Testimplementation einer Audiodatenbank auf dem Hyperwave Information Server

Diplomarbeit

an der

Technischen Universität Graz

vorgelegt von

Wolfgang Götzinger

Institut für Informationsverarbeitung und Computergestützte neue Medien (IICM)
Technische Universität Graz
A-8010 Graz

Graz, im Juni 1998

© Wolfgang Götzinger

Begutachter: o.Univ.-Prof. Dr. Dr.h.c. Hermann Maurer
Betreuer: Dipl. Ing. Christian Gütl

Zusammenfassung

Im Rahmen dieser Diplomarbeit wird eine Audiodatenbank auf dem Hyperwave Information Server implementiert. Der Hyperwave Information Server ist ein Internet/Intranet Dokumentenverwaltungssystem der zweiten Generation und unterscheidet sich von herkömmlichen Web Servern durch eine Reihe von Konzepten, wie z.B. Volltextsuche, Strukturelemente, Linkkonsistenz und Usermanagement. Bei der Implementation der Audiodatenbank ermöglichen diese Konzepte die Realisierung von Funktionen, die dem Anwender vielfältige Möglichkeiten bei der Nutzung der Audiodatenbank bieten; so kann unter anderem mit Hilfe von Metadatensets und Audiolinks ein rascher Überblick über den Datenbestand gewonnen werden. Der theoretische Teil der Arbeit beschreibt die derzeitigen Möglichkeiten zur Verwendung von Audio im Internet, gibt einen Überblick über den Einsatz von Metadaten und behandelt die digitale Audiotechnik im Hinblick auf Datenreduktion. Weiters wird der Realisierungsprozeß der Audiodatenbank unter Hyperwave beschrieben. Dabei werden im Vorfeld die Konzepte existierender Audiodatenbanken bewertet, eigene Ideen eingebracht und aus den daraus entstehenden möglichen Lösungsvarianten ein Realisierungsansatz gebildet. Der Abschluß der Arbeit beschreibt die Funktionsweise der Testimplementation.

Abstract

This Master's thesis describes the implementation of an audio database for the Hyperwave Information Server. The Hyperwave Information Server is a next generation Internet/Intranet document management system and differs from common web servers by many concepts, e.g. full text search, link consistency and user management. These concepts enable the realisation of features, which bring many facilities to the user when accessing the audio database, among other things a survey of the data is given with the help of metadata and audio links. The theoretical part of this thesis describes the possibilities of using audio on the Internet, gives a survey of the use of metadata and describes digital audio techniques with regard to data reduction. The process of realising the audio database on Hyperwave is also described. In doing so the concepts of existing audio databases were evaluated, further ideas concieved and a realisation concept built. The second part of the thesis describes the implementation in detail.

Inhaltsverzeichnis

1. Einleitung 1
2. Das Internet 2
   2.1 Allgemeines 2
   2.2 Geschichte des Internet 2
   2.3 Dienste im Internet 4
      2.3.1 E-Mail 4
      2.3.2 Usenet / Newsgroups 4
      2.3.3 File Transfer Protocol (FTP) 4
      2.3.4 Telnet 5
      2.3.5 Hypertext und Hypermedia im Internet 5
      2.3.6 Multimedia im Internet 6
      2.3.7 Internet Browser 6
      2.3.8 Audio im WWW 7
   2.4 Hyperwave Information Server 8
3. Metainformationen 11
   3.1 Metainformationen im Internet 11
   3.2 Überblick über verschiedenartige Metadaten 12
   3.3 Metadaten Elemente-Sets 13
      3.3.1 Dublin Core 13
      3.3.2 Name und Beschreibung der Elemente 14
      3.3.3 Nutzung von Dublin Core 17
      3.3.4 Dublin Core für Audiodaten 17
      3.3.5 Beispiel eines Dublin Core Elemente-Set 18
4. Digitale Audiosignale 19
   4.1 Informationsumsetzung 19
   4.2 Digitalisierung 19
      4.2.1 Abtastung 20
      4.2.2 Quantisierung 20
      4.2.3 Kodierung 21
      4.2.4 Datenraten gebräuchlicher digitaler Audiosysteme 21
      4.2.5 Digitale Datenformate 21
      4.2.5.1 AIFF (Audio Interchange File Format) 22
      4.2.5.2 MIDI (Musical Instrument Digital Interface) 22
      4.2.5.3 WAVE (Waveform Audio File Format) 22
      4.2.5.4 SND 23
   4.3 Datenreduktion 23
      4.3.1 Hörphysiologie 24
      4.3.1.1 Verdeckung 25
      A) Simultane Verdeckung 25
      B) Temporäre Verdeckung 26
      4.3.2 Datenraten gebräuchlicher datenreduziereter Audiosysteme 27
      4.3.3 Datenreduzierte Datenformate 27
      4.3.3.1 MPEG (Moving Picture Experts Group) 27
      4.3.3.2 Real Audio 29
      4.3.3.3 ASF (Advanced Streaming Format) 29

5. Konzepte von Audiodatenbanken 31
5.1 Music Boulevard [musicblvd.com] 31
5.2 IRCAM Soundbase [ircam.fr] 33
5.3 Subaudio [subaudio.net] 34
5.4 Universität Utrecht Audiodatabase [wwwots.let.ruu] 36
5.5 Performing Arts Data Service [arts.gla] 37
5.6 Zusammenfassung 37
6. Realisierungsprozeß einer eigenen Audiodatenbank unter Hyperwave 38
6.1 Bewertung der vorhandenen Audiodatenbanken 39
6.1.1 Music Boulevard 40
6.1.1.1 Bewertung Music Boulevard 40
6.1.2 IRCAM Soundbase 40
6.1.2.1 Bewertung IRCAM Soundbase 41
6.1.3 Subaudio 41
6.1.3.1 Bewertung Subaudio 41
6.1.4 Universität Utrecht Audiodatabase 41
6.1.4.1 Bewertung Universität Utrecht Audiodatabase 42
6.1.5 Zusammenfassung 42
6.1.6 Lösungsvorschlag 42
6.2 Einbringung eigener Ideen 43
6.2.1 Beispiele für Metadatensets 43
6.2.2 Konfigurierbare Metadatensets 44
6.2.2.1 Nicht konfigurierbar 44
6.2.2.2 Teilweise konfigurierbar 45
6.2.2.3 Voll konfigurierbar 45
6.2.3 Links 46
6.2.4 Audiolinks 47
6.2.4.1 Audiolinks auf eine Audioresource 47
6.2.4.2 Audiolinks von einer Audioresource 48
6.2.5 Verkettung von Audioresourcen 48
6.3 Lösungsvarianten für die Realisierung einer eigenen Audiodatenbank 48
6.4 Realisierungsansatz 49
6.4.1 Prüfung auf Realisierbarkeit 50
6.4.1.1 Zu realisierende Punkte 50
6.4.1.2 Realisierbarkeit unter Hyperwave 50
6.4.1.3 Aufwandsabschätzung 51
7. Testimplementation 53
7.1 Nutzung der Testimplementation 53
7.1.1 Übersichtsdarstellung 53
7.1.2 Suche 54
7.1.3 Infofenster 56
7.1.3.1 Infoansicht 56
7.1.3.2 Textansicht 58
7.1.4 Audioplayer 59
7.2 Administration der Testimplementation 59
7.2.1 Aufbau der Datenbankhierarchie 59
7.2.2. Einfügen einer Aufnahme 60
7.2.3 Editieren des Metadatensets 63
7.2.4 Editieren des Metaattributes „Text“ 64
7.2.5 Einfügen eines Audiolinks 65

7.2.6 Löschen eines Audiolinks 66
7.3. Funktionsweise der Testimplementation 67
8. Zusammenfassung und Ausblick 70
8.1 Zusammenfassung 70
8.2 Ausblick 70
Literaturverzeichnis 72
Abbildungsverzeichnis 74
Tabellenverzeichnis 75

1. Einleitung

Der Hyperwave Information Server ist ein Internet/Intranet Dokumentenverwaltungssystem der zweiten Generation. Gegenüber herkömmlichen Web Servern unterscheidet sich der Hyperwave Server unter anderem durch folgende Konzepte:
Objektorientierte Datenbank: Informationen werden in einer integrierten, objektorientierten Datenbank gespeichert. Objekttypen sind z.B. Dokumente, Links, Collections, Cluster oder Sequences.
Strukturierung: Dokumente können innerhalb von Collections, Clustern oder Sequences strukturiert werden. Zur zusätzlichen Beschreibung können zu jedem Objekt Metadaten eingefügt werden.
Linkkonsistenz: Links zwischen Dokumenten werden als eigene Objekte behandelt, nach dem Bearbeiten von Dokumenten (z.B. verschieben oder löschen) werden diese automatisch aktualisiert.
User und Gruppenrechte: Ein System von Zugriffsberechtigungen erlaubt die selektive Vergabe von Zugriffsrechten auf Dokumente für bestimmte User und Gruppen.
Annotationen: Dokumente oder Teile von Dokumenten können mit Anmerkungen versehen werden.
Suchfunktion: Die Suche ist ein integraler Bestandteil von Hyperwave. Sie erstreckt sich sowohl auf Links und Metadaten als auch auf die eigentlichen Inhalte der gespeicherten Dokumente.

Im Rahmen dieser Diplomarbeit wird eine Audiodatenbank auf dem Hyperwave Information Server implementiert. In Kapitel 2 wird ein Überblick über das Internet gegeben. Weiters wird der derzeitige Stand von Multimedia Anwendungen sowie deren Bedeutung - insbesondere Audioapplikationen - im WWW dargestellt. Kapitel 3 behandelt den Einsatz von Metadaten im Internet, gibt eine Übersicht über das Dublin Core Metadaten Elemente-Set und beschreibt die Möglichkeiten der Nutzung von Dublin Core für Audiodaten. In Kapitel 4 wird ein Überblick über digitale Audiosignale gegeben. Es werden Digitalisierungstechniken behandelt und gebräuchliche Datenformate vorgestellt. Besonderes Augenmerk wird auf die für Audioübertragung im Internet nötige Datenreduktion gelegt. In Kapitel 5 werden Konzepte existierender Audiodatenbanken vorgestellt. Diese Konzepte werden in Kapitel 6 analysiert und bewertet. Weiters werden mögliche eigene Ansätze zur Realisierung einer Audiodatenbank erarbeitet. Aus der Bewertung und den eigenen Ansätzen wird dann ein Realisierungsansatz gebildet. In Kapitel 7 wird die Testimplementation beschrieben. Kapitel 8 gibt einen Ausblick auf weitere Aspekte, die im Rahmen einer weiteren Ausbaustufe enthalten sein könnten.

2. Das Internet

2.1 Allgemeines

Das Internet ist heute in aller Munde. War es vor ein paar Jahren noch ein Metier, das einzig vom Wissenschaftsbereich dominiert wurde, so durchdringt das Internet heute alle Medien und erreicht auf diese Weise selbst Leute, die bisher mit Computertechnik oder angewandter Informatik nichts zu tun hatten. Auch im Bereich der Audiotechnik, in dem jahrzehntelang nur analog „gedacht“ wurde, findet zur Zeit ein großer Umbruch statt. Die Übertragung von digitalem Audio über das Internet eröffnet völlig neue Perspektiven und daher wird sich auch dieser Fachbereich auf das Internet einstellen und davon profitieren.

Obwohl man in fast jedem der heute zahlreich auf dem Markt existierenden Internet-Büchern etwas zur Entstehung des weltumspannenden Computernetzwerkes lesen kann, scheint es angebracht, auf die längere Geschichte des Internet zu verweisen. Denn so mancher hält das Internet für eine Erfindung der 90er Jahre, was man insbesondere der engen Assoziation mit dem World Wide Web verdanken kann, das 1990 am Europäischen Institut für Teilchenphysik CERN als die bisher universellste Internet-Anwendung entwickelt worden ist.

2.2 Geschichte des Internet

Der Ursprung des Internet liegt im ARPAnet, das 1969 vier Rechner (drei davon in Kalifornien, einer in Utah, USA) miteinander verband [KYAS96]. ARPA, eine Forschungsabteilung des US-Verteidigungsministeriums, hatte die Aufgabe, ein ausfallsicheres Computernetzwerk zu entwickeln. Schließlich kam die neue Technologie der packet-orientierten Datenübertragung zum Einsatz. Damit konnten Verbindungen zur Nachrichtenübermittlung von Rechner A nach Rechner C nicht nur über Rechner B sondern auch über Rechner D geleitet werden. Die Funktionalität des Netzwerkes ist demnach auch noch dann gewährleistet, wenn beispielsweise Rechner B ausfällt.

Ende der 70er Jahre wurden erstmals testweise verschiedene Computernetze (zu dem auch das ARPAnet in seiner ursprünglichen Form zählte) miteinander verbunden. Das daraus resultierende Gesamtnetzwerk nannte man ein Internet, ein Netz, das Netze verband. Das zugrunde liegende Netzwerkprotokoll (die gemeinsame "Sprache" der Rechner zur Datenübertragung) hieß TCP / IP (Transport Control Protocol / Internet Protocol). Die anfänglichen "Zwischennetze" hatten verschiedene Namen wie ARPAnet, TCP / IP - Internet etc., ehe sich später "Internet" als Eigenname für den immer größer werdenden Verbund aus Rechnernetzen mit dem gemeinsamen Netzwerkprotokoll TCP / IP etablierte.

1983 avancierte TCP / IP zum offiziellen Standard für das Internet [KROL97]. Gleichzeitig wurde der militärische Teil als "Milnet" ausgegliedert. Der übriggebliebene Teil fungierte als Forschungsnetz. Da das 1983 an einer kalifornischen Universität (Berkley) entwickelte Betriebssystem UNIX BSD 4.2 bereits TCP / IP enthielt und durch freie Verfügbarkeit schnelle Verbreitung fand, dominierten UNIX-Rechner über Jahre hinweg das Internet, nachdem es in seinen Anfangsjahren hauptsächlich von Rechnern der Firma DEC (Digital Equipment Company) mit dem Betriebssystem VMS getragen wurde. Es gibt zwar schon seit vielen Jahren TCP / IP - Anpassungen für weitere Rechnerarten, die Integration in deren Betriebssystem fand jedoch erst in jüngerer Zeit statt. Microsoft vollzog die Einbettung von TCP / IP beispielsweise mit seinem Betriebssystem Windows 95 (als Nachfolger von Windows 3.1) [COMPWO96].

Mit dem Entstehen des World Wide Web (WWW) und der Entwicklung grafischer Benutzeroberflächen für den Zugriff aufs Web vollzog das Internet einen Wandel von einer rein textuellen zur multimedialen Informations- und Kommunikationstechnologie. In die gleiche Zeit fällt auch die Kommerzialisierung des Internet. Heute stehen nicht nur Hochschulen und anderen Bildungs- und Forschungseinrichtungen die vielfältigen Leistungsmerkmale des Internet zur Verfügung, sondern auch Organisationen und Unternehmen aller Art. Dementsprechend rasant steigert sich die Anzahl der Hosts im Internet. Die folgende Übersicht von [nw.com] zeigt die weltweite Entwicklung von Hosts von Anfang 1991 bis Anfang 1998.

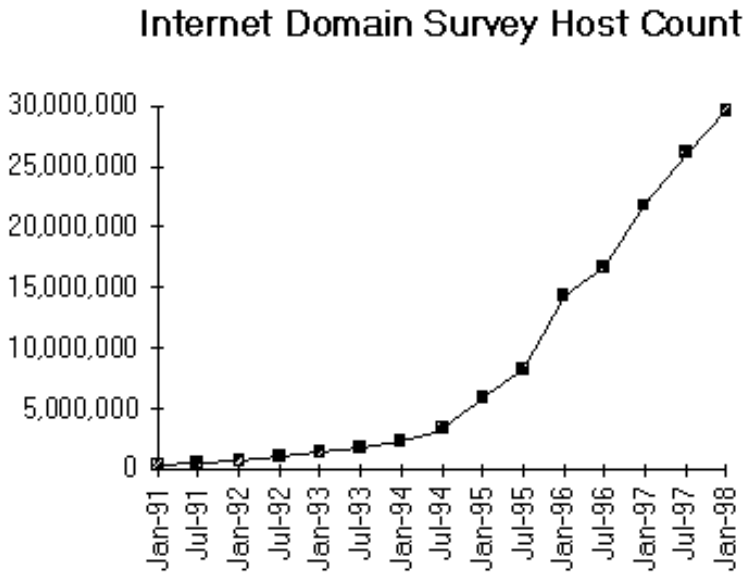

Abb. 2.2 Weltweite Entwicklung von Hosts von Anfang 1991 bis Anfang 1998.

In Zahlen ausgedrückt gab es im Jänner 1998 insgesamt 29.670.000 Hosts, davon 109.154 in Österreich [nw.com].

2.3 Dienste im Internet

Zur Nutzung des Internet wurde eine Vielzahl von Diensten entwickelt. Anfangs recht spartanisch, wurden diese Dienste jedoch mit der Zeit immer komfortabler. Gleichzeitig kamen neue Dienste dazu und mittlerweile ist es gar nicht mehr einfach, den Überblick zu behalten. Da jedoch für jeden Internet-Nutzer ein Basiswissen über diese Dienste unabkömmlich ist, folgt nun eine Übersicht über die wichtigsten Dienste im Internet nach [KROL97].

2.3.1 E-Mail

E-Mail (Electronic Mail) gehört zu den ältesten und am häufigsten genutzten Diensten im Internet. Mit Hilfe dieses Dienstes kann jeder Internet-Nutzer, der auch über eine eigene E-Mail-Adresse verfügt, Nachrichten an andere Netzteilnehmer versenden. Zusätzlich kann man mit den meisten Mail-Programmen auch Dateien an ein E-Mail anhängen und an einen Empfänger versenden. Die Übertragung dauert je nach Größe der Nachricht, der angehängten Datei und der Netzauslastung zwischen wenigen Sekunden und einigen Minuten.

2.3.2 Usenet / Newsgroups

Im Usenet können Netznutzer an Diskussionen über die unterschiedlichsten Themen teilnehmen. Diese Diskussionen können von einem Newsgroupleiter moderiert werden oder frei sein. Jeder Nutzer mit einem News-Client-Programm kann einzelne Newsgroups, die zu seinem Interessenbereich passen, anfordern und die Nachrichten der anderen Teilnehmer lesen und sich zu dem Gelesenen auch äußern, indem er seine Ansicht an die Newsgroup zurückschickt. Das Spektrum reicht von computer-bezogenen Themen, wie Hardware, Betriebssysteme und Anwendungsprogramme über Buchbesprechungen, Jobangebote, bis hin zu psychologischen, esoterischen und religiösen Diskussionen.

2.3.3 File Transfer Protocol (FTP)

Auch das File Transfer Protocol (FTP) ist ein schon recht früh entwickelter Dienst (1973). Ein FTP-Server kann jede Art von Daten, wie Texte, Programme, Graphiken oder Sound- bzw. Video-Files, zum Abruf über das Netz zur Verfügung halten. Mit einem FTP-Programm oder heute auch mit einigen WWW-Browsern kann jeder, der die Adresse des Servers kennt und die Berechtigung besitzt, auf diesen zugreifen und die angebotenen Daten herunterladen. Weiters kann auf einem FTP Server ein sogenannter Anonymous-Zugang eingerichtet werden, der den Zugang ohne Paßwort erlaubt, um Daten und Programme auf den Server zu legen und allen Nutzern zur Verfügung zu stellen.

2.3.4 Telnet

Mit Hilfe des Telnet Dienstes kann man sich in einen entfernten Rechner „einloggen“, und auf diesem arbeiten, als säße man an einem Terminal, das direkt an den Rechner angeschlossen ist. Heute wird dieser Dienst vorwiegend für Wartungs- und Pflegearbeiten in besonderen Anwendungsfällen sowie für die Programmentwicklung genutzt.

2.3.5 Hypertext und Hypermedia im Internet

Der Begriff Hypertext wurde schon 1965 von Ted Nelson geprägt [wired.com]. Anders als der sequentielle Text in einem Buch, ist Hypertext ein Text mit Verweisen zu anderen Textstellen (siehe auch Abb. 2.3.5). Mit dem WWW kam die Hypertext-Fähigkeit ins Internet. In einem Hypertext können neben Textstellen auch Links vorkommen. Links, auch Hypertext-Verbindungen genannt, sind in ein Dokument integrierte Verweise auf andere Dokumente. Wenn der Nutzer diesen Link mit der Maus anklickt, wird das Dokument geladen, auf das der Link zeigt. So können weiterführende Informationen, Querverweise und Quellen problemlos in jedes Dokument integriert werden [MAURER96].

