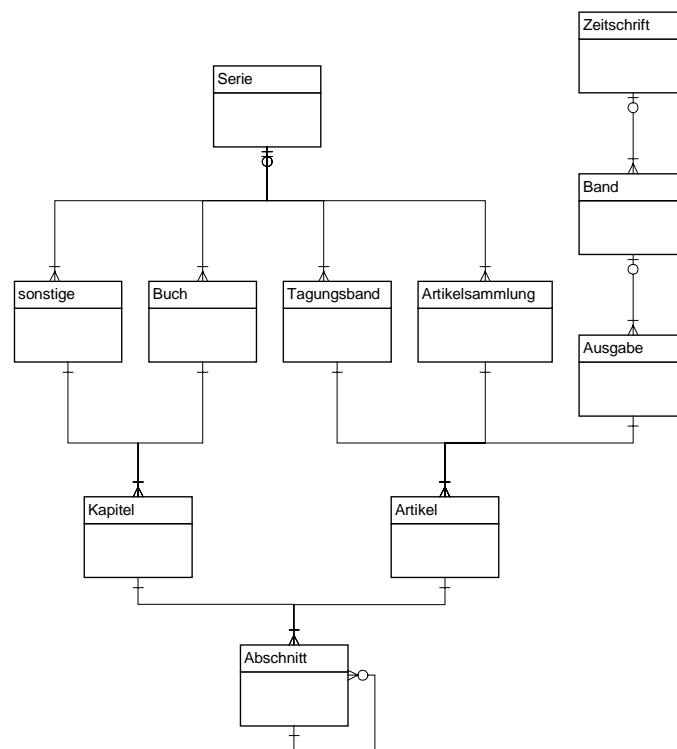


Abbildung 2: Dokumentenmodell

Zu jedem der im Dokumentenmodell aufgeführten Dokumenttypen werden die entsprechenden bibliographischen Angaben wie Titel, Autor, Verlag, etc. (angelehnt an die BibTeX-Attribute) gespeichert.

Das eVerlage-System sieht vor, dass auf jeder Hierarchieebene im Dokumentenmodell Lizenzen definiert werden können, d.h. über das System können sowohl ganze Zeitschriften als auch Bände, Ausgaben, einzelne Artikel oder Abschnitte zum Kauf angeboten werden. Dass Abschnitte rekursiv wiederum aus Abschnitten bestehen, erlaubt eine noch feinere Lizenzierung, beispielsweise für einzelne Definitionen in einem Lexikon.

Für das eVerlage-System wurde das Dokumentenmodell auf ein relationales Datenbankschema abgebildet. Zusätzlich zu den durch das Dokumentenmodell vorgegebenen Daten enthält die relationale Datenbank Informationen zur physikalischen Struktur eines Dokumentes (Zuordnung von Dateien zu einem Dokument). Der Zugriff auf den Dokumentenbestand wird vom sogenannten Anbieteragenten gesteuert, der in Abschnitt 5.3 beschrieben wird.

5.2 Verwaltung der Volltexte

Neben der Möglichkeit, in den bibliographischen Daten des Dokumentenbestandes zu recherchieren, bietet das eVerlage-System ebenfalls die Recherche in den Volltexten der Dokumente. Zu diesem Zweck wird zusätzlich zu den bibliographischen Informationen der komplette Volltext eines jeden Dokuments in das angeschlossene Datenbanksystem eingespielt.

Zur Verwaltung der Volltextinformationen ist im Datenbankschema lediglich eine weitere Relation notwendig, die zum einen die Volltexte indiziert und zum anderen den Bezug zu den zugehörigen bibliographischen Daten herstellt. Dies geschieht in Form einer Fremdschlüsselbeziehung auf Basis der DokumentenID, die die bibliographischen Daten eines Teildokumentes auf jeder Hierarchiestufe des oben beschriebenen Dokumentenmodells eindeutig identifiziert. Die Initiierung einer Volltextanfrage sowie die Kombination der Ergebnisse mit den bibliographischen Daten werden vom Anbieteragenten übernommen, der im folgenden Abschnitt beschrieben wird.

5.3 Anbieteragent

Der Anbieteragent ist die dokumentenverwaltende und dokumentenliefernde Komponente des eVerlage-Systems. Seine wichtigste Aufgabe besteht darin, das Datenbanksystem zur Speicherung des Dokumentenbestandes vom übrigen System zu kapseln. Anfragen, die Dokumente betreffen, werden vom Nutzeragenten an den Anbieteragenten geschickt, ausgewertet und die Ergebnisse wieder zurückgeliefert. Die Architektur des Anbieteragenten wird in Abbildung 3 skizziert.

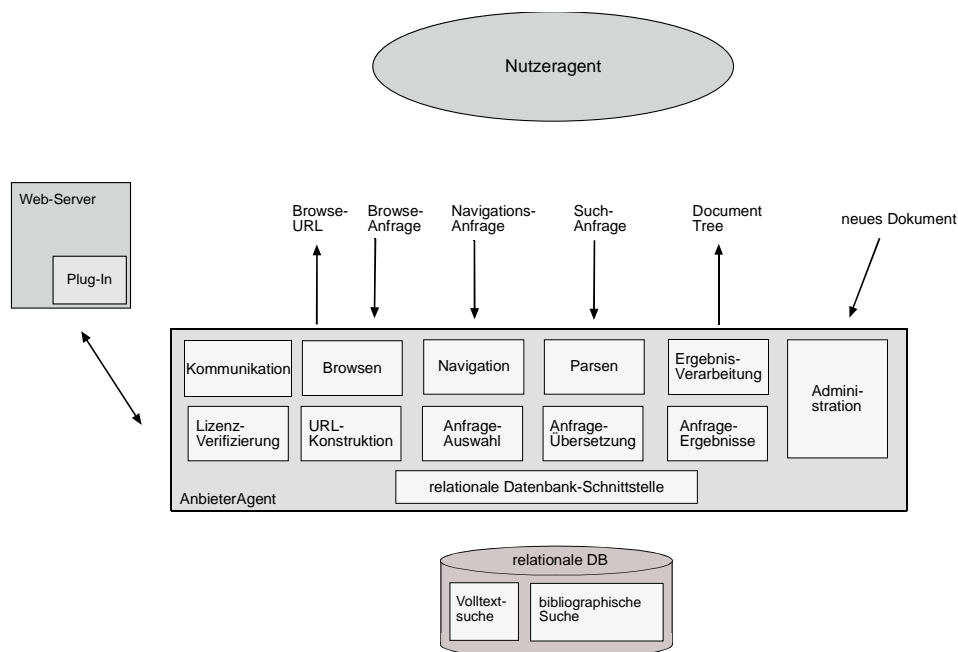


Abbildung 3: Architektur des Anbieteragenten

Die Kapselung des Datenbanksystems wird im wesentlichen von der *DB-Schnittstelle* geleistet. Da die Datenbankanbindung über JDBC erfolgt, ist das zugrunde liegende Datenbankmanagementsystem austauschbar. Anfragen nach Dokumenten (*Such-Anfrage*) werden vom Nutzeragenten an den Anbieteragenten geschickt. Der Anbieteragent übersetzt die in der eVerlage-Anfragesprache vorliegende Suchanfrage in die spezifische Anfragesprache des angeschlossenen Datenbanksystems

(Anfrage-Übersetzung). Nach Abarbeitung der Anfrage durch das Datenbanksystem überführt der Anbieteragent die mengenorientierten Anfrageergebnisse in hierarchische Dokumentenstrukturen (Ergebnis-Verarbeitung), die sogenannten *Document-Trees*. Diese liefert er dann für jedes Dokument, welches die Anfragebedingungen erfüllt, zurück an den Nutzeragenten. Ein Document-Tree rekonstruiert gewissermaßen das in Abschnitt 5.1 vorgestellte Dokumentenmodell.

Neben verschiedenen Suchmöglichkeiten in den bibliographischen Daten des Dokumentenbestandes bietet der Anbieteragent ebenfalls die Recherche in den Volltexten der Dokumente.

Empfängt der Anbieteragent eine aus Volltextsuche und Suche über bibliographische Attribute kombinierte Suchanfrage wird parallel zum Suchprozess in den bibliographischen Daten ein Suchprozess im Volltextdatenbestand angestoßen. Anschließend werden die Anfrageergebnisse gemäß den Verknüpfungsoperatoren in Beziehung gesetzt und das daraus resultierende Ergebnis zur Rückgabe an den Nutzeragenten in der oben beschriebenen Art und Weise aufbereitet.