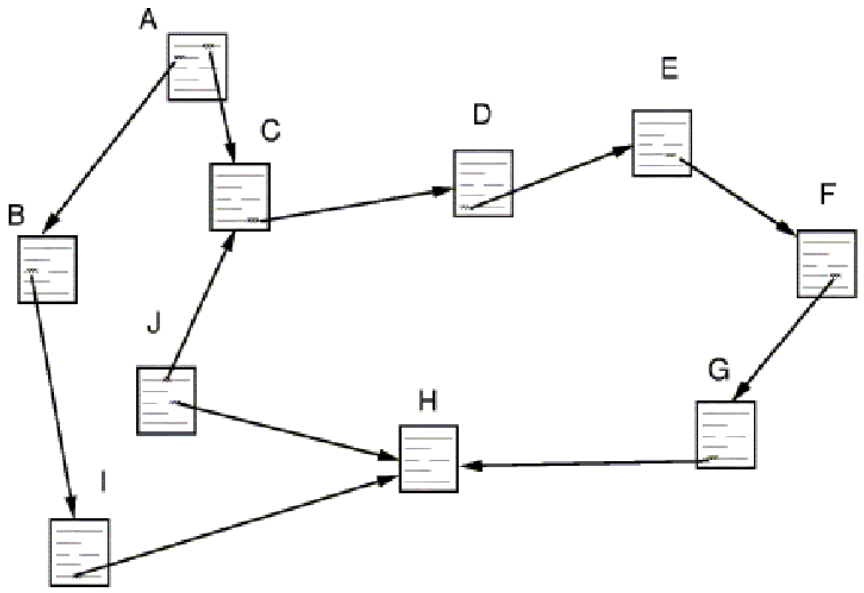

Abb. 2.3.5 Hypertext

Hypermedia ist eine Erweiterung von Hypertext. Links zeigen hier nicht nur auf Textdokumente sondern auch auf Grafiken, Ton- oder Videodokumente. Umgekehrt können diese Hyperlinks natürlich auch aus jeder dieser Informationsquellen heraus auf ein anderes Dokument zeigen.

Ein Hyperlink kann aus einem markierten Text, oder einer Grafik bestehen, und auf Texte, Grafiken, Tondokumente oder Videos verweisen, die irgendwo im Internet stehen können. Auf diese Weise sind Dokumente im WWW miteinander verflochten, so daß sie für den

Nutzer wie eine einzige, riesige Informationsquelle erscheinen, die auf der ganzen Welt verteilt dezentral erstellt und erweitert wird. Die dezentrale Verwaltung von Informationen ermöglicht eine schnelle Aktualisierung, da die Informationen dort bereitgestellt werden können, wo sie anfallen. Bedingt durch die Größe des WWW wird allerdings das Auffinden von relevanten Informationen zusehends erschwert. Aufgrund Hardware- und Funktionseinschränkungen (Frames, Imagemaps sowie dynamischen Webseiten) kann ein Informationsdienst nie alle Webseiten gleichzeitig erfassen. (Alta Vista versucht, nur sehr frequentierte Websites vollständig zu erfassen, trotzdem liegt die durchschnittliche Indexaktualisierung eines bestimmten Webangebotes bei einem Monat.) [GÜTL98]

2.3.6 Multimedia im Internet

Das Schlagwort Multimedia ist in den letzten Jahren von der Computerindustrie als das Modewort schlechthin benutzt worden. Interessanterweise existiert jedoch bisher keine allgemein anerkannte Definition des Begriffs Multimedia. In der Literatur findet man daher oft verschiedene Interpretationen. Eine Definition, die diesem Begriff wahrscheinlich noch am ehesten gerecht wird, lautet: Multimedia ist die "Integration verschiedener Medien, wie Text, Ton, Grafik, Bilder, Video, Animation und Sprache", wobei alle Funktionen "interaktiv einsetzbar" und "gegenseitig synchronisiert" werden können [MEYER91].

Im WWW gehört das Feature "multimediale Anwendungen" zu den herausragenden Leistungsmerkmalen. Anfangs wurde Hypertext durch mit Grafik versehenen Hypertext erweitert. Mit den ersten Bildern wurde auch schon das Zeitalter des Internet als "Multimedium" eingeleitet. Inzwischen sind jene anfänglichen Versuche, von reinen Textanwendungen wegzukommen, durch eine Vielzahl von weiteren Konzepten ergänzt worden, wie z. B. Sprache, Musik (bis hin zu Live-Radiosendungen), Videos und virtuellen Welten. Im Folgenden werden die Darstellungsmöglichkeiten multimedialer Elemente im Internet näher beleuchtet.

2.3.7 Internet Browser

Mit dem Browser Mosaic war es seit 1993 möglich, auf Resourcen im Internet zuzugreifen [RICHTER94]. Kurz darauf wurde von der Softwarefirma Netscape mit dem Navigator die erste graphische Benutzeroberfläche für den Zugriff aufs WWW entwickelt. Da der Download des Navigators kostenlos war, wurde er innerhalb kurzer Zeit sehr populär und zum Standard im Internet. Die Version 2.0 des Netscape Navigators (aktuelle Version seit November 97: 4.x) brachte einige Neuerungen. So wurde z. B. erstmals ein eigenständiger E-mail-Client in den Browser integriert, mit dem das Versenden und Empfangen von E-mails möglich wurde. Des weiteren kam das Konzept der Plug-Ins hinzu [CASSELBERRY97]. Plug-Ins sind zusätzliche Programme, die den Netscape Navigator in seiner Funktionalität erweitern. Über eine von Netscape speziell zur Verfügung gestellte Programmierschnittstelle (API, dem Application Programming Interface) können Hersteller aus aller Welt ihre Ergänzungsbausteine entwickeln. Obwohl Netscape plattform-unabhängig ist und auf den verschiedensten Betriebssystemen wie MS-Windows, UNIX (Derivaten), Macintosh-Oberfläche etc. läuft, sind Plug-Ins z.T. nur auf bestimmten Betriebssystemen lauffähig. Die meisten Plug-Ins stehen für MS-Windows zur Verfügung. Zu den populärsten Plug-Ins zählt beispielsweise das von der Firma Progressive Networks entwickelte "Real-Audio-Plug-In" [real.com]. Neuere Version von Netscape (ab 4.0) fordern

beim Antreffen eines noch nicht installierten Plug-Ins automatisch zum Download des entsprechenden Steckmoduls auf. Netscape gibt auf seiner Site einen Überblick zu einer Reihe von Plug-Ins[1].

Aber auch ohne Plug-Ins ist es möglich, die Funktionalität einer Resource im Internet zu erhöhen. Programmiersprachen wie Javascript oder Java ermöglichen die Integration von Funktionen und Programmen in Dokumente und bieten dem Ersteller dieser Dokumente unzählige Möglichkeiten zur Erhöhung der Attraktivität seiner Homepage. Eine zusätzliche Erweiterung des Funktionsumfanges des Web-Browsers wird durch den Zugriff auf Hilfsanwendungen (Helper Applications) ermöglicht. Hierbei wird die entsprechende Anwendung nicht im Fenster des Browsers geöffnet, sondern ein externes Programm gestartet (z.B. Aufruf eines Word-Dokumentes).

Das Konkurrenzprodukt, der Internet-Explorer von Microsoft [microsoft.com] kam erst 1995 auf den Markt, seitdem tobt ein heftiger Kampf um Marktanteile.

2.3.8 Audio im WWW

Mit der Etablierung des Multimedia-PCs, der mit Hilfe einer integrierten Soundkarte das Abspielen von Sprache und Musik ermöglicht, wurde auch im Internet der Ruf nach der Darstellbarkeit akustischer Anwendungen laut. Populäre Audioformate sind SND, WAVE, AIFF und MIDI (siehe auch Kapitel 4.2.5). Diese Formate können bereits von dem von Netscape mitgelieferten NAPlayer (Netscape Audio Player) gelesen werden. Von kurzen Sequenzen abgesehen, eignen sich diese Formate - mit Ausnahme von MIDI - jedoch weniger zum Online-Abspielen, da alleine der Download-Vorgang sehr zeitraubend sein kann, ehe das Abspielen der Sounddatei auf dem lokalen Rechner erst möglich wird. So benötigt die WAVE Datei eines Audiotracks von 1 min Länge (Quantisierung 16 bit und Samplerate 22 kHz in Mono) 2,64 Mbyte Speicher auf der Festplatte. Bei einer 28,8 Kbit/s – Verbindung würde der Download dieser Datei über 12 min dauern. Beinhaltet der Audiotrack gesprochene Information, stellt sich die Frage ob es nicht effektiver wäre, eine Textversion der Aufnahme zu übertragen. Die dabei zu übertragende Datei wäre bedeutend kleiner, allerdings wäre auch der Informationsgehalt deutlich geringer, da Zusatzinformationen wie z.B. Stimmelodik verloren gehen. Zur näheren Betrachtung wird im Folgenden ein Überblick über die verschiedenen Speichermöglichkeiten einer einminütigen gesprochenen Information gegeben:

Speicherformat	Speicherumfang	Dateigröße
ASCII – Text	Ca. 200 Wörter d.h. ca. 1000 Buchstaben	1 Kbyte
Winword Dokument	Ca. 200 Wörter d.h. ca. 1000 Buchstaben	15 Kbyte
GIF Grafik	Ca. 200 Wörter d.h. ca. 1/3 Din A4	10 Kbyte
WAVE Audio	1 min mit 8kHz, 8 bit, mono	480 Kbyte
WAVE Audio	1.min mit 22 kHz, 16 bit mono	2,64 Mbyte

Tab. 2.3.8 Speichermöglichkeiten im Überblick.

Eine Lösung des Problems bietet die innovative Technologie "Real Audio" von Progressive Networks [real.com], die eine hohe Datenkompression und somit eine Audioübertragung in

[1] URL: http://home.netscape.com/comprod/products/navigator/version_2.0/plugins/index.html

"Echtzeit" gestattet. Die Sounddatei wird ohne vorherige Ladezeit auf dem lokalen Rechner abgespielt. Je kleiner die Übertragungskapazität, desto größer die Kompressionsrate und der damit verbundene Qualitätsverlust. Mit dem Real-Audio-Encoder von Progressive Networks [real.com] wird die Datei im WAVE-Format in das Real Audio Format überführt, wobei eine Optimierung auf vorhandene Bandbreiten stattfindet. Um einen Audio-Stream (Echtzeitübertragung) zu gewährleisten, wird ein sogenanntes Real-Audio-Metafile (.ram) erzeugt, das auf die Datei im Real Audio Format verweist. Es besteht auch die Möglichkeit, in diesem Meta-File auf unterschiedlich stark komprimierte Real Audio Dateien zu verweisen. Bei automatischer Bandbreitenerkennung kann dann die Sounddatei mit der jeweils optimalen Klangqualität übertragen werden. Um Real Audio Streams im Web-Browser anzuhören, muß man allerdings das entsprechende Plug-In downloaden.

Diese Streaming-Technologie ist nicht nur für Audio, sondern auch im Hinblick auf Video-Streams eine der innovativsten Entwicklungen, die das Internet erlebt. Die Möglichkeit, einen Stream nicht nur einem Browser auf Anfrage zuzusenden, sondern den Stream unabhängig von Anfragen beliebig vielen Browsern anzubieten, macht das Internet mit klassischen elektronischen Medien vergleichbar [uu.net].

Angesichts der Möglichkeiten, die das Internet in Bezug auf die Verbreitung und Vervielfältigung von Audiodaten bietet, stellt sich die Frage nach den rechtlichen Aspekten von Audio im WWW. Für den Bereich Musik existieren mächtige Interessensverbände zum Schutz des geistigen Eigentums der Komponisten und Interpreten. Die AustroMechana in Österreich vertritt z.B den Standpunkt, daß das Kopieren einer Audio CD auf eine Festplatte eine gebührenpflichtige Vervielfältigung des Datenträgers ist. Die Entwicklung im Internet hat diese Interessensverbände völlig überrannt. Zur Zeit ist Niemand in der Lage, allgemeingültige Aussagen zu Lizenzgebühren oder zur Rechtssituation zu treffen. Auch die Frage, ob die AustroMechana überhaupt zuständig ist, wenn die Tondatei eines von Ihr vertretenen Künstlers auf einem Server im Ausland gelagert wird, ist offen. Vorläufig ist nur sicher, daß Ausschnitte eines Musikstückes mit einer Länge unter 50 Sekunden frei im Internet verwendet werden können [ix.de]. Längere Ausschnitte werden daher vielfach nur in geschützter Form angeboten. Das von Intertrust entwickelte System E-Commerce bietet die Möglichkeit, jegliche Form von Information zusammen mit definierten Zugriffsrechten in einer sogenannten „Digibox“ zu kapseln. Der Nutzer hat dann die Möglichkeit, bestimmte Zugriffsrechte zu kaufen [intertrust.com].

2.4 Hyperwave Information Server

Der Hyperwave Information Server ist ein Internet/Intranet Dokumentenverwaltungssystem der zweiten Generation. Der Server basiert auf einer objektorientierten Datenbank und ist voll kompatibel zu normalen WWW-Browsern. Gegenüber herkömmlichen Web Servern unterscheidet sich der Hyperwave Server unter anderem durch folgende Konzepte [MAURER96],[HYPERWAVE98]:

- **Objektorientierte Datenbank:** Informationen werden in einer integrierten, objektorientierten Datenbank gespeichert. Objekttypen sind z.B. Dokumente, Links, Collections, Cluster oder Sequences. Die Links in der Datenbank werden konsistent gehalten. Jedem Objekttyp können bestimmte Zugriffsrechte zugewiesen werden.

- **Strukturierung:** Dokumente können innerhalb von Collections strukturiert werden. Jedes Dokument muß mindestens zu einer Collection gehören, kann aber zugleich auch in anderen Collections vorkommen. Die Navigation wird entlang der Collectionshierarchie (die einem azyklischen Graphen entsprechen muß) vollzogen. Dadurch ist es auch möglich, Dokumente sichtbar zu machen, zu denen kein eigentlicher Link existiert. Zur zusätzlichen Beschreibung können zu jedem Objekt Metadaten eingefügt werden.

- **Linkkonsistenz:** Links zwischen Dokumenten werden als eigene Objekte behandelt. Dadurch ist eine prinzipielle Vernetzung mit jedem und auch von jedem Dokumenttyp aus möglich. Links können daher auch z.B mitten in ein Audiofile zeigen. Nach dem Bearbeiten von Dokumenten (z.B. verschieben oder löschen) werden die Links automatisch aktualisiert, „broken links“ sind daher nicht möglich.

- **User und Gruppenrechte:** Es werden verschiedene Stufen der Identifikation unterstützt, so daß jeder berechtigte Benutzer Objekte in den Server einbringen und verwalten kann. Ein System von Zugriffsberechtigungen erlaubt die selektive Vergabe von Zugriffsrechten auf Dokumente für bestimmte User und Gruppen.

- **Annotationen:** Dokumente oder Teile von Dokumenten können mit Anmerkungen versehen werden. Diese Anmerkungen sind wiederum eigene Objekte. Der Ersteller der Anmerkung braucht dabei keine Schreibrechte auf das anzumerkende Dokument haben.

- **Suchfunktion:** Die Suche ist ein integraler Bestandteil von Hyperwave. Sie erstreckt sich sowohl auf Links und Metadaten als auch auf die eigentlichen Inhalte der gespeicherten Dokumente. Der Gültigkeitsbereich der Suche kann sowohl auf eine Collection des Servers eingeschränkt, als auch auf Collectionen anderer Hyperwave Server[2] (Server Pool) erweitert werden.

Der Hyperweave Server ist in mehrere Schichten aufgeteilt, die ihrerseits wieder aus mehreren Komponenten bestehen können [MAURER96],[HYPERWAVE98]:

- **Protocol Conversion Layer:** Die Protokoll Umsetzungsschicht dient zur Umsetzung des intern verwendeten HG-CSP Protokolls in die jeweils geforderten Protokolle. Damit wird der Hyperwave Server zum Multi-Protokoll Server.

- **Session Layer:** Die Sitzungsschicht verarbeitet die eingehenden Anfragen über die Gateways und liefert die gewünschten Informationen aus der Datenbankschicht zurück. Zusätzlich kommuniziert die Sitzungsschicht mit anderen Hyperwave Servern auf entfernten Rechnern.

[2] Mehrere Hyperwave Server können zu einem Server Pool zusammengefasst werden, für den Nutzer verhält sich dieser Server Pool dann so wie ein Hyperwave Server.

- **Database Layer:** Die Datenbankschicht verwaltet die auf dem Hyperwave Server abgelegten Informationen. Die Informationen werden in drei unterschiedlichen Datenbanken aufbewahrt, die jeweils durch folgende Komponenten repräsentiert werden:

  - ftserver
  - dbserver
  - dcserver

  Der ftserver indiziert alle Dokumente und ermöglicht somit eine Volltextsuche. Der dbserver speichert alle Informationstypen wie Objekte, Links und Metainformationen. Der dcserver verwaltet den Inhalt der Dokumente, zusätzlich kommuniziert auch er mit entfernten Server.

Alle Komponenten sind als eigene Prozesse konzipiert. Die Prozesse in der Datenbankschicht existieren pro Server nur einmal, die Prozesse in der Protokoll Umsetzungsschicht und in der Sitzungsschicht existieren öfters; d.h. für jeden Client.

Veränderungen bezüglich Aussehen und Funktionalität[3] des Hyperwave Userinterfaces können durch Ändern der Templates vorgenommen werden. Die Verwendung von Placeholders in den Templates und seit der Version V4.0 Release auch die Verwendung von SS Javascript ermöglichen große Flexiblität bei der Anpassung des Userinterfaces [HYPERWAVE98].

Die aktuelle Version V4.0 Release (Mai 1998) bietet viele zusätzliche Tools (HODMA, CD-ROM Tool, Virtual Folders) und erlaubt eine Anbindung an Oracle Datenbanken.

[3] Dem Hyperwave Konzept entsprechend werden Informationsinhalt und Informationsdarstellung getrennt. Für die Darstellung und Funktionsweise sind die Hyperwave Templates verantwortlich.

3. Metainformationen

In den letzten Jahren hat die Fülle an Informationen, die im Internet angeboten wird, enorme Ausmaße angenommen. Zu jedem erdenklichen Thema existiert eine große Zahl von Fachinformation, die weltweit innerhalb von Sekunden abrufbar ist. Das Internet hat sich in der Tat zur „größten Bibliothek der Menschheit“ entwickelt. Leider fehlt dieser „Bibliothek“ ein zentrales Inhaltsverzeichnis. Durch die dezentrale Struktur (die ja eigentlich ein großer Vorteil ist) ist es nicht möglich, eine zentrale Übersicht über alle angebotenen Informationen zu geben. Das Dilemma läßt sich wie folgt beschreiben: „Das Internet ist wie eine Bibliothek, in der alle Bücher verstreut auf dem Boden liegen und das Licht ausgeschaltet ist. Besucher dürfen sich mit einer Kerze in der Hand auf die Suche nach einem gewünschten Buch machen ...“

Zum Glück hat jedes Buch einen Umschlag, auf dem Autor, Titel des Buches, Verlag und eine Inhaltsangabe abgedruckt sind; dadurch hat der Besucher der Bibliothek wenigstens die wichtigsten Informationen über das Buch, das er gerade in der Hand hält. Außerdem kann er ja hoffen, daß jemand genügend Zeit und Muße aufbringt, und notiert, wo denn nun welches Buch liegt.

Im Internet gibt es Suchmaschinen, die rund um die Uhr die riesigen Datenbestände durchforsten und Verzeichnisse erstellen, aber leider haben viele Daten keinen „Umschlag“. Die relevanten Schlagwörter, mit denen ein Mensch nach diesen Daten suchen würde, müssen daher von den Suchmaschinen aus dem gesamten Inhalt erstellt werden; ein Vorgang der nicht immer sinnvolle Schlagwörter ergibt. Bei Textdaten gibt es noch gewisse Auswahlkriterien (z.B. Überschriften), bei Audiodaten ist für Suchmaschinen zur Zeit einzig und allein der Dateiname informativ. Daher ist es bei solchen Daten besonders wichtig, mit Hilfe von zusätzlichen Informationen (Metainformationen = Information über die Information) genaue Auskunft über den Inhalt der Audiodaten zu geben.

Metainformationen sind nicht nur für Suchmaschinen eine wichtige Zusatzinformation zu den eigentlichen Daten, auch der Nutzer der Daten gewinnt mit Hilfe der Metainformationen bei der Datendarstellung einen guten Überblick. Auch bei großen Datenbeständen in Datenbanken ist der Einsatz von Metainformationen unumgänglich, im Folgenden wird daher näher auf Metainformationen eingegangen.

3.1 Metainformationen im Internet

Daten, die im Internet zur Verfügung stehen, sind mitunter mit Zusatzinformationen über den Inhalt der Daten versehen. Diese Zusatzinformationen werden Metainformationen oder auch Metadaten genannt, sie sind sozusagen ”Daten über Daten” bzw. Daten, die Daten beschreiben.

Metadaten sind Informationen über andere Daten (Dokumenten, Datensammlungen, Audiodaten etc.), die in einer Form gehalten werden, daß sie das Auffinden und die Nutzung

der Primärdokumente erleichtern. Durch Metadaten gewinnen formale und inhaltliche Angaben zu den Werken an Struktur, die Suche nach diesen Werken wird vereinfacht und bei der Recherche können nicht relevante Resultate vermieden werden. Auch für die Darstellung der Daten sind Metadaten wichtig, sie ermöglichen zum Teil erst eine korrekte Interpretation der Daten.

Metadaten dienen der Datenbeschreibung, indem sie Aussagen über diese Daten machen. Eine Aussage besteht aus einem Attribut und einem zugewiesenen Wert. Die einzelnen Aussagen beschreiben jeweils einen Aspekt der Daten und sind voneinander unabhängig. Metadaten bestehen somit aus verschiedenen Aussagen, die Auskunft über die einzelnen Aspekte der Daten geben. Zusammen ergeben diese einzelnen Attribute ein Elemente-Set, das die Daten beschreibt. Welche Attribute dabei sinnvoll verwendet werden, hängt von der Art der Daten ab.

Im Internet können Metadaten in folgenden Arten vorliegen [DEMPSEY97]:

- eingebettet in den Header der HTML-Datei
- in HTML-Dateien, die nur aus Metadaten bestehen, und die auf Nicht-HTML-Dateien (Bild-, Tondateien etc.) verweisen
- in Datenbankstrukturen, die die Zielinformationen enthalten oder auch nicht

Der erste Fall ermöglicht das "Einsammeln" der Metadaten durch eine Suchmaschine. Dies gilt auch für den zweiten Fall. Im dritten Fall kann eine Suchmaschine die Metadaten nicht auflesen, da diese in internen Strukturen verborgen sind. Datenbanken ermöglichen jedoch "domänspezifisches" Suchen, das heißt, Metadaten, die in einer fachspezifischen Datenbank eingetragen werden, werden sowohl über die Datenbankrecherchestruktur auffindbar als auch über eine datenbankinterne Suchmaschine. So ist die Qualitätssteigerung der Recherche mittels Metadaten auch abhängig von den gegebenen Strukturen, in denen die Metadaten gespeichert sind.

Die Metadaten können u.a. in verschiedenen Formaten dargestellt werden [KROL97]:

- **SOIF (Summary Object Interchange Format):** Format zum Datenaustausch zwischen Harvest Gatherer und Broker Komponenten.
- **XML (Extensible Markup Language):** Metasprache zur Beschreibung von Sprachen (z.B HTML). Ermöglicht die Erweiterung dieser Sprachen, auch im Hinblick auf Metainformation.
- **MCF (Meta Content Framework):** Format zur Einbettung von Metadaten in den Inhalt von Dokumenten.

3.2 Überblick über verschiedenartige Metadaten

Metadaten haben verschiedene Zwecke und Funktionalität. David Bearman und Ken Sochats [SOCHATS96] beschreiben die Anforderungen an Metadaten in einer Systemumgebung (wie z.B. im Internet) und definieren sechs verschiedene Arten von Metadaten:

- Metadaten zur Identifikation und zu Nachweiszwecken (z.B. Titel, Autor)
- Metadaten zu Nutzungs- und Beschaffungskonditionen (z.B. Preis, Originalquelle)
- Metadaten zu strukturellen Aspekten (z.B. Art des Dokuments, Datenformat)
- Metadaten zum Kontext (z.B. Querverweise, ähnliche Quellen)
- Metadaten zum Inhalt (z.B. Schlagwörter, Zusammenfassung)
- Metadaten zur Nutzungs- und Wirkungsgeschichte (z.B. andere Dokumente, in denen auf die Daten verwiesen wird)

Bearmans Metadatenmodell soll dem Zweck akzeptabler geschäftlicher Kommunikation und Vermittlung von digitalen Werken über verteilte Informationssysteme und Netze dienen. Es baut auf vorhandenen und angelaufenen Standardisierungsdiskussionen auf und versucht, alle Aspekte, die digitale Daten betreffen könnten, zu berücksichtigen.

3.3 Metadaten Elemente-Sets

Verschiedene Metadaten Elemente-Sets für die inhaltliche und formale Erschließung von digitalen Dokumenten und dokument-ähnlichen Ressourcen wurden in den letzten Jahren von Experten im Rahmen von Digitalisierungsprojekten wie z.B. der NSF-Digital Libraries Initiative (DLI) in USA entwickelt. Einige dieser Sets sind aufgrund ihrer nahezu universellen Einsetzbarkeit komplex und verwenden eine Vielzahl von Attributen. Andere sind aus den Bedürfnissen entstanden, digitalisierte Objekte im Internet eindeutig und mit sehr einfachen Attributen zu beschreiben, wobei zwangsläufig eine Beschränkung auf einen kleinen Bereich von Objekten nötig war. Insgesamt existiert daher eine Vielzahl von Elemente-Sets mit sehr unterschiedlichen Anwendungsbereichen.

Neben den herkömmlichen Katalogisierungsregeln und Klassifikationssystemen werden aufgrund der Anforderungen nach Vereinfachung und internationaler Standardisierung neue Metadaten-Elementensätze konzipiert. Der bekannteste hiervon, der Dublin Core, stammt größenteils aus der bibliothekarischen Welt und wurde nach dem Ort der ersten Tagung zu diesem Zweck, Dublin, Ohio, dem Sitz des Bibiothekslieferanten OCLC, benannt. Im Folgenden wird eine Übersicht über den Dublin Core Elementesatz gegeben.

3.3.1 Dublin Core

Der Dublin Core repräsentiert den "Kern" der inhaltlichen und formalen Erschließungsmerkmale, die sonst für die bibliothekarische und inhaltliche Erschließung benutzt werden. Ursprüngliches Ziel der Bemühungen von Bibliothekaren, Informationswissenschaftlern, Informatikern und Systemspezialisten war es, einen Minimalsatz von Erschließungselementen zu definieren, der zu einer verbesserten Auffindung und Nutzung digitaler Dokumente bei Recherchen im Internet, z.B. mittels Suchmaschinen, verhilft. Obwohl der Dublin Core zunächst an bibliothekarischen Objekten orientiert war, wurde sehr schnell festgestellt, daß Bild-, Ton- und andere Dateiarten ebenfalls mit dem Dublin Core zufriedenstellend indexiert werden können. Seit dem 12. Dezember 1996 gibt es die Version 1.0 des Dublin Core, welche aus 15 Elementen besteht. Diese wurden nach einer ca. 18-monatigen Diskussionsphase in drei internationalen Dublin-Core-Konferenzen, die von der National Science Foundation (NSF) und anderen

Gremien zur Erforschung, Verbreitung und Erschließung digitaler Werke unterstützt wurden, festgelegt. Außerdem einigte man sich auf folgende Richtlinien, die alle Dublin Core Elemente erfüllen sollen: [purl.org]

- Einfach: Die Dublin Core Elemente sollen ohne Vorbildung benützbar sein.
- Universell: Die Elemente sollen in verschiedenen Disziplinen einsetzbar sein.
- International: Weltweite Akzeptanz der Richtlinien soll erreicht werden.
- Flexibel: Erweiterungen sollen möglich sein.

Durch die universelle Einsetzbarkeit ergibt sich auch die Möglichkeit zur Nutzung als Elementeset in einer Audiodatenbank (siehe Kapitel 6.2.1).

3.3.2 Name und Beschreibung der Elemente

Im Folgenden sind die 15 Dublin Core Elemente mit einer kurzen Beschreibung aufgelistet. In den Klammern steht der englische Begriff des Elements (Label) des Dublin Core Element Reference Set [mpib-berlin]. Zum besseren Verständnis wird in Kapitel 3.3.5 ein Beispiel eines Dublin Core Elemente-Sets aufgelistet.

1. Titel (DC.TITLE)

Titel der Quelle (der vom Verfasser, Urheber oder Verleger vergebene Namen der Ressource).

2. Verfasser oder Urheber (DC.CREATOR)

Die Person bzw. Personen oder Organisation bzw. Organisationen, die den intellektuellen Inhalt verantworten (z.B. Autoren bei Textdokumenten; Künstler, Photographen bzw. auch andere Bezeichnungen wie Komponist und Maler bei graphischen Dokumenten).

3. Thema und Stichwörter (DC.SUBJECT)

Thema, Schlagwort, Stichwort. Das Thema der Ressource bzw. Stichwörter oder Phrasen, welche das Thema oder den Inhalt beschreiben. Die beabsichtigte Spezifizierung dieses Elements dient der Entwicklung eines kontrollierten Vokabulars. Das Element kann sowohl systematische Daten nach einer Klassifikation (SCHEME), wie Library of Congress Klassifikations-Nummer oder UDC-Nummer oder Begriffe aus anerkannten Thesauri (wie **ME**dical **S**ubject **H**eadings (MESH) und **A**rt and **A**rchitecture **T**hesaurus (AAT) Deskriptoren) enthalten.

4. Inhaltliche Beschreibung (DC.DESCRIPTION)

Dieser Punkt stellt eine textliche Beschreibung des Ressourceninhalts inklusive einer Kurzfassung (Abstract) bei dokumentähnlichen Ressourcen oder Inhaltsbeschreibungen bei graphischen Ressourcen dar. Künftige Metadaten-Sammlungen können auch numerische Inhaltsbeschreibungen (z.B. Spektralanalyse einer graphischen Ressource) enthalten, die eventuell noch nicht in bestehende Netzsysteme eingebettet werden können. In solchen Fällen, kann dieses Feld einen Link zu einer solchen Beschreibung enthalten statt der Beschreibung selbst.

5. Verleger bzw. Herausgeber (DC.PUBLISHER)

Die Einrichtung, die verantwortet, daß diese Ressource in dieser Form zur Verfügung steht, wie z.B. ein Verleger, ein Herausgeber, eine Universität oder eine korporatives Unternehmen. Der Zweck bei der Benutzung dieses Felds ist es, die Einrichtung oder Einheit zu identifizieren, die den Zugang zur Ressource gewährt. Damit kann gleichzeitig auch Auskunft über die Qualität der Ressource gegeben werden.

6. Weitere beteiligten Personen und Körperschaften (DC.CONTRIBUTORS)

Zusätzliche Person(en) und Organisation(en) zu jenen, die im Element 2 (DC.CREATOR) genannt wurden, die einen bedeutsamen intellektuellen Beitrag zur Ressource geleistet haben, deren Beitrag aber sekundär im Verhältnis zu denen im Element 2 (DC.CREATOR) zu betrachten ist (z.B. Übersetzer, Illustratoren, auch Konferenzleiter, Moderatoren).

7. Datum (DC.DATE)

Das Datum, an dem die Ressource in der gegenwärtigen Form zugänglich gemacht wurde. Der empfohlene Eintrag des Datums wäre eine achtstellige Zahl JJJJMMTT, wie es in ANSI X3.30-1985 definiert ist, beispielsweise: 19961213 (13. Dezember 1996). Andere Darstellungsschemata für das Datum sind erlaubt, werden sie angewandt, so sollten sie eindeutig identifiziert werden (über SCHEME), damit keine Fehlinterpretationen auftreten. Beim 4. Dublin Core Workshop (DC) wurde entschieden, dieses Element etwas offener, jedoch noch stark restriktiv zu definieren, und zwar in einer Unterteilung von zumindest DC.DATE.CREATION (Datum der Herstellung bzw. erstmaligen Veröffentlichung im Netz) und DC.DATE.LASTMODIFIED (Datum der letzten Änderung). Eine DC4-Arbeitsgruppe prüft z. Z., ob weitere Unterteilungen zu empfehlen sind.

8. Ressourcenart (DC.TYPE)

Die Art der Ressource, wie z.B. Homepage, Roman, Gedicht, Arbeitsbericht, technischer Bericht, Essay, Wörterbuch, wird hier eingetragen. Es wird erwartet, daß die Angabe für dieses Feld aus einer definierte Liste von Ressourcenarten entnommen wird. Die offizielle Liste von zulässigen Angaben zum "TYPE" wird noch in einer in Canberra konstituierten Arbeitsgruppe zusammengestellt, und zwar auch in Hinblick auf bestehende Formschlagwort-Normdateien.

9. Format (DC.FORMAT)

Hier wird das datentechnische Format der Ressource eingetragen, z.B. Text/HTML, ASCII, Postscript-Datei, ausführbare Anwendung, JPEG-Bilddatei etc. Die Angabe in diesem Feld gibt die erforderlichen Informationen, die Menschen oder Maschinen ermöglichen, über die Verarbeitungsmöglichkeiten der kodierten Daten zu entscheiden (z.B. welche Hard- und Software benötigt werden, um diese Ressource anzuzeigen bzw. auszuführen). Ähnlich wie bei der Ressourcensart (TYPE) solle die Angabe in diesem Feld (FORMAT) aus festdefinierten Listen, wie die der angemeldeten Internet Media Types (MIME Types) entnommen werden. Grundsätzlich können Formate auch physische Medieneinheiten wie Bücher, Zeitschriften oder andere nichtelektronische Medien mit einschließen.

10. Ressourcen-Identifikation (DC.IDENTIFIER)

Hier wird eine Zeichenkette oder Zahl eingetragen, die diese Ressource eindeutig identifiziert. Beispiele für vernetzte Ressourcen sind URLs und URNs[4] (wo verwendet). Andere weltweit eindeutige Identifikationen, wie die International Standard Book Number (ISBN) oder andere formale Namen wären weitere mögliche Bezeichnungen, die in dieses Element eingetragen werden können.

11. Quelle (DC.SOURCE)

In diesem Element wird - falls nötig - das gedruckte oder elektronische Werk, aus dem diese Ressource stammt, eingetragen. Zum Beispiel, bei einer HTML-Datei eines Sonetts von Shakespeare könnte hier die gedruckte Version der Sonetts als Quelle angeben werden, die als Vorlage für die Erstellung der HMTL-Datei benutzt wurde.

12. Sprache (DC.LANGUAGE)

Hier wird die Sprache bzw. werden die Sprachen des intellektuellen Inhalts dieser Ressource vermerkt. Der Inhalt dieses Felds sollte möglichst mit den von der NISO (National Information Standards Organisation) entwickelten dreistelligen Z39.53-Sprachcodes [sil.org] übereinstimmen (z.B. ENG=English, GER=German, ICE=Icelandic).

13. Beziehung zu anderen Ressourcen (DC.RELATION)

Die Angabe in diesem Feld ermöglicht es, Verbindungen unter verschiedenen Ressourcen darzustellen, die einen formalen Bezug zu anderen Ressourcen haben, aber als eigenständige Ressourcen existieren. Beispiele sind Bilder in einem Dokument, Kapitel eines Buches oder Einzelstücke einer Sammlung. Eine formale Spezifikation (Referenzliste) für dieses Element ist in Arbeit. Benutzer und Entwickler sollten dieses Feld noch als experimentell betrachten.

14. Räumliche und zeitliche Maßangaben (DC.COVERAGE)

Hier werden Angaben zur räumlichen Bestimmung (z.B. geographische Koordinaten) und zeitlichen Gültigkeit eingetragen, die die Ressource charakterisieren (z.B. Ablaufdatum). Formale Spezifikationen für COVERAGE werden z.Zt. erarbeitet. Benutzer und Entwickler sollen dieses Element ebenfalls noch als experimentell betrachten.

15. Rechtliche Bedingungen (DC.RIGHTS)

Vorgesehen für den Inhalt dieses Elements ist ein Link (URL oder andere passende URI soweit zutreffend) zu einem Urhebervermerk, ein "Rights-Management"-Vermerk über die rechtlichen Bedingungen oder ggf. zu einem Server, der solche Informationen dynamisch erzeugt. Die Angaben in diesem Feld ermöglichen es Informationsanbietern, die Verbindung von Bezugs- und Benutzungsbedingungen sowie rechtlichen und abrechnungsbedingten Voraussetzungen oder Urhebervermerken mit einer entsprechenden Ressource oder einer

[4] Eine URL gibt den Zugang zu einer Internetressource als direkter Link auf diese Resource an. Da diese Art der Adressierung jedoch das Problem der referenziellen Integrität mit sich bringt, wurden URN‘s eingeführt. Die URN verweist auf einen Eintrag in einem Name Resolution Server , der die URL enthält [KROL97].

Sammlung von Ressourcen herzustellen. Nutzer dieses Feldes sollten vermeiden, Schlußfolgerungen jeglicher Art zu ziehen, wenn dieses Feld leer oder nicht vorhanden ist.

3.3.3 Nutzung von Dublin Core

Jedes Dublin Core Element ist so definiert, daß es leicht verständlich ist und nicht weiter spezifiziert werden muß, um sowohl für Erschließungs- als auch für Recherche-Zwecke allein funktionieren zu können. Ein Qualifier (ein sogenanntes Unterfeld) - falls vorhanden - verfeinert den Wert und die Bedeutung des Elements ggf. durch den inhaltlichen Bezug zur Benutzung in einer bestimmten Fachcommunity. Qualifiers entstammten zunächst bekannten bibliothekarischen Standards bzw. Standardbegriffen eines Fachgebiets, zu dem die Ressource gehört. Die Flexibilität der nicht obligatorischen Qualifiers ermöglicht eine "Tiefe" u.a. der sachlichen Erschließung, die auch den Anforderungen anspruchsvollerer Forschungs- und wissenschaftlichen Recherchezwecken genügt. Qualifiers sind wichtig, weil sie eine selbst zu definierende Spannweite zwischen Alltagsbenutzern und wissenschaftlichen Nutzern zulassen [WEIBEL95].

Allerdings teilen sich die Meinungen über die Anwendungen von Qualifiers. Anhänger des ursprünglichen Ziels des Dublin Core, einen einfachen Satz von Elementen zu benutzen, die der entsprechenden Erschließung und dem Retrieval von digitalen Ressourcen im Internet höhere Qualität verleihen könnte, behaupten, der Kernsatz von 15 Dublin Core Elementen müßte für eine exakte Suche ausreichen. Andere vertreten die Meinung, daß ohne zusätzliche Qualifiers das inhaltliche Ziel des Dublin Core, nämlich eine verbesserte Suchpräzision, kaum erreichbar wäre, und damit eine Reihe von Qualifiers für einzelne Elemente notwendig und wünschenswert sei, um Eindeutigkeit des Suchverfahrens zu erzielen und gleichzeitig der Komplexität des Fachgebiets bzw. den Ansprüchen der Wissenschaft gerecht zu werden. So einigte man sich während des 4. Dublin Core Workshop (März 1997), daß mit einigen wenigen Ausnahmen alle Elemente auch unqualifiziert ausreichend recherchierbar sein sollen [RUSCH-FEJA97]. Dagegen besteht die Möglichkeit der Elementverfeinerung durch zusätzliche Qualifiers, die aber in einem definierten und beschriebenen Satz von erlaubten Qualifiers für das jeweilige Element enthalten sind.

Die Einfachheit des Dublin Core Elementenschemas ermöglicht es, daß nicht nur Bibliothekare (d.h. Kenner eines oder mehrerer komplexer Erschließungssysteme) diese Metadaten anwenden und "vergeben" können. Mit Hilfe von Eingabemasken mit Auswahlfeldern wird die Vergabe von Metadaten auch für „normale“ Nutzer möglich. Diese Einfachheit wird wahrscheinlich auch ein Pluspunkt für die Akzeptanz solcher Metadaten durch kommerzielle Suchmaschinenhersteller bilden. Allerdings sollte in Betracht gezogen werden, daß Metadaten, die ein „normaler“ Nutzer angelegt hat, nicht unbedingt zuverlässig und richtig sind.

3.3.4 Dublin Core für Audiodaten

Mit dem Dublin Core können unter anderem auch Audiodaten zufriedenstellend indexiert werden. In einem Workshop des PADS (Performing Arts Data Service) [DUFFY97] wurden die Anforderungen von Soundressourcen an den Dublin Core definiert und entsprechende Qualifiers in diesen integriert. Besonderes Augenmerk wurde dabei auf die Einbindung der

Informationen über Herkunft, Art, Rechte und Technologie der Ressource gelegt. Bei gewissen Werken wurde eine Zusammenlegung der Elemente DC.CREATOR und DC.CONTRIBUTORS empfohlen, da es im Bereich der Musik oftmals keinen eindeutig nennbaren Komponisten, dafür aber sehr viele am Werk Beteiligte gibt. Ausserdem wurde für einfache Anwendungen der Qualifier SOUND im Element DC.TYPE definiert, der die Werte Hintergrundgeräusch, Soundeffekt, Musik, Erzählung oder Rede haben kann.

3.3.5 Beispiel eines Dublin Core Elemente-Set

Zum besseren Verständnis wird im Folgenden der Dublin Core Elemente-Set einer Aufnahme aus der „Performing Dickens Collection“ der BBC Educational Resources [bbc.co] aufgelistet.

Dublin Core Element	**Element content**
DC.TITLE	Hard Times
DC.CREATOR.MUSIC	Guillaume.Joannes
DC.CREATOR.WORDS	W.L. Edmonds
DC.CREATOR.PERFORMER	Celia Duffy
DC.CREATOR.PERFORMER	Catherine Owen
DC.SUBJECT	Charles Dickens 1812-1870, comic song, voice and piano, Victorian music hall
DC.SUBJECT.ORIGINATION	London, Players Theatre
DC.DESCRIPTION	Amateur Performance recorded live at the Players Theatre, London. First line begins : "Dickens has lately come out..."
DC.PUBLISHER.THIS FORMAT	Performing Arts Data Service
DC.PUBLISHER.FORMAT=DVD	BBC Educational Resources
DC.DATE.PRINTED SOURCE	1854
DC.DATE.PERFORMANCE	1997
DC.DATE.THIS FORMAT	1998
DC.DATE.FORMAT=DVD	1997
DC.TYPE	sound recording
DC.TYPE.SOUND	Musik
DC.FORMAT	wav.file, 2 Mb, 1 min 30 s
DC.IDENTIFIER.PADS CAT	000004
DC.SOURCE.ISBN DVD	0975497312x
DC.LANGUAGE	English
DC.RELATION	Performing Dickens, BBC Educational Resources, 1997
DC.RIGHTS	Players Theatre Trust, London W1, 0171 333 4444

Tab. 3.3.5 Beispiel eines Dublin Core Elemente-Set.