Eine weitere Aufgabe des Anbieteragenten besteht darin, die in Abschnitt 2.1 beschriebene Navigationsfunktionalität zur Verfügung zu stellen. Hierbei wird zwischen expliziter und impliziter Navigation unterschieden. Die *explizite Navigation* durchläuft den Dokumentenbestand anhand bestimmter Klassifikationshierarchien, an deren Enden sich jeweils die Dokumente befinden. Solch eine explizite Navigation beispielweise nach Autoren startet bei den Buchstaben des Alphabets, gelangt dann bei Auswahl eines Buchstaben zu den entsprechenden Autorennamen und endet nach Auswahl eines Namens bei den Werken des betreffenden Autors. Die *implizite Navigation* kann nur in Verbindung mit bereits vorliegenden Anfrageergebnissen durchgeführt werden. Sie bewirkt eine Auflistung aller Dokumente, die mit dem Navigationskriterium übereinstimmen, z.B. alle Werke eines bestimmten Autors. Angestoßen wird die implizite Navigation auf Seiten des Nutzeragenten, der das Navigationskriterium in eine Anfrage umsetzt (*Anfrage-Auswahl*).

Sowohl für Rechercheanfragen als auch für Navigationsanfragen gilt, dass der zurückgelieferte Document-Tree nur bibliographische Daten der entsprechenden Dokumente enthält. Für den Zugriff auf die Inhalte von Dokumenten müssen diese explizit vom Nutzer ausgewählt werden. Der Anbieteragent erhält dann entsprechende Aufrufe (*Browse-Anfrage*), überprüft, ob der Nutzer eine gültige Lizenz für das entsprechende Dokument besitzt, und ermittelt dann den Verzeichnispfad, in dem die zu dem Dokument gehörenden Dateien abgelegt sind. Ist der Nutzer zugriffsberechtigt, so wird aus diesem Verzeichnispfad eine sogenannte *Browse-URL* (siehe nächster Abschnitt) konstruiert (*URL-Konstruktion*), über die der Nutzer mit seinem WWW-Browser auf das Dokument zugreifen kann.

5.4 Zugriffsschutz auf Dokumente

Wie bereits in Abschnitt 3 dargelegt wurde, enthält das eVerlage-System überwiegend kostenpflichtige Dokumente, für deren Nutzung Lizenzen erworben werden müssen. Deshalb ist ein Schutz der Dokumente gegen einen illegalen Zugriff von großer Wichtigkeit. Zu diesem Zweck werden eine Reihe von Sicherheitsmechanismen implementiert, die einen Zugriffsschutz gewährleisten und ein illegales Downloaden von Dokumenten verhindern sollen.

Einen Schutzmechanismus gegen illegalen Zugriff stellt die bereits in Abschnitt 5.3 erwähnte Browse-URL dar. Diese unterscheidet sich von einer gewöhnlichen URL dadurch, dass sie einen Schlüssel enthält, der nur einem berechtigten Nutzer den Zugriff auf das Dokument ermöglicht. Die Dekodierung des Schlüssels erfolgt über ein Plug-In innerhalb des Web-Servers, der für die Dokumentenlieferung zuständig ist. Das Plug-In leitet die URL an den Anbieteragenten weiter, der den Schlüssel überprüft (*Lizenz-Verifizierung*). Dieses Verfahren greift jedoch nicht, wenn ein Nutzer seine Browse-URL an andere Nutzer weitergibt. Um den Schaden, der durch das Weitergeben der URL mit Schlüssel verursacht wird, möglichst gering zu halten, besitzt der Schlüssel nur eine kurze Gültigkeitsdauer. Ist diese Zeit abgelaufen, muss sich der Nutzer durch die Eingabe von Login und Passwort erneut identifizieren und bekommt dann automatisch einen neuen Schlüssel zugeteilt.