4. Digitale Audiosignale

Bei der Speicherung, Übertragung und Verarbeitung von Audiodaten in digitaler Form ist ein Problem immer präsent: Digitale Audiodaten sind „Speicherfresser“. Speziell bei der Speicherung vieler Aufnahmen in einer Audiodatenbank macht sich früher oder später „Speicherschwund“ bemerkbar, auch die Übertragung von Audiodaten über das Internet wird durch die großen Datenraten zum Problem. Durch Datenreduktion können diese großen Datenraten jedoch entscheidend verringert werden. Datenreduzierte Audiosignale benötigen weniger Speicherlatz und eine geringere Bandbreite bei der Übertragung, deshalb sollten in einer Audiodatenbank datenreduzierte Audiodaten eingesetzt werden. Zum besseren Verständnis der Materie wird im Folgenden ein Überblick über digitale Audiosignale und Datenreduktion gegeben.

4.1 Informationsumsetzung

Analoge Audioquellen (im hörbaren Bereich des Menschen) geben im Zeitbereich stetig verlaufende Signale mit unbegrenztem Wertevorat ab. Zur digitalen Verarbeitung müssen diese Signale zeitlich diskret abgetastet, quantisiert und codiert werden. Diese Vorgänge werden unter dem Begriff ”Digitalisierung” zusammengefaßt und durch Analog-Digital-Converter (ADC) ausgeführt. Die Rückführung in analoge Signale geschieht durch Digital-Analog-Converter (DAC). Im dazwischenliegenden Bereich werden die digitalen Signale in bitparaleller oder bitserieller Form gespeichert, übertragen und verarbeitet.

Aufgrund anatomischer Gegebenheiten kann das Gehör akustische Ereignisse mit einem Schalldruckpegel von 0 dB bis 120 dB innerhalb des Frequenzbereiches von 20 Hz bis 20 kHz wahrnehmen[5]. Bei der Umsetzung eines analogen Audiosignals in die digitale Domäne ist es daher nötig, in Bezug auf Samplerate und Quantisierung eine Datenrate auszuwählen, die eine Digitalisierung mit einer an diese Wahrnehmungsgrenzen angepaßten Qualität ermöglicht.

4.2 Digitalisierung

Bei der Digitalisierung wird der Signalverlauf eines analogen Audiosignals auf eine Folge von diskreten Digitalwerten abgebildet. Jeder Digitalwert (Sample) beschreibt die Amplitude des Analogsignals zu diesem Zeitpunkt, eine Folge von Samples beschreibt den Amplitudenverlauf des Analogsignals über einen gewissen Zeitraum hinweg. Die Abtastrate gibt an, mit welcher Häufigkeit die Abtastung des Analogsignals erfolgt, die Quantisierung gibt an, welche Wortbreite jeder Sample hat (siehe auch Abb. 4.2).

[5] Mit zunehmenden Alter verringert sich diese Fähigkeit (siehe auch Kapitel 4.3.1)

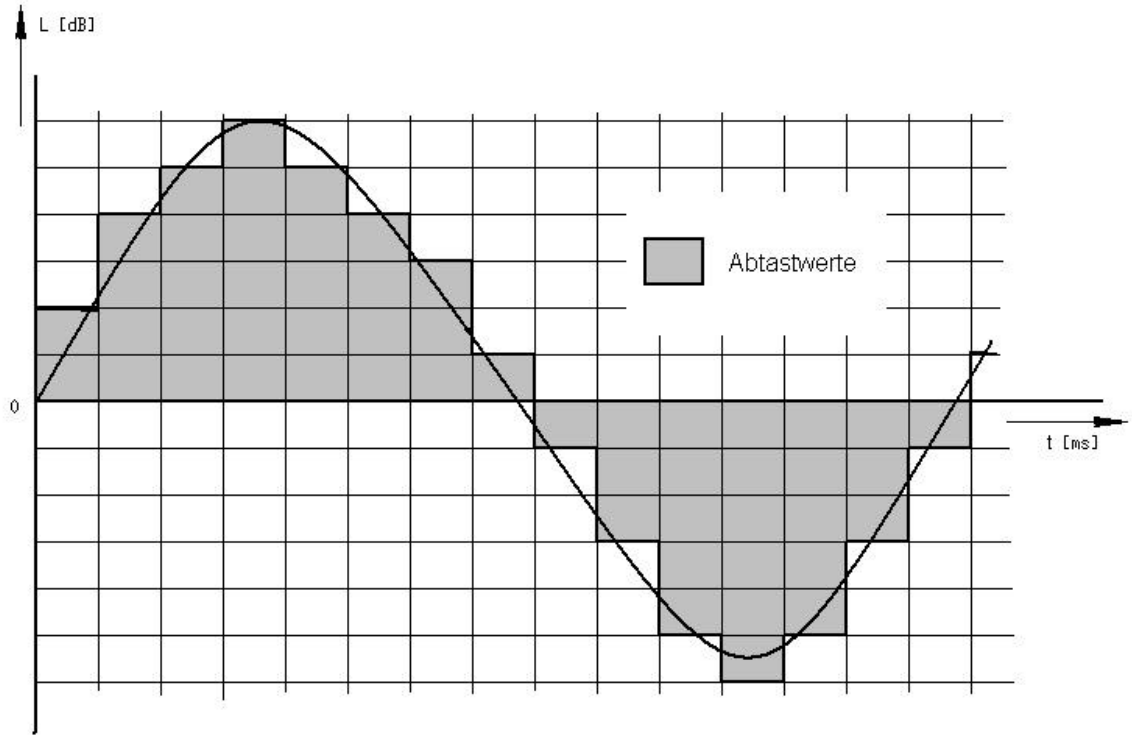

Abb. 4.2 Digitalisierung eines analogen Audiosignals.

4.2.1 Abtastung

Durch die Abtastung erfolgt eine Zeitdiskretisierung des analogen Audiosignals. Für die Frequenz des Abtasttaktes gilt die Nyquist-Bedingung, wonach das abzutastende Analogsignal nur Frequenzanteile besitzen darf, die kleiner als die Hälfte der Taktfrequenz sind, da ansonsten Alias-Komponenten[6] im digitalisierten Signal auftreten [SCHÖNFELDER88].

4.2.2 Quantisierung

Durch die Quantisierung erfolgt eine Amplitudendiskretisierung des analogen Audiosignals. Da das Analogsignal nur durch eine endliche Anzahl von Digitalwerten ausgedrückt werden kann, entstehen Rundungsfehler, die bei einer Wiedergabe des Signals als Quantisierungsrauschen hörbar werden. Dieses Quantisierungsrauschen ist indirekt proportional zur Wortlänge (Bitanzahl).

[6] Ähnlich wie beim Effekt der „stehenden Wagenräder“ in Western-Filmen kann bei der Abtastung von Audiosignalen bei zu geringer Abtastfrequenz die ursprüngliche Information (Drehung des Rades bzw. Schwankung des Audiosignals) verlorengehen. Statt dessen zeigen sich dann Alias Komponenten, diese äußern sich im Audiosignal durch Störanteile mit falschen Frequenzen.

4.2.3 Kodierung

Die durch Abtastung und Quantisierung entstandenen Digitalwerte werden nach bestimmten Regeln kodiert. Es existieren verschiedene Kodierungstechniken, die zum Teil die Datenrate durch Kompression verringern. Die wichtigsten Kodierungstechniken sind (siehe auch Kapitel 4.3) [SCHÖNFELDER88]:

- **PCM (Pulse Code Modulation):** Bei der PCM wird das Analogsignal in Bezug auf die Amplitudenstufen linear kodiert.
- **ADPCM (Adaptive Differential Pulse Code Modulation):** Bei der ADPCM werden die Schwankungen des Analogsignals kodiert.
- **A-law Kodierung:** Der Dynamikbereich des Analogsignals wird nicht linear, sondern nach europäischen Standard logarithmisch kodiert.
- **µ-law Kodierung:** Der der A-law Kodierung entsprechende amerikanische Standard.

4.2.4 Datenraten gebräuchlicher digitaler Audiosysteme

Im Folgenden wird eine Gegenüberstellung der Datenraten von Digital Sattelite Radio (DSR), Compact Disc (CD) und Digital Audio Tape (DAT) gegeben.

Anwendung	Abtastrate	Quantisierung	Datenrate
DSR	32 kHz	16 bit	1,024 Mbit/s
CD	44,1kHz	16 bit	1,412 Mbit/s
DAT	48 kHz	16 bit	1,536 Mbit/s

Tab. 4.2.4 Datenraten digitaler Audiosysteme.

Die Taktfrequenzen 32 kHz und 48 kHz leiten sich aus Vielfachen der Abtastfrequenz von Fernsprechsignalen (8 kHz) her. Die Taktfrequenz 44,1 kHz steht in relativ einfachen Zahlenverhältnissen zu den Zeilenfrequenzen der Fernsehsysteme sowohl mit 625 Zeilen als auch mit 525 Zeilen, weil als CD-Masterbänder vielfach Videocasetten verwendet werden.

4.2.5 Digitale Datenformate

Für die Übertragung von Multimedia über das Internet wurde 1993 das MIME (Multimedia Interchange Mail Extension) Protokoll definiert. Der MIME Typ gibt Auskunft über das Datenformat der übertragenen Files. Zusätzliche Informationen über das Format (Samplerate, Bitbreite, usw.) müssen jedoch durch Auslesen des File Headers gewonnen werden. Zum besseren Verständnis wird im Folgenden ein Überblick über die wichtigsten digitalen Datenformate gegeben [echo.lu].

4.2.5.1 AIFF (Audio Interchange File Format)

MIME: audio/x-aiff
Extention: .aif .aiff .aifc

AIFF ist ein von Apple Computer entwickeltes Audiodatenformat. Ein AIFF File besteht mindestens aus einem Form Chunk, einem Common Chunk und mehreren Sound Data Chunks. Zusätzlich können noch Marker Chunks, Comment Chunks und Instrument Chunks in das AIFF File integriert werden.

- **Form Chunk:** Der Form Chunk beschreibt die Größe und das Format des Files.

- **Common Chunk:** Der Common Chunk beschreibt die Anzahl der Kanäle, die Anzahl der Sample Frames, die Samplerate und die Wortbreite der Sound Data Chunks.

- **Sound Data Chunk:** Jeder Sound Data Chunk beinhaltet Sample Frames. Jeder Sample Frame enthält die Samples für einen oder mehrere (bis zu 6) Audiokanäle, d.h. die eigentliche Audioinformation.

Die zusätzliche Verwendung von Marker Chunks, Comment Chunks und Instrument Chunks ermöglicht die Integration von MIDI-Daten, Sample-Player Daten und zusätzlichen benutzerdefinierten Informationen zu den Audiodaten.

Das AIFF Format wurde zur Speicherung von kodierten Audiodaten um das AIFF-C Format erweitert. Mit diesem Format können unter anderem µ-law, A-law und ADPCM kodierte Daten gespeichert werden (siehe Kapitel 4.2.3).

4.2.5.2 MIDI (Musical Instrument Digital Interface)

MIME: audio/x-midi
Extention: .mid

Ursprünglich wurde das MIDI Protokoll als Standard zum Datenaustausch zwischen Synthesizern entwickelt. Bald jedoch wurde dieser Datenaustausch auch als Dateiformat standardisiert. Das MIDI Format überträgt keine Audiodaten, sondern Kontrollsignale, die ein Synthesizer wiedergibt. Die Datenraten der Kontrollsignale sind sehr gering, dadurch eignet sich MIDI im Bereich von Instrumentalmusik gut als Musikformat für das Internet. Der General MIDI Standard definiert eine Klangbibliothek mit 128 verschiedenen Instrumenten. Dadurch wird erreicht, daß ein bestimmtes Kontrollsignal auf jedem Synthesizer in etwa gleich klingt. Echte Audiodaten können jedoch nur mit dem MIDI Sample Dump Format übertragen werden.

4.2.5.3 WAVE (Waveform Audio File Format)

MIME: audio/x-wav
Extention: .wav

Das WAVE Datenformat wurde von Microsoft und IBM entwickelt und ist Teil des RIFF (Resource Interchange File Format) für Windows 3.1. Aufgrund der großen Verbreitung von Windows ist das WAVE Format ebenso weit verbreitet. Ein WAVE File besteht aus einem Format Chunk und einem Data Chunk. Zusätzlich können noch Cue Chunks, und Playlist Chunks in das WAVE File integriert werden.

- **Format Chunk:** Der Format Chunk beschreibt das Format der Sample Blocks, die Anzahl der Kanäle, die Samplerate und den Blockaufbau im Data Chunk.

- **Data Chunk:** Der Data Chunk beinhaltet die Sample Blocks. Jeder Sample Block enthält Audiodaten in dem im Format Chunk angegebenen Format.

Die zusätzliche Verwendung von Cue Chunks und Playlist Chunks ermöglicht die Definition von Bereichen, die übersprungen oder öfters abgespielt werden können.

Im WAVE Format werden hauptsächlich PCM kodierte Audiodaten gespeichert, es können aber auch µ-law, A-law und ADPCM kodierte Daten gespeichert werden (siehe Kapitel 4.2.3).

4.2.5.4 SND

MIME: audio/basic
Extention: .au, .snd

Das SND Datenformat wird hauptsächlich von NeXT und Sun verwendet. Ein SND File besteht aus Kopf (SNDSoundStruct) und Daten. Auf NeXT Systemen ist der SNDSoundStruct direkt in eine Standardbibliothek (SoundKit) eingebunden. Dadurch ist es möglich, einen Klang abzuspielen oder auch einfache Editierfunktionen bereitzustellen, ohne Kenntnis vom internen Aufbau einer SND Datei zu haben.

- SNDSoundStruct: SNDSoundStruct beschreibt das Format der Daten, die Anzahl der Kanäle, die Samplerate und beinhaltet optionale Textinformation.

Im SND Format werden PCM kodierte Audiodaten oder µ-law, A-law oder ADPCM kodierte Daten gespeichert (siehe Kapitel 4.2.3).

4.3 Datenreduktion

Die Datenraten gebräuchlicher digitaler Datenformate bewegen sich in einer Größenordnung, die eine Verwendung in Audiodatenbanken und eine Übertragung mit im Internet -mit den zur Zeit üblichen Übertragungsraten[7]- ausschließt. Audiodaten in Datenbanken im Internet sind daher meistens datenreduziert. Bei der Datenreduktion wird

[7] Neue Übertragungstechniken werden in Zukunft die derzeit möglichen Übertragungsraten um ein vielfaches steigern; z.B. planen das Unternehmen Teledesic und Bill Gates ein aus 840 Satelliten bestehendes globales Kommunikationssystem, das ab 2002 weltweit breitbandigen Anschluß ans Internet anbieten soll (genaue Angaben über die Datenraten liegen noch nicht vor).

die Datenrate und damit meistens auch die Qualität der Audiodaten verringert. Komplexe Datenreduktionsverfahren mit wahrnehmungsangepasster Datenreduktion reduzieren die Audiodaten auf 1/10 ihrer ursprünglichen Größe (MPEG I Layer 3), ohne die Qualität merklich zu verschlechtern. Der Grundsatz wahrnehmungsangepasster Datenreduktion ist, das jene Anteile des Audiosignals, die vom menschlichen Ohr nicht wahrgenommen werden können, nicht kodiert werden. Im folgenden wird daher ein kurzer Überblick über die Leistungsfähigkeit und die Wahrnehmungsgrenzen des menschlichen Gehörs gegeben.

4.3.1 Hörphysiologie

Aufgrund anatomischer Gegebenheiten kann das Gehör akustische Ereignisse mit einem Schalldruckpegel zwischen 0 dB und 120 dB innerhalb des Frequenzbereiches von 20 Hz und 20 kHz wahrnehmen. Ein Blick auf die Funktion des Innenohrs zeigt, daß das Frequenzspektrum eines eintreffenden akustischen Ereignisses mittels einer Frequenz-Orts-Transformation auf die Basilarmembran abgebildet und so analysiert wird. Die Leistungsanteile des Spektrums werden dabei nicht mit einer linearen Frequenzskala ausgewertet, sondern je nach Frequenzbereich mit einer unterschiedlichen Lautstärke empfunden [ZWICKER82].

Die Lautstärkeempfindung eines akustischen Ereignisses mit einem gewissen Schalldruckpegel hängt daher von dessen Frequenz ab. Ein Sinuston mit einem gewissen Schalldruckpegel wird wie aus Abb. 4.3.1 ersichtlich je nach Tonhöhe verschieden laut empfunden, umgekehrt hat ein Sinuston mit einer bestimmten Lautstärkeempfindung je nach Frequenz einen anderen Schalldruckpegel.

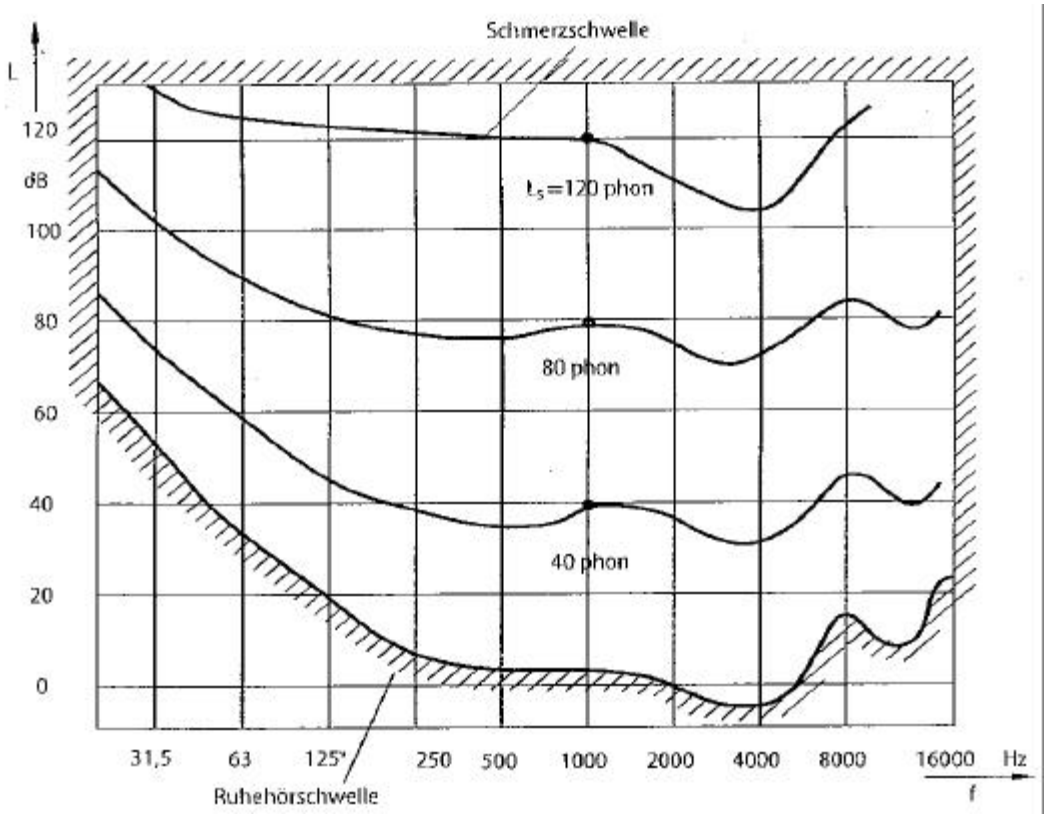

Abb. 4.3.1 Kurven gleicher Lautstärkeempfindung.

Die Kurven gleicher Lautstärkeempfindung sind in Abb. 4.3.1 dargestellt. Der Bereich zwischen Ruhehörschwelle und Schmerzschwelle wird als Hörfläche bezeichnet.

4.3.1.1 Verdeckung

Die Verdeckung ist ein aus dem Alltag bekannter Effekt. Ein akustisches Ereignis, das in ruhiger Umgebung gut hörbar ist, ist in lauter Umgebung nicht mehr hörbar, es wird sozusagen verdeckt (maskiert). Die Psychoakustik unterscheidet zwischen zwei Arten der Verdeckung, der simultanen und der temporären Verdeckung.

A) Simultane Verdeckung

Die Simultane Verdeckung ist ein Phänomen der Frequenzdomäne. Ein akustisches Ereignis mit niedrigen Pegel (maskiertes Signal) kann von einem gleichzeitig stattfindenden starken Signal (Maskierer) unhörbar gemacht werden, falls Maskierer und maskiertes Signal frequenzmäßig nicht zu weit auseinander liegen (siehe auch Abb. 4.3.1.1.a bzw. Abb. 4.3.1.1.b). Die Mithörschwelle bezeichnet jenen Pegel, bei dem ein leises Signal gerade noch durch den Maskierer verdeckt wird. Alle unter der Mithörschwelle liegenden Signale werden vom Gehör nicht mehr wahrgenommen.