Neben programmiertechnischen Maßnahmen im Anbieteragenten und im Web-Server besteht auch die Möglichkeit, „Hürden“ in die Dokumente selbst zu integrieren, um einen illegalen Download durch entsprechende Tools zumindest zu erschweren. Eine solche „Hürde“ stellen z.B. Links dar, die für die Nutzer unsichtbar sind, ein Download-Tool jedoch verwirren und zum Abbruch des Downloadvorgangs führen. Dies können z.B. rekursive Links sein, die den Download-Vorgang in Zyklen festhalten, oder Links zu Dateien, die sehr große Mengen an „Datenmüll“ beinhalten, oder Links, die der Anbieteragent als illegal identifiziert und unmittelbar jeden weiteren Zugriff sperrt.

5.5 Administrationswerkzeug

Neben den überwiegend funktionalen Aufgaben des Anbieteragenten, die in den vorangegangenen Abschnitten beschrieben wurden, sind auch die administrativen Aufgaben nicht zu vernachlässigen. Hierbei ist in erster Linie das Einspielen neuer Dokumente zu nennen. Diese Tätigkeit ist mitunter sehr komplex, da die einzuspielenden Dokumente häufig aus einer Vielzahl von Dateien bestehen. Für einen zuverlässigen Zugriff auf die Dokumente ist es von großer Wichtigkeit, dass sowohl die bibliographischen Daten als auch die zu den einzelnen Abschnitten gehörenden Dateien in korrekten Beziehungen zueinander in die Datenbanken eingespielt werden. Erleichtert wird diese Aufgabe durch ein Werkzeug (*Administration-Modul*), welches bequem und schnell die Aufteilung des Dokuments in lizenzierbare Abschnitte, die Eingabe der bibliographischen Daten sowie die Zuordnung der entsprechenden Dateien unterstützt. Abbildung 4 zeigt einen Snapshot des Werkzeug-Prototypen.

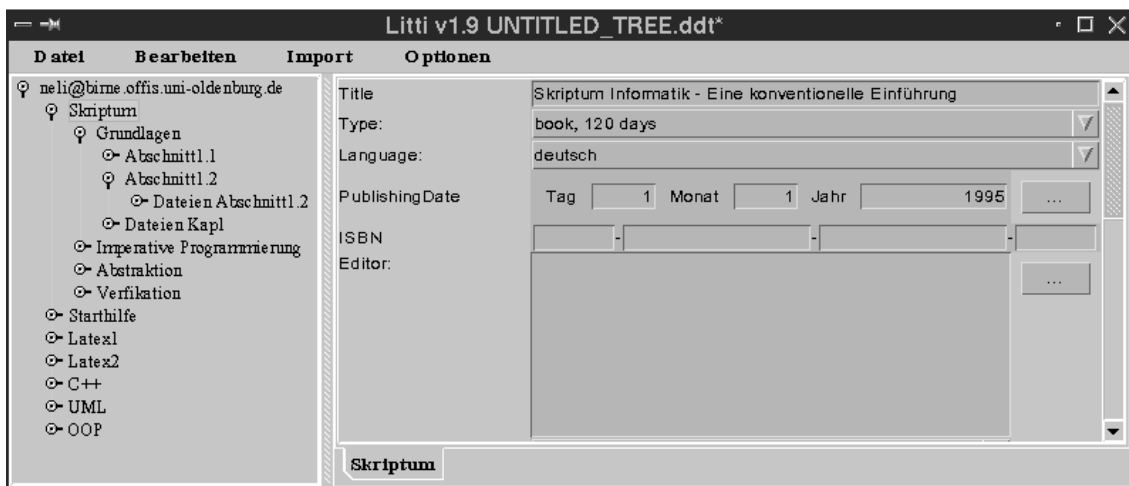


Abbildung 4: Prototyp des Administrationswerkzeugs für den Dokumentenbestand

Die Benutzungsschnittstelle weist im wesentlichen zwei Bereiche auf. Der linke Bereich stellt eine Art Navigationsfenster dar, welches alle Dokumente des Datenbestandes eines Anbieteragenten mit ihren Teildokumenten wie in einem Dateisystem anzeigt. Jedes Dokument wird in Form eines Ordners dargestellt, der die oberste Hierarchiestufe für ein Dokument bildet. Als Name des Ordners wird ein Kurztitel des Dokuments gewählt. Innerhalb eines Ordners lassen sich beliebig viele weitere Unterordner anlegen. Mit dem Ordnerart Teildokument können die Kapitel bzw. Artikel oder die Abschnitte eines Dokuments beschrieben werden. Der Ordnerart Dateireferenz dient der Zuordnung von Dateien zu einem Teildokument.