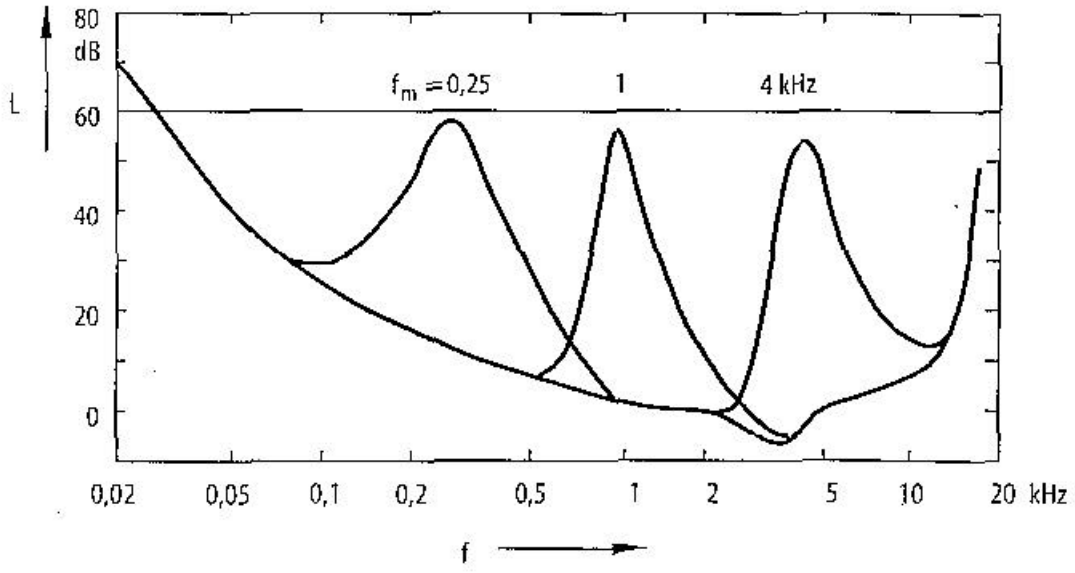

Abb. 4.3.1.1.a Mithörschwellen durch Schmalbandrauschen 60 dB.

In Abb. 4.3.1.1.a sind die Mithörschwellen durch maskierendes Schmalbandrauschen mit 60 dB bei verschiedenen Frequenzen dargestellt.

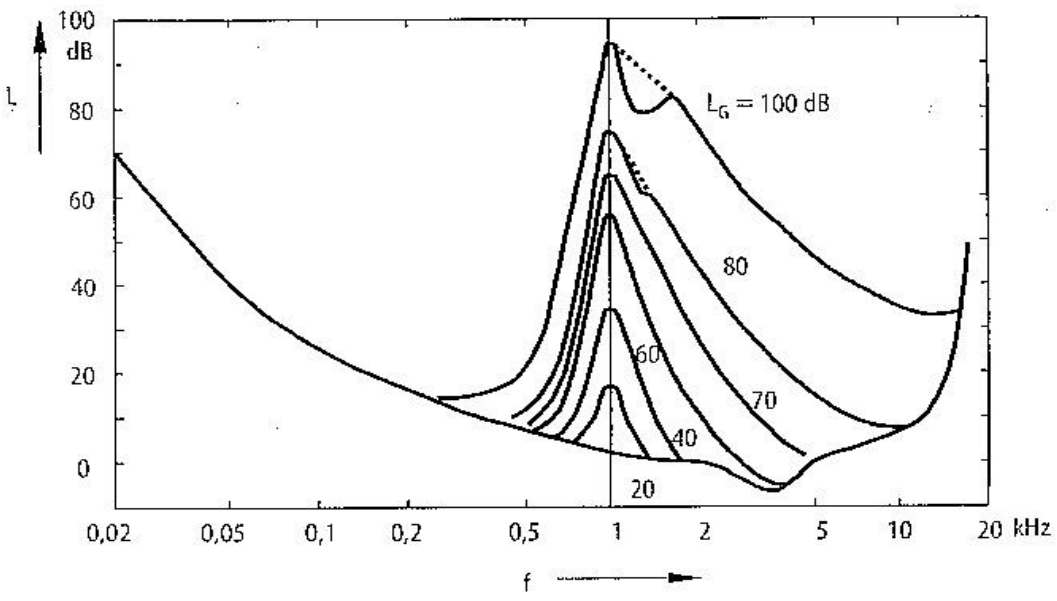

Abb. 4.3.1.1.b Mithörschwellen durch Schmalbandrauschen mit 1 kHz.

In Abb. 4.3.1.1.b sind die Mithörschwellen durch maskierendes Schmalbandrauschen mit 1 kHz bei verschiedenen Schalldruckpegeln dargestellt.

B) Temporäre Verdeckung

Temporäre Verdeckung tritt auf, wenn zwei akustische Ereignisse annähernd gleichzeitig stattfinden. Das stärkere Signal kann dabei das schwächere Signal auch dann verdecken, wenn dieses kurz nach, bzw. sogar kurz vor dem Maskierer eintrifft (siehe Abb. 4.3.1.1.c). Die Zeit, in der Nachverdeckung auftreten kann, beträgt bis zu 200 ms, bei Vorverdeckung sind es nur 50 ms.

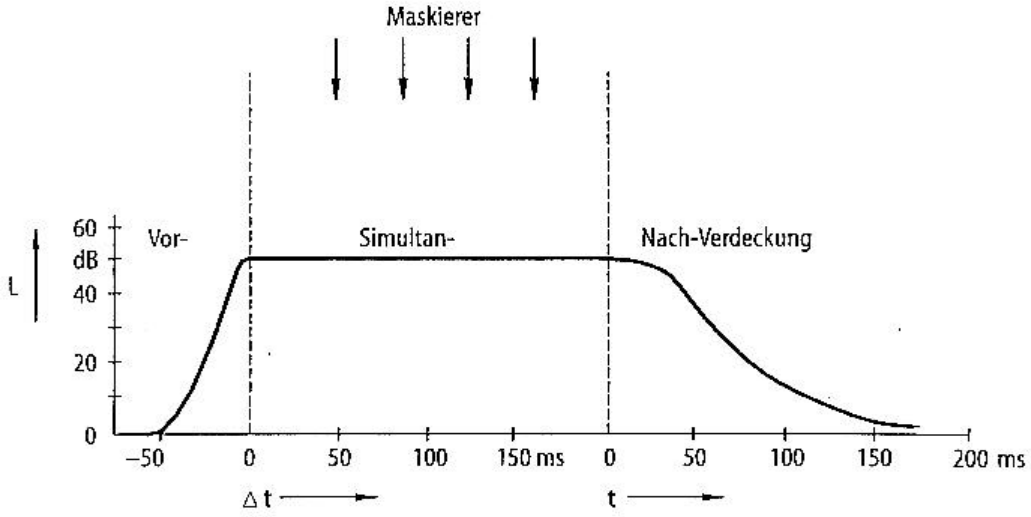

Abb. 4.3.1.1.c Temporäre Verdeckung.

Durch simultane und temporäre Verdeckung maskierte Signalanteile werden vom Gehör nicht wahrgenommen und müssen daher auch nicht abgespeichert bzw. übertragen werden. Die Datenrate des Signals kann daher abgesenkt werden, ohne daß eine merkliche Qualitätsminderung auftritt.

4.3.2 Datenraten gebräuchlicher datenreduziereter Audiosysteme

Im Folgenden wird eine Gegenüberstellung der Datenraten von Digital Audio Broadcast (DAB), Mini Disc (MD) und Digital Compact Casette (DCC) gegeben.

Anwendung	Reduktionsfaktor	Verfahren	Datenrate
DAB	7,35	MPEG Layer II	192 kbit/s
MD	4,83	ATRAC	292 kbit/s
DCC	3,67	MPEG Layer I	384 kbit/s

Tab. 4.3.2 Datenraten datenreduzierter Audiosysteme.

Audiosysteme mit Datenreduktion bringen also eine Ersparnis an Speicher- und Übertragungskapazität um den Faktor 3 bis 7 im Vergleich zu einer herkömmlichen CD. Da diese Ersparnis auch im Hinblick auf eine Verwendung dieser Audioformate in einer Audiodatenbank interessant ist, folgt im Folgenden ein Überblick über die wichtigsten datenreduzierten Datenformate.

4.3.3 Datenreduzierte Datenformate

4.3.3.1 MPEG (Moving Picture Experts Group)

MIME: audio/x-mpeg
Extention: .mpg .mus

Die unter MPEG standardisierten Codierverfahren beruhen alle auf einer Trennung des Verfahrens in Frequenzanalyse und Quantisierung mit dynamischer Bitzuweisung. Zur Frequenzanalyse wird das Audiosignal in eine Anzahl von Teilbänder zerlegt. Die Quantisierung der Teilbänder wird im wesentlichen von zwei Charakteristika eines zugrundeliegenden psychoakustischen Hörmodells beeinflußt. Diese sind simultane und temporäre Verdeckung. Die Quantisierung wird so vorgenommen, daß für das Gehör irrelevante Signalanteile nicht übertragen werden.

Der MPEG I Standard umfasst drei Layer, die sich in der Komplexität des Codierverfahrens unterscheiden. Layer II entspricht dem MUSICAM-Codierverfahren[8], das vom Forschungszentrum CCETT entwickelt wurde. Dieser Layer hat die Anwendungsempfehlung für den Rundfunkbereich bekommen und ist daher weit verbreitet.

[8] Masking-pattern UniversalSubband Integrated Coding And Multiplexing

MPEG I unterstüzt Abtastfrequenzen von 32, 44.1 und 48 kHz und Bitraten zwischen 32 und 192 kbit/s für Monosignale und zwischen 128 und 384 kbit/s für Stereosignale.

MPEG II ist eine Erweiterung von MPEG I hinsichtlich der Abtastfrequenzen (zusätzlich 16, 22.05 und 24 kHz) und mehrkanaliger Codierung.

Das Ziel der Entwicklung von MPEG IV ist die Codierung von synthetischen und natürlichen Audio- und Videosignalen bei niedrigen Bitraten. Bei MPEG IV soll nicht nur die Codierung von zusammengesetzten Audio- und Videosignalen möglich sein, sondern darüberhinaus auch die Codierung einzelner Objekte, aus denen das Signal zusammengesetzt ist. Dadurch soll es möglich sein, die heute immer häufiger geforderte Interaktivität zu unterstützen und höhere Kompressionsraten zu erreichen.

Das Codierverfahren aller drei Layer zerlegt das Eingangssignal in Frequenzbänder und bewertet jedes dieser Subbänder innerhalb eines kurzen Zeitrahmens nach den psychoakustischen Eigenschaften des menschlichen Gehörs. Aufgrund dieser Bewertung wird jedem Subband die nötige Wortlänge in dem diesem Zeitrahmen zugehörigen Datenrahmen zugewiesen (Irrelevanzreduktion). Außerdem wird der Amplitudenbereich, den das Signal innerhalb dieses Zeitrahmens einnimmt, als Skalenfaktor übertragen. Dadurch muß der innerhalb des Zeitrahmens nicht genutzte Amplitudenbereich nicht übertragen werden (Redundanzreduktion). Die folgende Abbildung zeigt den Datenrahmen eines MPEG II codierten Audiosignals.

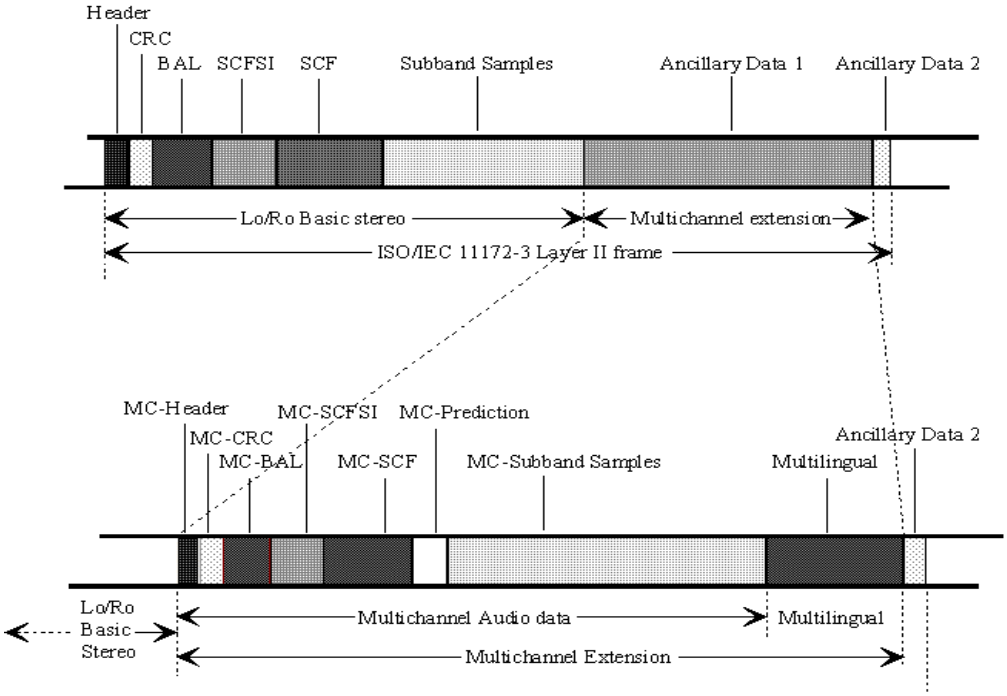

Abb. 4.3.3.1 Datenrahmen eines MPEG II codierten Audiosignals.

Der Datenrahmen (siehe Abb. 4.3.3.1 oben) beinhaltet neben den codierten Subbändern auch noch Seiteninformation, die die Wortlängen und Skalenfaktoren der einzelnen Subbänder angibt und Ancillarydaten (siehe Abb. 4.3.3.1 unten), die die Audiodaten von Mehrkanalaufnahmen enthalten.

Die Unterschiede zwischen den drei Layern der Codierverfahren betreffen in der Hauptsache die Anzahl der Subbänder, die Länge des Rahmens und den Umfang der Filterbänke. Layer I arbeitet mit 32 Subbändern mit einer Bandbreite von 750 Hz, die psychoakustische Bewertung erfolgt für einen Zeitrahmen von 8 ms. Layer II arbeitet mit einem Rahmen von 24 ms und Subbändern mit unterschiedlicher Bandbreite. Bei Layer III wird eine weitere Anpassung der Subbänder an die Frequenzgruppen des Ohres vorgenommen, die Bewertung erfolgt in variablen Zeitrahmen zwischen 8 und 24 ms.

4.3.3.2 Real Audio

MIME: audio/x-pn-realaudio
Extention: .ra

Das Real Audio Format wurde von Progressive Networks [real.com] entwickelt und wird zur Übertragung von Audiodaten im Internet verwendet. Bemerkenswert ist, daß nicht die ganze Audiodatei im Voraus übertragen werden muß. Das Real Audio Format ermöglicht Streaming Audio, d.h. die Wiedergabe startet sofort nach Aufruf der Datei. Möglich wird dies durch eine geringe Datenrate und einer Übertragung über das Internet mit dem UDP (User Datagram Protocol). Jedes UDP Datenpaket enthält einen Teil der Audiodaten der letzten 4 Sekunden. Der Verlust eines Datenpaketes führt daher nicht zu Unterbrechungen der Audiowiedergabe.

Progressive Network gibt über das Codierverfahren keine Auskunft, die geringe Datenrate wird aber nicht durch komplexe Datenreduktionsalgorithmen sondern durch Verzicht auf Audioqualität erreicht. Bei stark datenreduzierten Signalen würde der Verlust von Datenpaketen die Audiowiedergabe stark verschlechtern, daher ist eine komplexe Reduktion nicht möglich. Den Vorteil der sofortigen Wiedergabe ohne vorherigen Download der Audiodatei erkauft man sich also mit einer etwas schlechteren Audioqualität bei der Wiedergabe.

4.3.3.3 ASF (Advanced Streaming Format)

Das Advanced Streaming Format wurde von Microsoft [microsoft.com] entwickelt und ermöglicht die Speicherung und Übertragung von synchronisierten Multimediadaten. Vorhandene Audio- und Videodaten können synchronisiert und gemeinsam datenreduziert werden, wobei verschiedene Datenraten gewählt werden können. Der Kompressionsfaktor beeinflußt dabei die Wiedergabequalität. Ähnlich wie bei Real Audio (siehe Kapitel 4.3.3.2) erfolgt die Wiedergabe sofort nach Aufruf der Datei, allerdings ist im Gegensatz zu Real Audio die Spezifikation des Datenformates öffentlich [asia.microsoft] zugänglich. Ein ASF File besteht aus einem Header Object, einem Data Object und einem Index Object:

- **Header Object:** Das Header Object beschreibt den Inhalt des Files.

- **Data Object:** Das Data Object beinhaltet die einzelnen Audio- und Videodaten in Form von ASF Data Units. Jede ASF Data Unit beinhaltet dabei die Daten für einen bestimmten Zeitabschnitt.

- **Index Object:** Das Index Object beinhaltet die Zeitinformation der einzelnen ASF Data Units.

Softwareunternehmen wie z.B Vivo Audio [vivo.com] haben Userinterfaces entwickelt, die eine komfortable Konvertierung von Audio- und Videodaten in ASF ermöglichen. Die Wiedergabe eines ASF-Streams erfolgt im Netscape Navigator mit einem Plugin (z.B. dem Vivo-Player), der Internet Explorer benötigt kein Plugin.

5. Konzepte von Audiodatenbanken

Im WWW existiert bereits eine Vielzahl von Audiodatenbanken für verschiedenste Zwecke. Der Hauptanteil dieser Datenbanken bietet Zugriff auf Musik, meistens haben die Betreiber das kommerzielle Ziel, CD´s zu verkaufen. Andere, hauptsächlich von Universitäten betriebene Datenbanken dienen zur Veröffentlichung von wissenschaftlich interessanten Audiomaterial[9]. Natürlich gibt es auch Datenbanken, die einfach ”aus Spaß an der Freude” entstanden sind. Für die Implementation einer eigenen Audiodatenbank im Rahmen der vorliegenden Arbeit eignen sich die bestehenden Datenbanken hervorragend als Anregung, als Vorbild oder auch als Mahnung. Im Folgenden wird daher ein kurzer Überblick über verschiedene Audiodatenbanken mit deren Vor- und Nachteilen gegeben. Die dabei getroffene Beurteilung bildet zusammen mit den im nachfolgenden Kapitel vorgestellten eigenen Ansätzen eine Basis für die Realisierung einer eigenen Audiodatenbank.

5.1 Music Boulevard [musicblvd.com]

Music Boulevard ist einer der bekanntesten ”Music Stores” im Internet. Über 200.000 CD´s sind katalogisiert, kurze Passagen einiger Titel können vorgehört werden. Von jeder CD sind alle wichtigen Daten (Artist, Label, CatalogID, Title, Price, ...) gespeichert. Die CD´s sind mehrfach nach Stilrichtung unterteilt, innerhalb einer Stilrichtung sind sie alphabetisch sortiert. Zusätzlich ist ein Interface zur Suche in den Datenbeständen integriert (siehe auch Abb. 5.1). Dabei wird bei den Suchfeldern zwischen Suche nach klassischer und nicht klassischer Musik unterschieden.

[9] z.B. entstand am IICM der TU-Graz das Projekt AEIOU (Annotierbares Elektronisches Interaktives Oesterreichisches Universal-Informationssystem). In diesem Kultur-Informationssystem wurden alle zum Österreichischen Millenium 1996 vorgeschlagenen Projekte mit dem Themenschwerpunkt "Kultur-Daten in elektronisch abrufbarer Form" zu einem gemeinsamen Projekt zusammengeführt. Unter anderem wurde auch die österreichische Musikgeschichte mit Tonbeispielen dokumentiert [aeiou.at].