Der rechte Bereich der Benutzungsschnittstelle dient der Eingabe von Inhalten. Die Eingabefelder variieren in Abhängigkeit vom gewählten Ordnerart auf der linken Seite. Abbildung 4 zeigt den ausgewählten Ordnerart der obersten Hierarchiestufe ("Skriptum") und die entsprechenden Eingabefelder für bibliographische Daten wie Titel (title), Herausgeber (editor) oder Erscheinungsdatum (publishing date). Des Weiteren werden Metainformationen wie z.B. Dokumententyp (type) oder Sprache (language) festgelegt. Bei der Auswahl des Ordnerart Teildokument sind die Eingabefelder ähnlich, jedoch wird auf diejenigen verzichtet, die auf unteren Hierarchiestufen redundant wären, wie z.B. Herausgeber oder Verlag. Der Ordnerart Dateireferenz wird auf der rechten Seite durch komfortable Datei-Auswahl-Dialoge unterstützt.

Nachdem der Benutzer alle neue Daten zum Einspielen eines Dokumentes eingegeben oder ggf. Änderungen an bestehenden Datensätzen durchgeführt hat, kann er diese auch in der angeschlossenen Datenbank wirksam machen. Dies kann auf zwei Arten geschehen: Entweder direkt über eine JDBC-Verbindung oder indirekt durch ein SQL-Skript, welches exportiert werden kann und bei Ausführung in der Datenbankanwendung die Änderungen vornimmt. Die zweite Variante gibt dem Benutzer eine größere Flexibilität, da er die Möglichkeit hat, das SQL-Skript nachträglich zu editieren, um eventuell Aspekte umzusetzen, die mit der starren Vorgabe der Benutzungsschnittstelle nicht realisiert werden können.

6. Andere digitale Bibliotheken

Andere kostenpflichtige digitale Bibliotheken, die über das Internet nutzbar sind, die jedoch sehr unterschiedliche Architekturen besitzen, verschiedene Nutzerkreise ansprechen und meist nur eine sehr eingeschränkte Auswahl an Lizenzmodellen bieten, sind bspw. die Bibliotheken der ACM (<http://www.acm.org/dl/>), der IEEE Computer Society (<http://computer.org/publications/dlib/>) und Springer-LINK (<http://link.springer.de>). Die Bibliothek der ACM [ACM 1999a], [ACM 1999b] beinhaltete im September 1999 bereits mehr als 31000 Artikel aus den Zeitschriften und Konferenzbänden der ACM. Die Computer Society [Computer 2000a], [Computer 2000b] bietet 18 Zeitschriften und über 350 Konferenzbände online an. Beide Systeme bieten eine Suche über bibliographische Daten (Titel, Abstract, Autor, ...) sowie eine Volltextsuche an. Die Dokumente liegen jeweils im Format PDF vor. Die Suche ist kostenlos, erst beim Zugriff auf den Volltext eines Dokumentes wird der Nutzer nach Login und Passwort gefragt. Als Lizenzmodell unterstützt ACM pay-per-view (pro Artikel) und Subskription. Die IEEE Computer Society unterstützt ebenfalls die Subskription des gesamten Dokumentenbestandes oder aber die Subskription ausgewählter Zeitschriften. LINK, die digitale Bibliothek der Springer-Verlagsgruppe, ist von der Funktionalität ungefähr vergleichbar mit den beiden vorher aufgeführten Systemen, allerdings beinhaltet LINK auch Dokumente aus anderen Fachbereichen als der Informatik.

Bei der grauen Literatur haben sich andere Ansätze durchgesetzt. Graue Literatur wird meist kostenlos über das WWW angeboten. An dieser Stelle sollen nur zwei Systeme kurz gegenüber gestellt werden: NCSTRL (<http://www.ncstrl.org>) und NZDL(<http://www.nzdl.org>). NCSTRL [Leiner 1998] ist ein verteiltes System interagierender Sites. Jede Site betreibt einen eigenen Server, der die Suche im gesamten Dokumentenbestand ermöglicht. Die Dokumente selbst liegen im Format PostScript im Dateisystem des Servers, es lassen sich jedoch auch andere Dokumentenformate integrieren. Die bibliographischen Daten zu einem Dokument liegen in einer separaten Datei im Filesystem entsprechend dem RFC 1807. Von dort aus werden sie entweder mit FreeWAIS oder mit einem in NCSTRL integrierten Indexer indiziert und sind damit suchbar. Unterstützt wird eine Suche in den bibliographischen Angaben (Autor, Titel, Abstract, ...). Einen anderen Ansatz verfolgt das New Zealand Digital Library Project mit seiner digitalen Bibliothek (NZDL). NZDL [Witten et al. 1996] sucht im WWW nach Technical Reports im PostScript-Format. Dabei wird von Standard-Sites mit Listen von Technical-Report-Servern ausgegangen, einschlägige NewsGroups werden durchsucht oder eine neue Site wird dem NZDL-Koordinator mitgeteilt. Hat die NZDL-Software einen Bericht gefunden, so wird aus dem PostScript-Text ein ASCII-Text und ein Bild der ersten beiden Textseiten erstellt und an den zentralen Server in Neuseeland geschickt. Dort wird der ASCII-Text indiziert. Eine Suche in NZDL entspricht dann einer Volltextsuche auf den ASCII-Texten. Die Suche nach bibliographischen Angaben wird approximiert (z.B. die Suche nach einem Autor entspricht der Suche nach dem Autorennamen auf der ersten Seite eines Dokumentes). Will der Nutzer auf den Volltext eines Dokumentes zugreifen, so wird er auf den Server verwiesen, auf dem das Dokument ursprünglich gefunden wurde.

7. Zusammenfassung und Ausblick

Das eVerlage-System realisiert eine verteilte digitale Bibliothek über die kostenpflichtige hypermediale Dokumente angeboten werden können. Es ist als flexibles anpassbares Testbett realisiert, das dazu dient, dass alle am Publikationsprozess beteiligten Nutzergruppen Erfahrungen im Bereich des Information Commerce, d.h. des Handels mit digitalen Produkten bzw. Informationen über das Internet, sammeln können.

Der erste Prototyp des eVerlage-Systems wurde im Frühjahr 2000 fertig gestellt und auf der Leipziger Buchmesse präsentiert. Erste Erfahrungen der Nutzer waren recht positiv. Es ist geplant, noch im Laufe des Jahres 2000 ein funktionsfähiges System für die Öffentlichkeit freizuschalten. Ob eine solche kostenpflichtige digitale Bibliothek ausreichend Akzeptanz bei den Nutzern findet, wird allerdings auch davon abhängen, wie sich elektronische Zahlungsverfahren in der Zukunft durchsetzen werden.

8. Literatur

[Barth et al. 1998] Barth, A.; Breu, M.; Endres, A.; de Kemp, A.: Digital Libraries in Computer Science: The MeDoc Approach. In Lecture Notes in Computer Science, Vol. 1392, Springer-Verlag, Heidelberg, 1998.

[Brüggemann-Klein et al. 1998]Brüggemann-Klein, A.; Endres, A.; Jessen, E.; Weber, R.; Werner, H.: Projekt Chablis - Abrechnungs- und Zahlungskonzepte für Dienstleistungen digitaler Bibliotheken. In: Informatik Forschung und Entwicklung, 1998, Band 13, Heft 3, S. 169 - 172

[ACM 1999a] ACM Digital Library FAQ , 1999. <http://www.acm.org/dl/faq.html>

[ACM 1999b] ACM Digital Library Content and Organization, 1999. <http://www.acm.org/dl/org.html>

[Computer 2000a] About the Computer Society Digital Library, 2000.
<http://computer.org/publications/dlib/about.htm>

[Computer 2000b] MDLS: Member Digital Library Subscription, 2000.
<http://www.computer.org/subscribe/MDLS.htm>

[Leiner 1998] Leiner, B.: The NCSTRL Approach to Open Architecture for the Confederated Digital Library. In: D-Lib Magazine, 1998.

[Witten et al. 1996] Witten, I.H.; Cunningham, S.J.; Apperley, M.D.: The New Zealand Digital Library Project. In: D-Lib Magazine, 1996.

Stichwörter: Digitale Bibliothek, hypermediale Dokumente, Dokumentenverwaltung, Information Commerce, elektronische Zahlungsverfahren, eVerlage