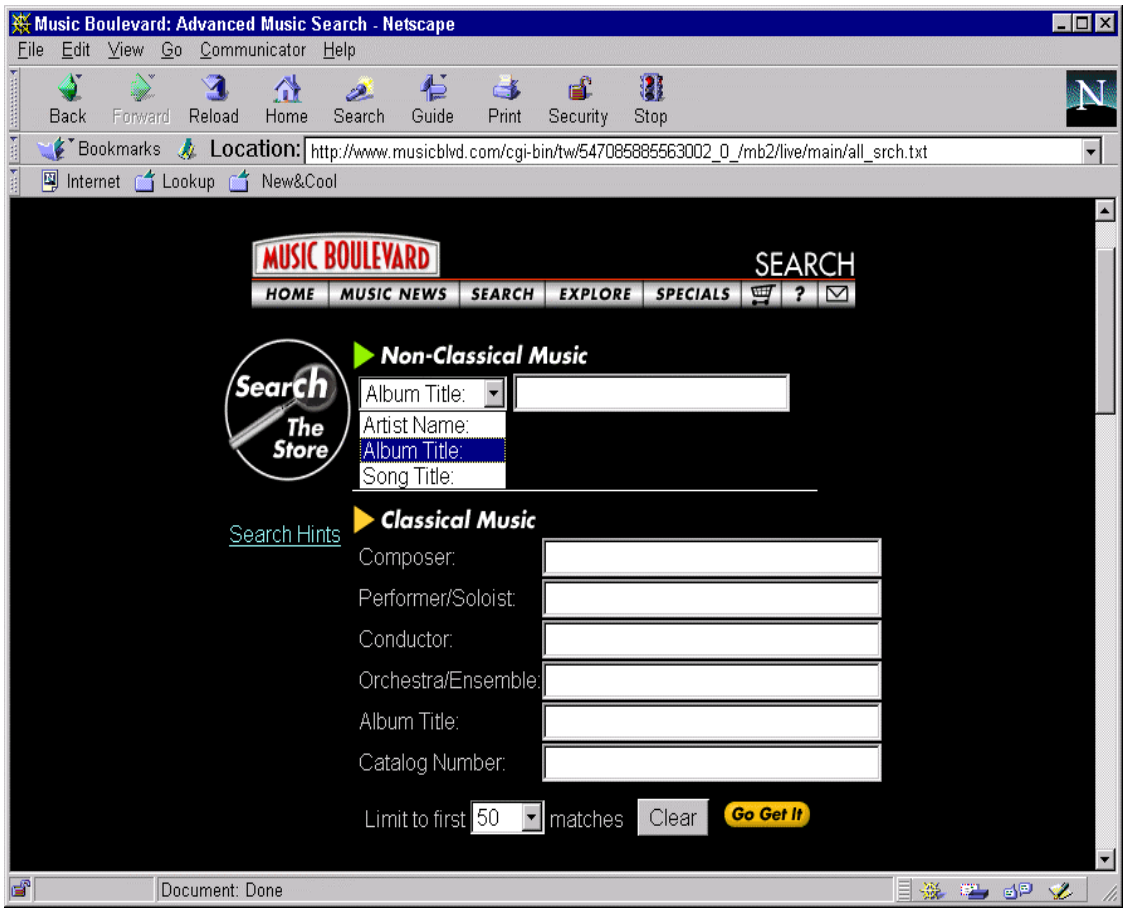

Abb. 5.1 Screenshot der Music Boulevard Audiodatenbank.

Suchkriterien

Nicht Klassische Musik
- Artist Name (z.B. Gary Larson)
- Album Title (z.B. Hip)
- Song Title (z.B. How Do You Do)

Klassische Musik
- Composer (z.B. Chopin)
- Performer/Soloist (z.B. Lena Mirkov)
- Conductor (z.B. Bill Seville)
- Orchestra/Ensemble (z.B. Vienna Symphonic Orchestra)
- Album Title (z.B. Violin Concert No. 1)
- Catalog Number (z.B. A790045)

Vorhören
- Real Audio (siehe Kapitel 4.3.3.2)
- Mpeg II (in verschiedenen Qualitäten, siehe Kapitel 4.3.3.1)

Herunterladen
- Liquid Audio[10] (nach Bezahlung)

[10] Format zur Kapselung von Daten [liquid]. Eine andere Möglichkeit bietet Intertrust (siehe Kapitel 2.3.8)

Vorteile
- Große Datenbank
- Verschiedene Formate und Qualitäten bei Vorhören

Nachteile
- Nur kurze Passagen vorhörbar
- Suchinterface bei nicht klassischer Musik zu wenig detailiert

5.2 IRCAM Soundbase [ircam.fr]

Das IRCAM (Institut de Recherche Musicale et Coordination Acoustique et Musique) bietet auf seiner Homepage unter anderem auch Zugriff auf die IRCAM Soundbase. In dieser Soundbase sind Klänge von 11 verschiedenen Instrumenten abgespeichert. Pro Instrument sind 700 bis 3000 Samples verfügbar, damit wird eine Fülle von Klangvarianten abgedeckt. Die Klänge können zum Vorhören mit niedriger Qualität heruntergeladen werden. Ausgewählte Klänge kann man dann in einem Soundmanager ablegen und am Ende der Session mit hoher Qualität im eigenen Filesystem abspeichern.

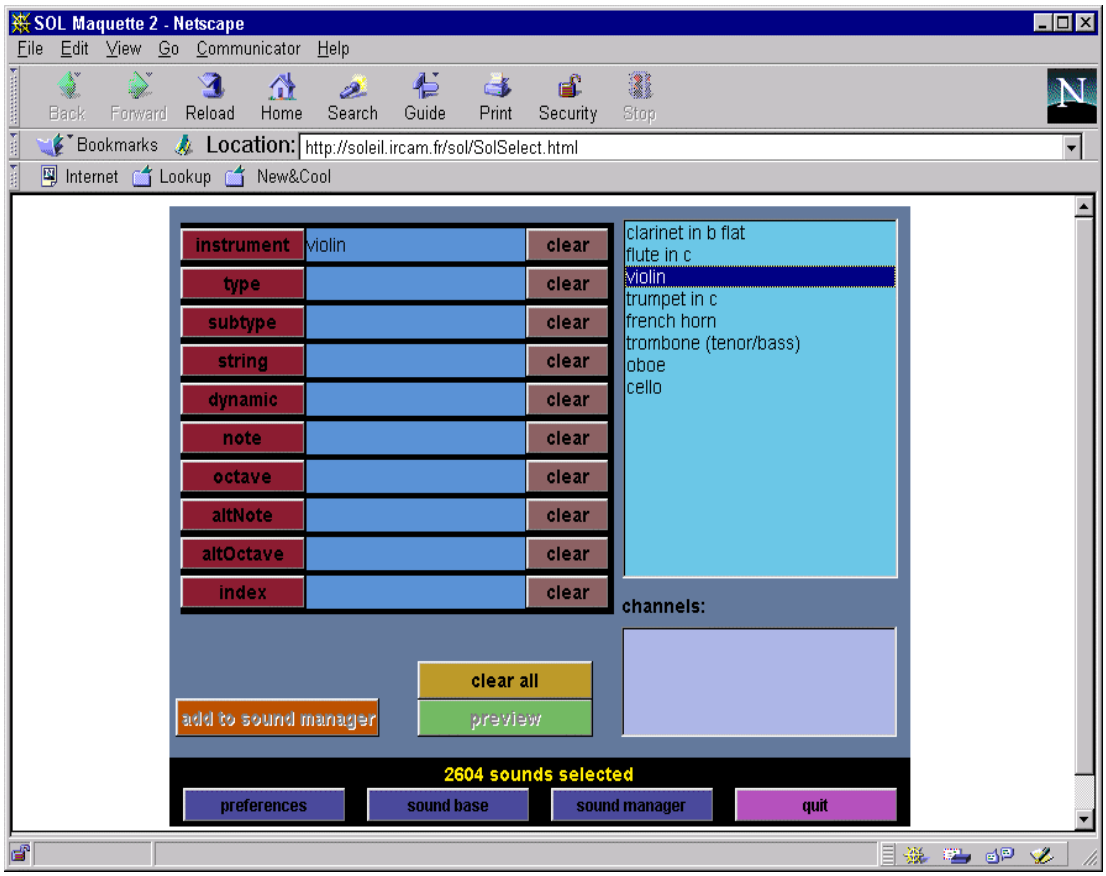

Abb. 5.2 Screenshot der IRCAM Audiodatenbank.

Suchkriterien
- Instrument (z.B. Violine)
- Type (z.B. gestrichen)
- String (z.B. e)
- Dynamic (z.B mezzoforte)
- Note (z.B. fis)
- Octave (z.B. 2)
- Index (z.B. 1312)

Vorhören
- AIFF (Samplerate einstellbar, siehe Kapitel 4.2.5.1)
- AIFF-C (Samplerate einstellbar, siehe Kapitel 4.2.5.1)
- IRCAM[11] (Samplerate einstellbar)

Herunterladen
- AIFF (Samplerate einstellbar)
- AIFF-C (Samplerate einstellbar)
- IRCAM (Samplerate einstellbar)

Vorteile
- Einstellbare Formate und Qualitäten bei Vorhören und Herunterladen
- Soundmanager während Session

Nachteile
- Samplelänge auf 2 sec. beschränkt

5.3 Subaudio [subaudio.net]

Auf den Webseiten von Subaudio befinden sich Informationen aller Art zum Thema Musik. Die zwei Hauptbereiche sind das Sendezentrum und der Plattenladen. Im Sendezentrum findet man Songs, Videos, Konzertmitschnitte und Radiosendungen, die alle mit dem Real-Player abgerufen werden können. Im Plattenladen kann man mit einem Suchinterface nach einem der 32.000 gespeicherten Tonträgern suchen, eine Suche nach Song Title und Vorhören ist hier aber nicht möglich.

[11] Audiodatenformat des IRCAM, ähnlich AIFF

Abb. 5.3 Screenshot der Subaudio Audiodatenbank.

Suchkriterien

- Artist Name (z.B. Rolling Stones)
- Album Title (z.B. Tatoo You)
- Label (z.B. Apple)
- Music Style (z.B. Rock)
- Format (z.B. CD)
- Price max. (z.B. 20$)

Vorteile

- Songs im Sendezentrum in verschiedenen Formaten vollständig vorhörbar
- Keine Beschränkung auf Audio, auch Video abrufbar

Nachteile

- Zweiteilung in Sendezentrum ohne Suchmöglichkeit und Plattenladen ohne Vorhören
- Keine Übersicht über Datenbestände möglich
- Keine Information über Song Titles

5.4 Universität Utrecht Audiodatabase [wwwots.let.ruu]

Das Institut für Linguistik der Universität Utrecht betreibt eine Audiodatenbank, die Zugriff auf viele linguistisch interessante Audiotakes bietet. Neben Aufnahmen von verschiedenen niederländischen Dialekten sind Hörbeispiele vorhanden, welche die Qualitätsminderung von Audiodaten durch Datenreduktion verdeutlichen. Bedingt durch die Verschiedenartigkeit der Aufnahmen erfolgte eine Aufteilung in mehrere Datenbestände. Durch Blättern in der Baumstruktur eines Datenbestandes gelangt man zu den einzelnen Aufnahmen, die zusätzlich durch ein umfangreiches Metadatenset beschrieben werden. Eine Suche mit Suchinterface ist nicht integriert.

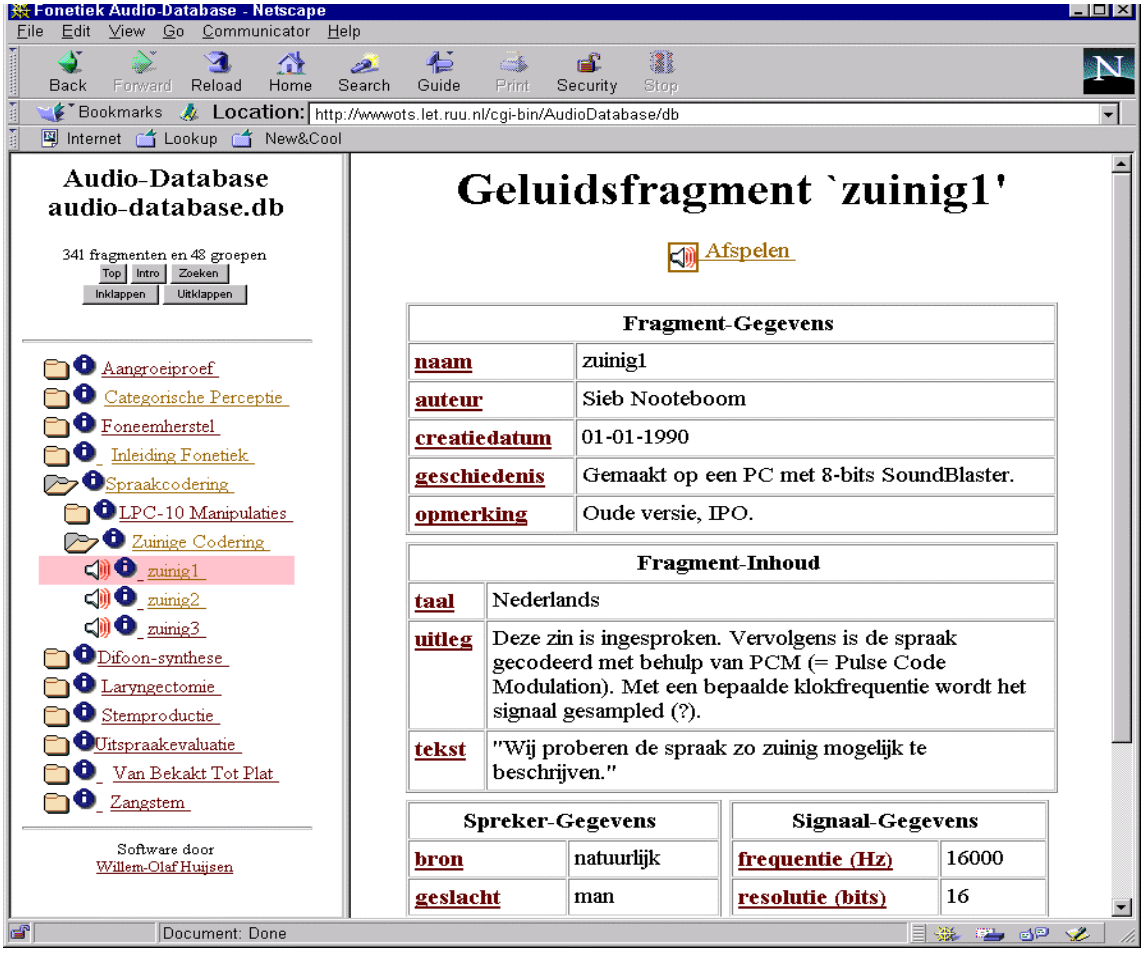

Abb. 5.4 Screenshot derUniversität Utrecht Audiodatenbank.

Herunterladen

- AIFF (siehe Kapitel 4.2.5.1)
- AIFF-C (siehe Kapitel 4.2.5.1)
- SND (siehe Kapitel 4.2.5.4)
- WAVE (siehe Kapitel 4.2.5.3)

Vorteile
- Gute Übersicht über den Datenbestand durch Baumstruktur
- Umfangreiches Metadatenset

Nachteile
- Kein Suchinterface
- Kein Vorhören
- Metadaten in niederländisch

5.5 Performing Arts Data Service [arts.gla]

Das Performing Arts Data Service (PADS) der Universität Glasgow hat das Ziel, digitale Datenresourcen zu sammeln und zu Recherche- und Bildungszwecken zur Verfügung zu stellen. Hauptsächlich sollen für die Kunsthochschule Glasgow interessante Daten gesammelt werden. Es ist keine Beschränkung auf Audiodaten vorgesehen, vielmehr sollen Daten jeglicher Art aus den Bereichen Musik, Film, Video, Theater und Tanz gesammelt werden. Zur Beschreibung der Daten soll das Dublin Core Metadatenset (siehe Kapitel 3.3.1) verwendet werden. Das Konzept der Datenbank wird zur Zeit gerade vorbereitet [arts.gla], detaillierte Informationen über Aufbau und Funktionalität liegen aber noch nicht vor.

5.6 Zusammenfassung

Die Konzepte dieser Audiodatenbanken sind sehr unterschiedlich und vielfältig und auf den jeweiligen Anwendungsfall zugeschnitten. Wirklich universell einsetzbar ist allerdings keine davon. Eine detaillierte Bewertung der soeben vorgestellten Audiodatenbanken folgt in Kapitel 6.1.

6. Realisierungsprozeß einer eigenen Audiodatenbank unter Hyperwave

Der Realisierungsprozeß der Audiodatenbank erfolgt in Anlehnung an das Sytems-Engineering Konzept (SE-Konzept) [HABERFEL92]. Das SE-Konzept enthält eine Reihe von Vorgehensrichtlinien, die sich bei der Realisierung eines Projektes bewähren. Ein Projekt ist unter anderem durch folgende Kennzeichen festgelegt:

- Zeitliche Begrenzung
- Komplexe Aufgabenstellung
- Mehrere Personen und Stellen beteiligt
- Neuartige Aufgabenstellung

Die vorliegende Arbeit kann daher ebenfalls als Projekt angesehen werden. Die wichtigsten Grundrichtlinien, die sich bei der Realisierung eines Projektes bewähren, sind:

- Vom Groben zum Detail

  Das Vorgehensprinzip „Vom Groben zum Detail“ empfiehlt, das Betrachtungsfeld zunächst weit zu fassen und dann schrittweise einzuengen. Dabei werden zuerst generelle Ziele festgelegt, deren Konkretisierungsgrad schrittweise vertieft wird. Konzepte auf höheren Ebenen dienen als Orientierungshilfe für eine detailliertere Lösung.

- Prinzip der Variantenbildung

  Das Prinzip der Variantenbildung empfiehlt, sich nicht mit der erstbesten Lösung zufriedenzugeben, sondern auch andere Varianten in Betracht zu ziehen. Dies ist besonders in frühen Phasen eines Projektes wichtig, da ansonsten andere Lösungsansätze erst in einem späteren Realisierungsstadium zur Sprache kommen und dann das bereits fortgeschrittene Projekt in Frage stellen.

- Abgrenzung der Projektphasen

  Das Prinzip der Abgrenzung von Projektphasen empfiehlt, einen stufenweisen Realisierungsprozeß mit vordefinierten Entscheidungs- und Korrekturpunkten zu definieren. Dadurch wird die Komplexität der Projektabwicklung reduziert.

- Problemlösungszyklus

  Der Problemlösungszyklus soll als Leitfaden zur Behandlung von wichtigen Teilproblemen in jeder Phase eines Projektes bewußt verwendet werden. Der Zyklus läßt

sich in folgende Denk- und Arbeitsschritte gliedern:

- Zielsuche: Wo stehen wir? Was wollen wir? Warum?
- Lösungssuche: Welche Möglichkeiten gibt es, die Ziele zu erreichen?
- Auswahl: Welche Möglichkeit ist die beste, zweckmäßigste?

Grundsätzlich sind beim Realisierungsprozeß der Audiodatenbank folgende relevante Bereiche zu diskutieren:

- Struktureller Aufbau der Audiodatenbank

  - Welche hierarchische Struktur soll der Datenbestand haben?
  - Welches Metadatenset soll die Daten beschreiben?

- Gestaltung der Userinterfaces

  - Wie wird der Datenbestand dargestellt?
  - Wie kann man den Datenbestand nach relevanten Daten durchsuchen?
  - Wie kann man auf relevante Daten zugreifen?
  - Wie wird der Datenbestand gewartet?

Für den Realisierungsprozeß der Audiodatenbank ergibt sich in Anlehnung an das SE-Konzept folgende Vorgehensweise:

1. Bewertung der vorhandenen Audiodatenbanken und Auswahl der positiven Aspekte
2. Einbringen eigener Ideen
3. Variantenbildung aus Punkt 1 und Punkt 2
4. Festlegung des Realisierungskonzeptes

Die vorhandenen Audiodatenbanken werden in Kapitel 6.1 bewertet. Kapitel 6.2 dient der Einbringung eigener Ideen. In Kapitel 6.3 werden aus den positiv bewerteten Punkten vorhandener Datenbanken und den eigenen Ideen mehrere Lösungsvarianten für die Realisierung einer eigenen Audiodatenbank erarbeitet. In Kapitel 6.4 wird aus diesen Lösungsvarianten ein Realisierungsansatz gebildet, der auf Realisierbarkeit unter Hyperwave geprüft und realisiert wird.

6.1 Bewertung der vorhandenen Audiodatenbanken

Die Bewertung der vorhandenen Audiodatenbanken erfolgt hauptsächlich hinsichtlich des strukturellen Aufbaus und der Gestaltung der Userinterfaces. Der strukturelle Aufbau beschreibt die Unterteilung des Datenbestandes in Teilbereiche (z.B. Musik und Sprache). Wie und in welche Teilbereiche die Unterteilung erfolgt, hängt vom Einsatzgebiet der Datenbank ab. Über die Userinterfaces hat der Nutzer Zugriff auf den Datenbestand. Je nach Möglichkeiten, die die Userinterfaces bieten, gestaltet sich für den Nutzer die Recherche nach relevanten Daten umständlich oder einfach.

6.1.1 Music Boulevard

Die hierarchische Struktur besteht aus einer Unterteilung des Datenbestandes in zwei Bereiche (Classical Music und Non-Classical Music). Der Bereich Non-Classical Music ist in Gruppen und Untergruppen - nach Music Style - unterteilt. Es existieren zwei Metadatensets, ein komplexes für Classic (Composer, Performer, Conductor, Orchestra, Album Title, Catalog Number) und ein einfaches für Non Classic (Music Style, Artist Name, Album Title, Song Title). Der Datenbestand ist also in zwei völlig getrennte Bereiche mit jeweils passenden Metadatenset aufgeteilt. Dieser strukturelle Aufbau scheint für eine universell einsetzbare Datenbank sehr gut geeignet zu sein, da eine Unterteilung in Untergruppen mit speziell angepaßten Metadatenset zu einer guten Übersicht über den Datenbestand führt.

Eine Übersichtsdarstellung des Datenbestandes, die alle Audiodaten einer Untergruppe zeigt, wird nicht gegeben. Ein einzelner Datensatz wird in Abhängigkeit der vorhandenen Metadaten flexibel dargestellt; d.h. falls z.B. kein Performer eingetragen ist, wird dieses Feld auch nicht angezeigt. Die Suche nach Daten erfolgt mittels eines zweigeteilten Suchinterfaces (Search for Classical Music or Non-Classical Music). Das Suchergebnis wird als Auflistung aller relevanten Datensätze gezeigt. Der Zugriff auf relevante Audiodaten erfolgt bei der Darstellung eines einzelnen Datensatzes.

6.1.1.1 Bewertung Music Boulevard

Bereich	**Bewertung**	**Anmerkung**
Übersicht Datenbestand	-	Keine Übersicht
Unterteilung Datenbestand	+/-	Unterteilung in 2 Gruppen
Metadatenset	+/-	Flexibel, aber wenige Metadaten
Suchkriterien	+/-	für klassische Musik nicht ausreichend
Audioformate	-	Nicht einheitlich

6.1.2 IRCAM Soundbase

Die hierarchische Struktur besteht aus einer Unterteilung des Datenbestandes in 11 Gruppen (unterteilt nach Instrument). Es existiert ein Metadatenset, das dem Verwendungszweck angepaßt ist (Instrument, Type, String, Dynamic, Note, Octave, Index). Der Datenbestand ist also in 11 Gruppen unterteilt, die alle das gleiche Metadatenset verwenden. Für eine universell einsetzbare Datenbank scheint ein einheitliches Metadatenset jedoch nicht geeignet zu sein, da ein einziges Metadatenset verschiedenartige Audiodaten nicht hinreichend genau beschreiben kann.

Eine Übersichtsdarstellung des Datenbestandes wird nicht gegeben. Ein selektierter Datensatz wird im Soundmanager dargestellt. Die Selektion dieses Datensatzes erfolgt mittels eines Suchinterfaces, in dem durch Auswahl von Werten der gewünschten Metaattribute die Anzahl der angezeigten Datensätze minimiert wird. Das Suchergebnis kann dann in den Soundmanager übertragen werden. Der Zugriff auf relevante Audiodaten erfolgt im Soundmanager.

6.1.2.1 Bewertung IRCAM Soundbase

Bereich	Bewertung	Anmerkung
Übersicht Datenbestand	-	keine Übersicht
Unterteilung Datenbestand	+/-	Unterteilung in 11 Gruppen
Metadatenset	+/-	Fixes, aber großes Metadatenset
Suchkriterien	+	Minimierung der angezeigten Datensätze
Audioformate	+/-	Unpopuläre Formate aber wählbare Samplerate

6.1.3 Subaudio

Wie bei Music Boulevard besteht die hierarchische Struktur aus einer Unterteilung des Datenbestandes in mehrere Gruppen (unterteilt nach Music Style). Es existiert ein Metadatenset, das dem Verwendungszweck angepaßt ist (Music Style, Artist Name, Album Title, Label, Format, Price max.). Der Datenbestand ist also in mehrere Gruppen unterteilt, die alle das gleiche Metadatenset verwenden. Für eine universell einsetzbare Datenbank ist ein einheitliches Metadatenset jedoch nicht geeignet, da ein einziges Metadatenset verschiedenartige Audiodaten nicht hinreichend genau beschreiben kann.

Eine Übersichtsdarstellung des Datenbestandes wird nicht gegeben. Ein einzelner Datensatz wird je nach vorhandenen Metadaten unterschiedlich dargestellt. Die Suche nach Daten erfolgt mittels eines Suchinterfaces. Das Suchergebnis wird als Auflistung aller relevanten Datensätze gezeigt. Der Zugriff auf relevante Audiodaten erfolgt bei der Darstellung eines einzelnen Datensatzes.

6.1.3.1 Bewertung Subaudio

Bereich	Bewertung	Anmerkung
Übersicht Datenbestand	-	keine Übersicht
Unterteilung Datenbestand	+	Unterteilung in mehrere Gruppen
Metadatenset	-	Unflexibel, wenige Metadaten
Suchkriterien	+/-	Einfaches Suchinterface
Audioformate	+	Einheitliche populäre Formatauswahl

6.1.4 Universität Utrecht Audiodatabase

Die hierarchische Struktur besteht aus einer Unterteilung des Datenbestandes in mehrere Bereiche (in linguistisch relevant und tontechnisch relevant). Die Bereiche sind in Gruppen und Untergruppen unterteilt. Jede Untergruppe hat ein eigenes, speziell angepaßtes Metadatenset. Der Datenbestand ist also in mehrere völlig getrennte Bereiche aufgeteilt, die wiederum in mehrere Gruppen und Untergruppen mit speziell angepaßten Metadatensets aufgeteilt sind. Dieser strukturelle Aufbau ist jedoch nur bedingt geeignet, da eine Unterteilung in viele Untergruppen mit speziell angepaßten Metadatenset zu einer unnötig großen Anzahl von Metadatensets führt.

Eine Übersicht des Datenbestandes wird in einer Baumstruktur dargestellt. Der gerade ausgewählte Datensatz wird rechts neben der Baumstruktur dargestellt. Die Suche nach relevanten Daten erfolgt durch blättern in der Baumstruktur. Der Zugriff auf relevante Audiodaten erfolgt bei der Darstellung eines ausgewählten Datensatzes.

6.1.4.1 Bewertung Universität Utrecht Audiodatabase

Bereich	Bewertung	Anmerkung
Übersicht Datenbestand	+	Gute Übersicht durch Baumstruktur
Unterteilung Datenbestand	+	Unterteilung in mehrere Gruppen
Metadatenset	+	Angepaßte Metadaten
Suchkriterien	-	Kein Suchinterface
Audioformate	+/-	Viele, nicht einheitliche Formate

6.1.5 Zusammenfassung

Bei der Bewertung der Audiodatenbanken wurde folgende hierarchische Struktur positiv bewertet: Aufteilung des Datenbestandes in mehrere völlig getrennte Bereiche mit jeweils passenden Metadatenset (z.B. Music Boulevard).

Innerhalb des Datenbestandes der Audiodatenbanken wurden folgende Punkte positiv bewertet:

1. Gute Übersicht über den Datenbestand durch Baumstruktur (z.B. Universität Utrecht Audiodatabase).
2. Unterteilung des Datenbestandes in mehrere Gruppen (z.B. Universität Utrecht Audiodatabase).
3. An die jeweilige Gruppe angepaßte Metadaten (z.B. Universität Utrecht Audiodatabase).
4. Suchinterface zum Eingrenzen der angezeigten Datensätze (z.B. IRCAM Soundbase).
5. Einheitliche populäre Formatauswahl im ganzen Datenbestand (z.B. Subaudio).

6.1.6 Lösungsvorschlag

Nach Analyse und Bewertung der vier existierender Audiodatenbanken wird vorgeschlagen, die Audiodatenbank mit einem in mehrere völlig getrennte Bereiche aufgeteilten Datenbestand zu realisieren, wobei jeder Datenbestand mit einem angepaßten Metadatenset beschrieben wird.

Eine Übersicht über den Datenbestand soll in einer Baumstruktur dargestellt werden. Das Suchinterface soll eine möglichst genaue Eingrenzung der relevanten Datensätze ermöglichen. Die angebotenen Audioformate sollen populär sein. Ausgehend von einem Format, das im gesamten Datenbestand gleich sein sollte, sollen verschiedene zusätzliche Formate anbietbar sein.

6.2 Einbringung eigener Ideen

6.2.1 Beispiele für Metadatensets

Im Folgenden werden mögliche Metadatensets für verschiedene Datenbestände zusammengestellt. Die Auswahl repräsentiert das absolut notwendige Minimum an erforderlichen Attributen und stützt sich auf die Attributsätze existierender Audiodatenbanken. Die danach folgenden Abschnitte beschreiben Möglichkeiten hinsichtlich der Konfigurierbarkeit von Metadatensets.

Musik:	Title	z.B. How Do You Do
	File Name	z.B. howdo.wav
	Mime Type	z.B. audio/x-wav
	Music Style	z.B. Pop
	Composer	z.B. Ritchie Havens
	Performer	z.B. Andy Larson
	Subject	z.B. How Do You Do, Radio Mix
	Description	z.B. Song aus dem Album „Best of Andy Larson“
	Keyword	z.B. Pop, Musik, Larson, Havens
	Text	z.B. Got home from camping last spring, saw many places and things, ...
Sprache:	Title	z.B. Der kleine Mann
	File Name	z.B. mann.wav
	Mime Type	z.B. audio/x-wav
	Speaker	z.B. Helmut Qualtinger, Karl Böhm
	Subject	z.B. Der kleine Mann, Sketch
	Description	z.B. Sketch aus der Reihe „Damals“
	Keyword	z.B. Sketch, Qualtinger, Böhm
	Text	z.B. Sagen Sie, ist Ihnen nicht auch schon aufgefallen daß der Wirt immer freundlicher wird und ...
Geräusch:	Title	z.B. Naturaufnahme
	File Name	z.B. natur.wav
	Mime Type	z.B. audio/x-wav
	Subject	z.B. Naturaufnahmen für Filmvertonung
	Description	z.B. Aufnahmen von Tieren, Gewitter, Wind
	Keyword	z.B. Natur, Film
	Details	z.B. 1. Vogelgezwitscher 2. Kühe am Bauernhof 3. Hundebellen

Aus dieser Auswahl ergibt sich ein minimales Metadatenset, das bei allen Datenbeständen gleich ist und zum Großteil aus Attributen besteht die auch im Dublin Core Element Set (siehe Kapitel 3.3.2) vorhanden sind.

Minimal-Set:	Title	DC.TITLE
	File Name	
	Mime Type	DC.FORMAT
	Subject	DC.SUBJECT
	Description	DC.DESCRIPTION
	Keyword	
	Text	

Dieses Minimal-Set wird je nach Datenbestand durch sinnvolle Attribute erweitert. Dadurch ergibt sich ein für jeden Datenbestand angepaßtes Metadatenset.

6.2.2 Konfigurierbare Metadatensets

In einer universell einsetzbaren Datenbank sollten verschiedene Metadatensets eingesetzt werden können (siehe Kapitel 6.1.6). Für jeden Datenbestand sollte ein passendes Metadatenset vorhanden sein. Durch eine ausreichend große Vielfalt an Metadatensets können zwar die unterschiedlichsten Einsatzmöglichkeiten abgedeckt werden, trotzdem ist der Weg, für jeden neuen Datenbestand aus den bereits vorhandenen Metadatensets ein passendes auszuwählen, nicht zielführend. Selten wird ein Metadatenset optimal für einen neuen Datenbestand geeignet sein, meistens wird man bei der Auswahl Kompromisse eingehen müssen.

Optimal wäre daher ein Metadatenset das man vor Errichten eines neuen Datenbestandes selbst konfigurieren kann, dadurch wäre eine hohe Flexiblität bezüglich der Errichtung neuer Datenbestände gegeben, das Metadatenset könnte sozusagen an den Datenbestand angepaßt werden. Bei Metadatensets sind verschiedene Grade der Konfigurierbarkeit vorstellbar. Im Folgenden wird eine Übersicht über die verschiedenen Möglichkeiten gegeben.

6.2.2.1 Nicht konfigurierbar

Die Meta-Attribute sind vorgegeben. Die den Attributen zuweisbaren Werte sind ebenfalls vorgegeben d.h. es gibt Attribute, bei denen man einen aus einer vorgegebenen Wertetabelle entnommenen Attributwert zuordnen muß. Des weiteren gibt es Attribute bei denen vorgegeben ist, daß man einen frei wählbaren Attributwert zuordnen muß; jedenfalls kann nichts frei konfiguriert werden (siehe Tab. 6.2.2.1).

Beispiel:

Attribut	Wertebereich	Beispiel
Attribut: Song Title	Wert: frei wählbar	Hard Times
Attribut: Song Creator	Wert: frei wählbar	Guillaume Joannes
Attribut: Music Style	Wertetabelle: Funk/Jazz/Soul	Jazz

... Nicht konfigurierbar

Tab. 6.2.2.1 Nicht konfigurierbares Metadatenset.

Die Meta-Attribute Song Title, Song Creator und Music Style sind vorgegeben, es ist nicht möglich diese Attribute zu ändern oder neue Attribute hinzuzufügen. Weiters ist vorgegeben, daß den Attributen Song Title und Song Creator ein frei wählbarer Attributwert zugeordnet werden muß und daß dem Attribut Music Style ein Attributwert aus einer vorgegebenen Wertetabelle (Funk/Jazz/Soul) zugeordnet werden muß.

6.2.2.2 Teilweise konfigurierbar

Die Meta-Attribute sind vorgegeben. Die den Attributen zugewiesenen Werte sind frei konfigurierbar d.h. man hat zwar fix vorgegebene Attribute, es ist jedoch frei konfigurierbar ob man diesen Attributen Werte aus Wertetabellen oder frei wählbare Werte zuordnen muß. Welche Attributwerte eine Wertetabelle enthält ist ebenfalls frei konfigurierbar (siehe Tab. 6.2.2.2).

Beispiel:

Attribut	Wertebereich	Beispiel
Attribut: Song Title	Wert: frei wählbar	Hard Times
Attribut: Song Creator	Wert: frei wählbar	Guillaume Joannes
Attribut: Music Style	Wertetabelle: Funk/Jazz/Soul	Jazz

... Nicht konfigurierbar

... Konfigurierbar

Tab. 6.2.2.2 Teilweise konfigurierbares Metadatenset.

Die Meta-Attribute Song Title, Song Creator und Music Style sind vorgegeben, es ist aber frei konfigurierbar ob dem z.B. dem Attribut Music Style ein Attributwert aus einer Wertetabelle oder ein frei wählbarer Attributwert zugeordnet werden muß (ob man also das Wort Jazz eintippt oder aus einer Tabelle auswählt). Ebenso ist konfigurierbar, welche Werte die Wertetabelle enthält (z.B. Funk/Jazz oder Funk/Jazz/Soul). Ebenso wäre es auch möglich das Metadatenset so zu konfigurieren, daß man dem Attribut Song Title einen Attributwert aus einer Wertetabelle zuordnen muß; dies ist jedoch nicht sinnvoll.

6.2.2.3 Voll konfigurierbar

Die Meta-Attribute sind frei konfigurierbar. Die den Attributen zugewiesenen Werte sind ebenfalls konfigurierbar d.h. man kann beim Errichten eines neuen Datenbestandes die Attribute des Metadatensets frei definieren. Zusätzlich ist frei konfigurierbar ob man diesen Attributen Werte aus Wertetabellen oder frei wählbare Werte zuordnen muß. Welche Attributwerte eine Wertetabelle enthält ist ebenfalls frei konfigurierbar (siehe Tab. 6.2.2.3).

Beispiel:

Attribut	Wertebereich	Beispiel
Attribut: Song Title	Wert: frei wählbar	Hard Times
Attribut: Song Creator	Wert: frei wählbar	Guillaume Joannes
Attribut: Music Style	Wertetabelle: Funk/Jazz/Soul	Jazz

... Konfigurierbar

Tab. 6.2.2.3 Voll konfigurierbares Metadatenset.

Die Meta-Attribute Song Title, Song Creator und Music Style sind vom Datenbankadministrator beim Errichten des Datenbestandes definiert worden, außerdem wurde definiert, ob z.B. dem Attribut Music Style ein Attributwert aus einer Wertetabelle oder ein frei wählbarer Attributwert zugeordnet werden muß. Ebenso wurde vom Datenbankadministrator definiert, welche Werte die Wertetabelle enthält. Die vom Datenbankadministrator definierten Meta-Attribute sollten möglichst sinnvoll ausgewählt werden, um den Nutzer den Umgang mit den Metadaten zu erleichtern. Bei voll konfigurierbaren Metadatensets wäre es prinzipiell auch möglich, z.B. in ein Metadatenset einer Audiodatenbank ein Meta-Attribut „Film Style“ mit der Wertetabelle „red/green/blue“ zu definieren; dies wäre jedoch nicht sinnvoll.

6.2.3 Links

In einer universell einsetzbaren Audiodatenbank sollten zu den Audiodaten auch Metainformationen über diese Audiodaten vorhanden sein. Diese Metainformationen beschränken sich in existierenden Audiodatenbanken jedoch auf Attribute, allenfalls wird eine Textversion der Aufnahme angeboten. Eine sinnvolle Erweiterung wäre, mit Links von dieser Textversion auf zusätzliche Informationen zu verzweigen. So könnten z.B. von der Textversion einer Nachrichtenaufnahme Links auf weiterführende Informationen verweisen (siehe Abb. 6.2.3).

Beispiel:

Attribut: Text	Wien - Die Einführung des Euro bringt für Österreich eine Kostenentlastung von 16 Milliarden Schilling pro Jahr auf Basis 1995. Bei einer BIP-Angleichung auf das Niveau von 1997 bedeutet dies eine Entlastung von rund 20 Milliarden Schilling jährlich. Versailles – Die französische Justiz hat den Parteichef der rechtsextremen Nationalen Front (FN), Jean-Marie Le Pen, wegen einer Prügelei im Wahlkampf 1997 für die Dauer von zwei Jahren für unwählbar erklärt. Die Entscheidung gab das Strafgericht Versailles heute bekannt.

Abb. 6.2.3 Textversion einer Nachrichtenaufnahme mit Links.

Durch anklicken des Links Wien könnte man z.B. zur Homepage der Stadt Wien gelangen.

6.2.4 Audiolinks

Unter Audiolinks sollen Links verstanden werden, die entweder von einer Informationsquelle auf einen Teilbereich einer Audioresouce zeigen, oder von einem Teilbereich einer Audioresource auf eine Informationsquelle zeigen. Grundsätzlich sind beide Möglichkeiten interessante Erweiterungen für eine Audiodatenbank. Im Folgenden werden daher die Möglichkeiten, die Audiolinks bieten, vorgestellt.

6.2.4.1 Audiolinks auf eine Audioresource

In einer universell einsetzbaren Audiodatenbank sollten die Audiodaten nicht nur in einem Stück vorhörbar sein. Eine denkbare Erweiterung wäre, von der Textversion der Aufnahme aus gewisse Teilbereiche der Audiodaten vorzuhören. So könnten z.B. von der Textversion einer Nachrichtenaufnahme Audiolinks auf vordefinierte Teilbereiche der Audiodaten verweisen. Durch anklicken des Audiolinks kann der zugeordnete Teilbereich der Nachrichtenaufnahme vorgehört werden. Die Audiolinks sind durch ein Lautsprechersymbol gekennzeichnet (siehe Abb. 6.2.4.1.a)

Beispiel:

Attribut: Text	Wien - Die Einführung des Euro bringt für Österreich eine Kostenentlastung von 16 Milliarden Schilling pro Jahr auf Basis 1995. Bei einer BIP-Angleichung auf das Niveau von 1997 bedeutet dies eine Entlastung von rund 20 Milliarden Schilling jährlich. Versailles – Die französische Justiz hat den Parteichef der rechtsextremen Nationalen Front (FN), Jean-Marie Le Pen, wegen einer Prügelei im Wahlkampf 1997 für die Dauer von zwei Jahren für unwählbar erklärt. Die Entscheidung gab das Strafgericht Versailles heute bekannt.

Abb. 6.2.4.1.a Textversion einer Nachrichtenaufnahme mit Audiolinks.

Durch anklicken des Audiolinks kann man die zugeordnete Nachrichtenmeldung vorhören. Es ist mit Hilfe von Audiolinks aber auch möglich, in einem Sprachtrainingssystem die richtige Aussprache von Fremdwörtern zu lernen. So könnten z.B. von der Textversion einer Sprachübung Audiolinks auf vordefinierte Teilbereiche der Audiodaten verweisen. Durch anklicken des Audiolinks kann der zugeordnete Teilbereich der Sprachübung vorgehört werden.

Beispiel:

Attribut: Text	**freeze** – froze – frozen **get** – got – got **give** – gave – given **go** – went – gone **grow** – grew – grown

Abb. 6.2.4.1.b Textversion einer Sprachübung mit Audiolinks.

Durch anklicken des Audiolinks kann man die zugeordnete Sprachübung vorhören.

6.2.4.2 Audiolinks von einer Audioresource

In einer universell einsetzbaren Audiodatenbank sollten die Audiodaten nicht nur vorhörbar sein. Zusätzlich könnten beim Abspielen der Audiodaten auch Links zu weiterführender Information angeboten werden. Eine denkbare Erweiterung wäre, im Audioplayer während dem Vorhören der Aufnahme zu gewissen Zeiten einen Linkbutton anzubieten. Von diesem Linkbutton aus könnte dann je nach Inhalt der Audiodaten zu weiterführender Information über das gerade Gehörte verzweigt werden.

6.2.5 Verkettung von Audioresourcen

In einer universell einsetzbaren Audiodatenbank sollten die Audiodaten nicht nur für sich abgeschlossen vorhanden sein. Zusätzlich könnte auch eine definierte Reihenfolge, die diese Audiodaten haben, definierbar sein. Wie auf einer CD sollten die einzelnen Tracks in einer vorgegebenen Reihenfolge vorhörbar sein. Eine denkbare Erweiterung wäre, Linkbuttons anzubieten, die auf den vorhergehenden bzw. folgenden Track verweisen.

6.3 Lösungsvarianten für die Realisierung einer eigenen Audiodatenbank

Nach dem Prinzip der Variantenbildung werden im Folgenden durch Anwenden des morphologischen Kasten[12] mehrere Lösungsvarianten für die Realisierung einer eigenen Audiodatenbank entwickelt. Grundsätzlich stehen folgende Varianten aus Kapitel 6.1 und Kapitel 6.2 zur Verfügung:

Datenbestand	Metadaten	Audioformate	Suchinterface	Audiolinks
(a) Baumstruktur (b) Auflistung (c) Baum&Liste	(a) Konfigurierbar (b) Unkonfigurierbar (c) Teilkonfigurierbar	(a) Einheitlich (b) Verschieden	(a) Ja (b) Nein	(a) Ja (b) Nein

[12] Der morphologische Kasten ermöglicht eine systematische Bildung von Varianten durch Kombination der verschiedenen Möglichkeiten von Teilaspekten.

Aus den 72 theoretisch möglichen Varianten werden zunächst folgende sinnvolle Lösungsvarianten gebildet:

- **Simple Tree (a,b,a,b,b):** Der Datenbestand ist in einer Baumstruktur dargestellt, ein unkonfigurierbares Metadatenset beschreibt die Audiodaten mit einheitlichem Format. Suchinterface und Audiolinks sind nicht integriert.

- **Simple List (b,b,a,b,b):** Der Datenbestand ist in einer Auflistung dargestellt, ein unkonfigurierbares Metadatenset beschreibt die Audiodaten mit einheitlichem Format. Suchinterface und Audiolinks sind nicht integriert.

- **Simple Search (b,b,a,a,b):** Der Datenbestand ist in einer Auflistung dargestellt, ein unkonfigurierbares Metadatenset beschreibt die Audiodaten mit einheitlichem Format. Mit Hilfe eines Suchinterfaces sind Audiodaten suchbar. Audiolinks sind nicht integriert.

- **Extended Tree (a,a,b,a,b):** Der Datenbestand ist in einer Baumstruktur dargestellt, ein konfigurierbares Metadatenset beschreibt die Audiodaten mit verschiedenen Formaten. Mit Hilfe eines Suchinterfaces sind Audiodaten suchbar. Audiolinks sind nicht integriert.

- **Extended List (b,c,b,a,b):** Der Datenbestand ist in einer Auflistung dargestellt, ein teilkonfigurierbares Metadatenset beschreibt die Audiodaten mit verschiedenen Formaten. Mit Hilfe eines Suchinterfaces sind Audiodaten suchbar. Audiolinks sind nicht integriert.

- **Extended Tree&List (c,c,b,a,a):** Der Datenbestand kann in einer Baumstruktur oder in einer Auflistung dargestellt werden, ein teilkonfigurierbares Metadatenset beschreibt die Audiodaten mit verschiedenen Formaten. Mit Hilfe eines Suchinterfaces sind Audiodaten suchbar. Audiolinks sind integriert.

- **Konfigurierbar Tree&List (c,a,b,a,a):** Der Datenbestand kann in einer Baumstruktur oder in einer Auflistung dargestellt werden, ein konfigurierbares Metadatenset beschreibt die Audiodaten mit verschiedenen Formaten. Mit Hilfe eines Suchinterfaces sind Audiodaten suchbar. Audiolinks sind integriert.

6.4 Realisierungsansatz

Ausgehend von den dargestellten Lösungsvarianten wird im Folgenden ein Realisierungsansatz gebildet, der auf Realisierbarkeit unter Hyperwave geprüft und nach eventuellen Abstrichen im Hinblick auf den vertretbaren Aufwand der vorliegenden Arbeit realisiert wird. Die Lösungsvarianten unterscheiden sich in Bezug auf Komplexität und Benutzerfreundlichkeit wesentlich voneinander. Die Variante „Extended Tree&List“ hat einen Großteil der Features integriert und scheint zugleich die Variante zu sein, die dem Nutzer umfassende Möglichkeiten zur Recherche von Audiodaten bietet und im Rahmen der hier vorgestellten Arbeit realisiert werden kann. Diese Lösungsvariante wird daher eingehend auf ihre Realisierbarkeit geprüft.

6.4.1 Prüfung auf Realisierbarkeit

6.4.1.1 Zu realisierende Punkte

Für die Lösungsvariante „Extended Tree&List“ müssen folgende Punkte realisiert werden:

- **Datenbestand:** Der Datenbestand muß sowohl als Auflistung aller Audiodaten eines einzelnen Bereichs als auch als Baumstruktur mit Übersicht über die Audiodaten aller Bereiche darstellbar sein. Der Nutzer muß jederzeit die Möglichkeit haben, zwischen diesen beiden Betrachtungsweisen wechseln zu können.

- **Metadatenset:** Das Metadatenset muß teilkonfigurierbar sein. Ausgehend vom fix vorgegebenen Minimal-Set (siehe Kapitel 6.2.1) muß eine Erweiterung des Metadatensets mit zusätzlichen Attributen möglich sein.

- **Audioformat:** Die Audiodaten müssen in verschiedenen Formaten vorliegen können. Ausgehend von einem Format, daß auf jeden Fall vorliegen sollte, muß eine beliebige Anzahl von zusätzlichen Formaten in der Datenbank ablegbar sein. Der Nutzer muß jederzeit die Möglichkeit haben, ein bevorzugtes Audioformat wählen zu können.

- **Suchinterface:** Ein Suchinterface muß die Möglichkeit bieten, im gesamten Datenbestand oder in einem frei definierbaren Teilbereich nach relevanten Datensätzen zu suchen. Weiters soll das Suchinterface eine verfeinerte Suche bieten, damit das Suchergebnis weiter eingeschränkt werden kann.

- **Audiolinks:** Es muß die Möglichkeit geben, mit Audiolinks von der Textversion einer Aufnahme auf vordefinierte Teilbereiche der Audiodaten zu verweisen. Ebenso muß es möglich sein, mit Links von den Audiodaten auf weiterführende Informationen zu verweisen.

6.4.1.2 Realisierbarkeit unter Hyperwave

Im Folgenden wird geprüft, ob diese Punkte unter Hyperwave realisiert werden können:

- **Datenbestand:** Unter Hyperwave V4.0 ist ein Datenbestand wahlweise als Auflistung aller Dokumente einer Collection (Darstellung: Liste) oder als Baumstruktur mit Übersicht über alle Collections (Darstellung: Baum) darstellbar. Zwischen diesen beiden Betrachtungsweisen kann jederzeit mit Hilfe des Hyperwave Userinterfaces gewechselt werden (Menüpunkt: Einstellungen). Audiodaten sind in einem Objektcontainer mit Metadaten, der Sounddateien und Textdateien enthält, darstellbar. Somit kann der Datenbestand der Audiodatenbank ohne zusätzlichen Aufwand als Auflistung aller Audioobjekte eines einzelnen Bereichs oder als Baumstruktur mit Übersicht über die Audioobjekte aller Bereiche dargestellt werden.

- **Metadatenset:** Unter Hyperwave können zu jedem Objekt frei definierte Attribute vergeben werden. Diese Attribute können die Metadaten einer Aufnahme enthalten. Ein

derart aufgebautes Metadatenset ist einfach konfigurierbar und erweiterbar. Die Attribute, die das fix vorgegebene Minimal-Set enthalten, müssen indiziert werden, um eine Suche nach diesen Metadaten zu ermöglichen. Unabhängig davon ist eine Erweiterung des Metadatensets mit zusätzlichen Attributen möglich. Zum komfortablen Einspielen eines Objektcontainers mit Metadaten, Sound- und Textobjekten muß das Hyperwave Userinterface angepaßt werden.

- **Audioformat:** Unter Hyperwave können die verschiedenen Formate einer Aufnahme in einer Collection liegen. Diese Collection kann für den Nutzer als Objektcontainer mit Metadaten dargestellt werden. Das bevorzugte Audioformat in dieser Collection kann vorgehört oder ins eigene Filesystem geladen werden. Zur Darstellung der Collection als Objektcontainer mit Metadaten und zum Vorhören muß das Hyperwave Userinterface angepaßt werden.

- **Suchinterface:** Unter Hyperwave kann im gesamten Datenbestand oder in einem frei definierbaren Teilbereich nach relevanten Aufnahmen gesucht werden. Auch eine verfeinerte Suche im Suchergebnis ist möglich. Allerdings muß das Suchinterface an das fix vorgegebene Metadaten Minimal-Set und an die Erweiterung mit zusätzlichen Attributen angepaßt werden. Ebenso muß die Darstellung der relevanten Audiodaten angepaßt werden.

- **Audiolinks:** Unter Hyperwave ist auch ein Link ein Objekt, dem Attribute zugeordnet werden können. Daher ist es möglich, einem Link, der von der Textversion einer Aufnahme auf Audiodaten zeigt, die Information über den Teilbereich der Audiodaten (Einsprungstelle und Dauer) als Attribut zuzuordnen. Um das Erzeugen solcher Links zu ermöglichen, muß der Hyperwave Link Wizard angepaßt werden. Außerdem muß ein Audioplayer implementiert werden, der die Information über den Teilbereich der Audiodaten verarbeiten kann.

6.4.1.3 Aufwandsabschätzung

Eine Realisierung der Lösungsvariante „Extended Tree&List“ unter Hyperwave erfordert folgenden Aufwand:

- Zum komfortablen Abspeichern eines Objektcontainers mit Metadaten, Sound- und Textobjekten muß das Hyperwave Userinterface angepaßt werden.

- Die Attribute, die das fix vorgegebene Minimal-Set enthalten, müssen indiziert werden. Indizierte Attribute unter Hyperwave müssen in einem Konfigurationsbereich eingegeben werden.

- Zur Darstellung des Objektcontainers als Aufnahme mit Metadaten und zum Vorhören muß das Hyperwave Userinterface angepaßt werden.

- Das Suchinterface muß an das fix vorgegebene Metadaten Minimal-Set und an die Erweiterung mit zusätzlichen Attributen angepaßt werden (gilt insbesonders für komfortables Arbeiten).

- Die Darstellung der nach einer Suche relevanten Aufnahmen muß angepaßt werden.

- Zum Erzeugen von Audiolinks muß der Hyperwave Link Wizard angepaßt werden.
- Ein Audioplayer, der die Information über den Teilbereich der Audiodaten verarbeiten kann, muß implementiert werden.

Die Realisierung der Lösungsvariante „Extended Tree&List“ unter Hyperwave ergibt einen vertretbaren Aufwand, der in Hinblick der Aufgabenstellung zu vorliegender Arbeit als Testimplementation umgesetzt werden soll. Nachfolgender Abschnitt beschreibt die Realisierung der Audiodatenbank unter Hyperwave.

7. Testimplementation

Bei der Testimplementation der Audiodatenbank wurde auf eine Erweiterung des Hyperwave Servers ohne Reduktion der Standardfunktionalität Wert gelegt. Neben den Audioaufnahmen können daher auch andere Daten (Textdokumente, Grafiken, Videos) in die Datenbank eingefügt werden. Dadurch ist die Testimplementation universell einsetzbar, die zu den Aufnahmen zusätzlich angebotene Information kann damit sehr umfangreich sein.

Bei der Testimplementation wurden alle in Kapitel 6.4.1.3 aufgelisteten Punkte realisiert. Dafür waren u.a. eine Einarbeitung in Funktion und Aufbau des Hyperwave Information Servers, umfangreiche Aktualisierungen der Templates in HTML, PLACE und SSJavascript und die Implementation eines Audioplayers in Java nötig.

Die Testimplementation ist zur Zeit am IICM der TU-Graz unter der URL http://www2.iicm.edu:9000/~Database zugänglich.

Die Beschreibung der Testimplementation ist in drei Abschnitte unterteilt. Kapitel 7.1 beschreibt den für den nicht eingeloggten Nutzer relevanten Bereich, Kapitel 7.2 den für den Datenbankadministrator relevanten Bereich. In Kapitel 7.3. wird ein Einblick in die Funktion der Testimplementation gegeben

7.1 Nutzung der Testimplementation

Die Testimplementation bietet dem nicht eingeloggten Nutzer umfangreiche Möglichkeiten beim Zugriff auf Audiodaten. Mit Hilfe der Übersichtsdarstellung oder des Suchinterfaces ist eine komfortable Recherche - durch Navigieren bzw. Suchen - nach relevanten Aufnahmen möglich. Mit Audiolinks können von der Textversion der Aufnahme aus gewisse Teilbereiche der Aufnahme vorgehört werden, gleichzeitig kann zu weiterführender Information verwiesen werden. Relevante Aufnahmen können in verschiedenen Audioformaten ins lokale Filesystem geladen werden.

7.1.1 Übersichtsdarstellung

Die Übersichtsdarstellung der Audiodatenbank zeigt den Datenbestand mit den darin enthaltenen Bereichen. Die Navigation im Datenbestand erfolgt im View- oder im Tree-Modus. Im View-Modus werden alle Inhalte des aktuellen Bereichs dargestellt, im Tree-Modus wird der gesamte Datenbestand angezeigt. Zwischen diesen beiden Betrachtungsweisen kann jederzeit mit Hilfe des Userinterfaces gewechselt werden.

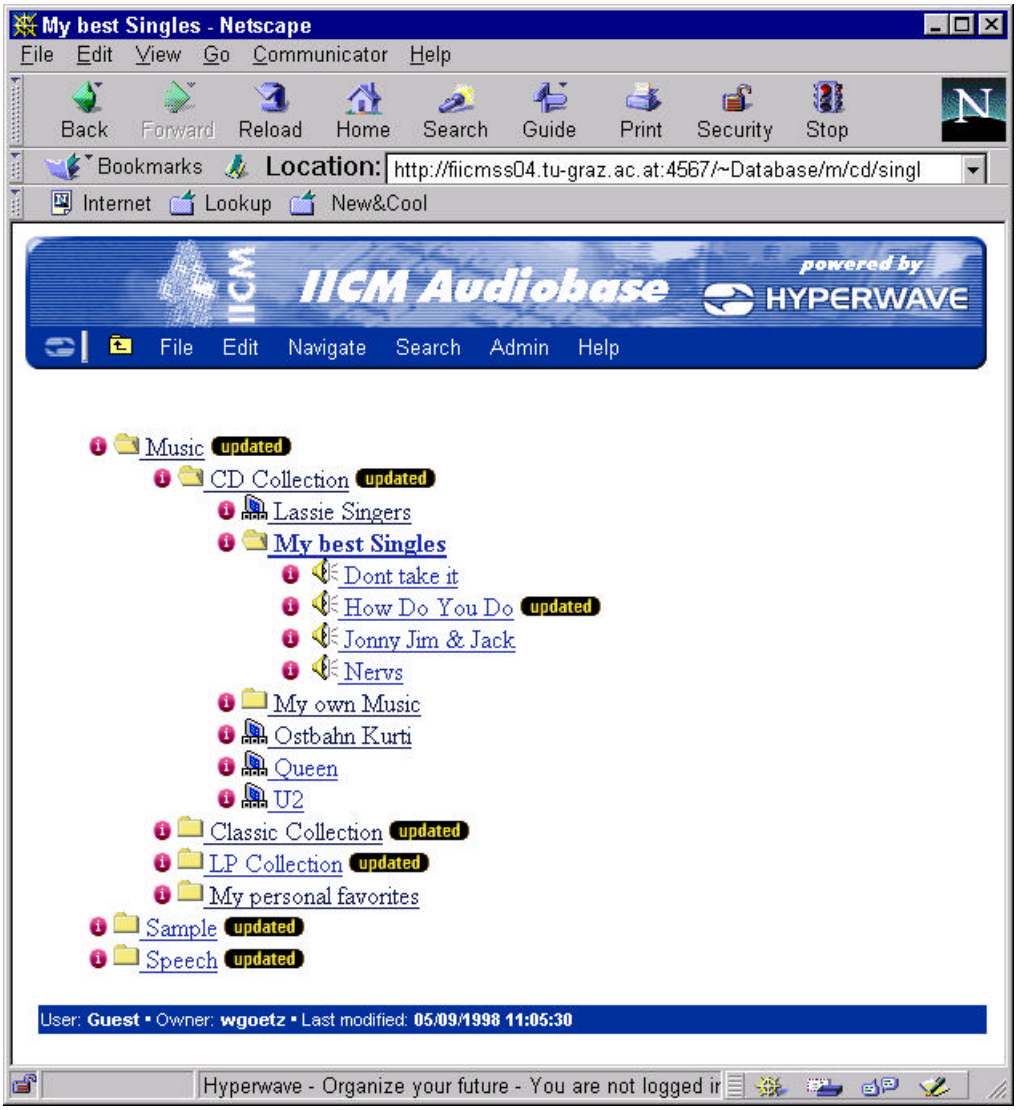

Abb. 7.1.1 Screenshot der Übersichtsdarstellung im Tree-Modus.

Die Testimplementation sieht vor, daß Audioaufnahmen in Collections oder in Sequences vorhanden sein können, nach anklicken einer Aufnahme gelangt man in das Infofenster (siehe Kapitel 7.1.3).

7.1.2 Suche

Zum komfortablen Suchen von Aufnahmen in der Datenbank wurde das Hyperwave Userinterface um den Menüeintrag „Audiosearch“ erweitert. Nach anklicken des Menüeintrages wird man zum Eingeben des Suchbegriffes aufgefordert (siehe Abb. 7.1.2.a). Zusätzlich kann man auswählen, in welchen Metadaten der Suchbegriff gefunden werden soll. Durch anklicken des „Search !“ Buttons wird die Suche gestartet.

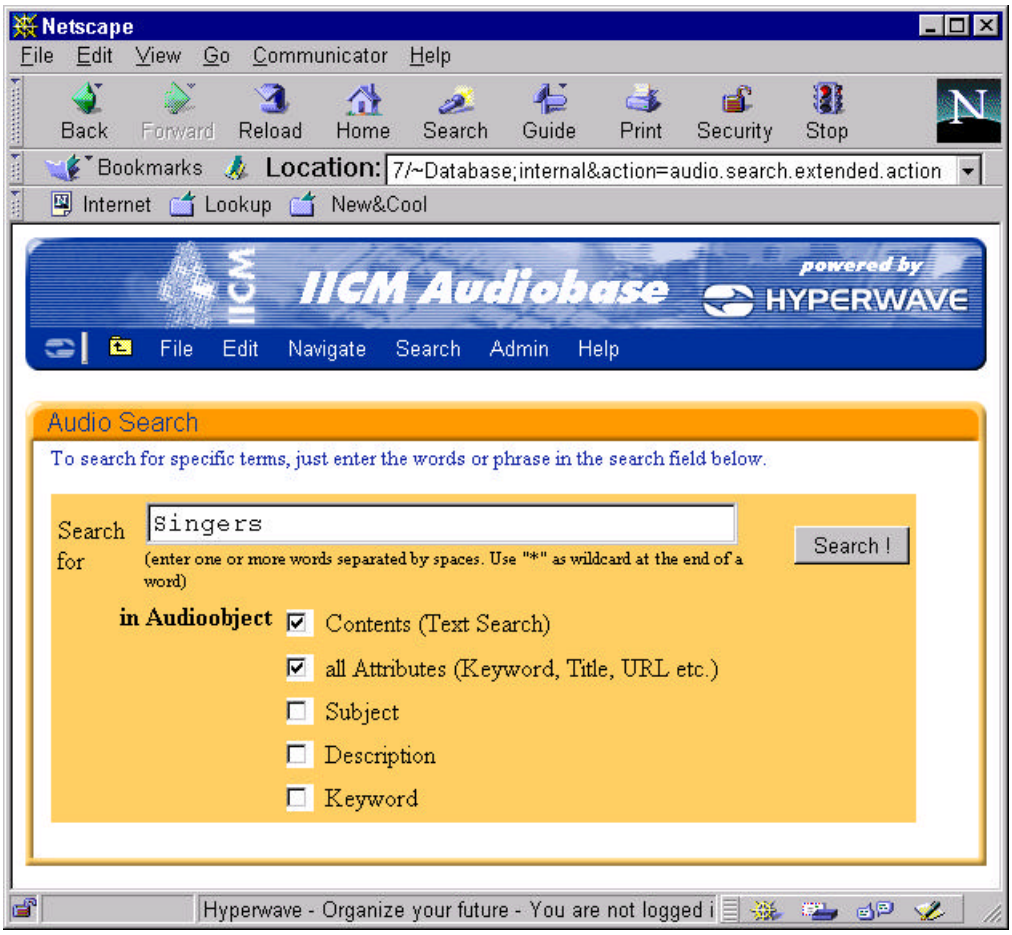

Abb. 7.1.2.a Screenshot der Audiosuche.

Nach Starten der Suche werden die relevanten Audioobjekte angezeigt. Durch anklicken einer im Suchergebnis aufgelisteten Aufnahme gelangt man in das Infofenster (siehe Kapitel 7.1.3). Durch anklicken des „Refine Query“ Buttons kann man neuerlich eine - verfeinerte - Suche starten (siehe Abb. 7.1.2.b).

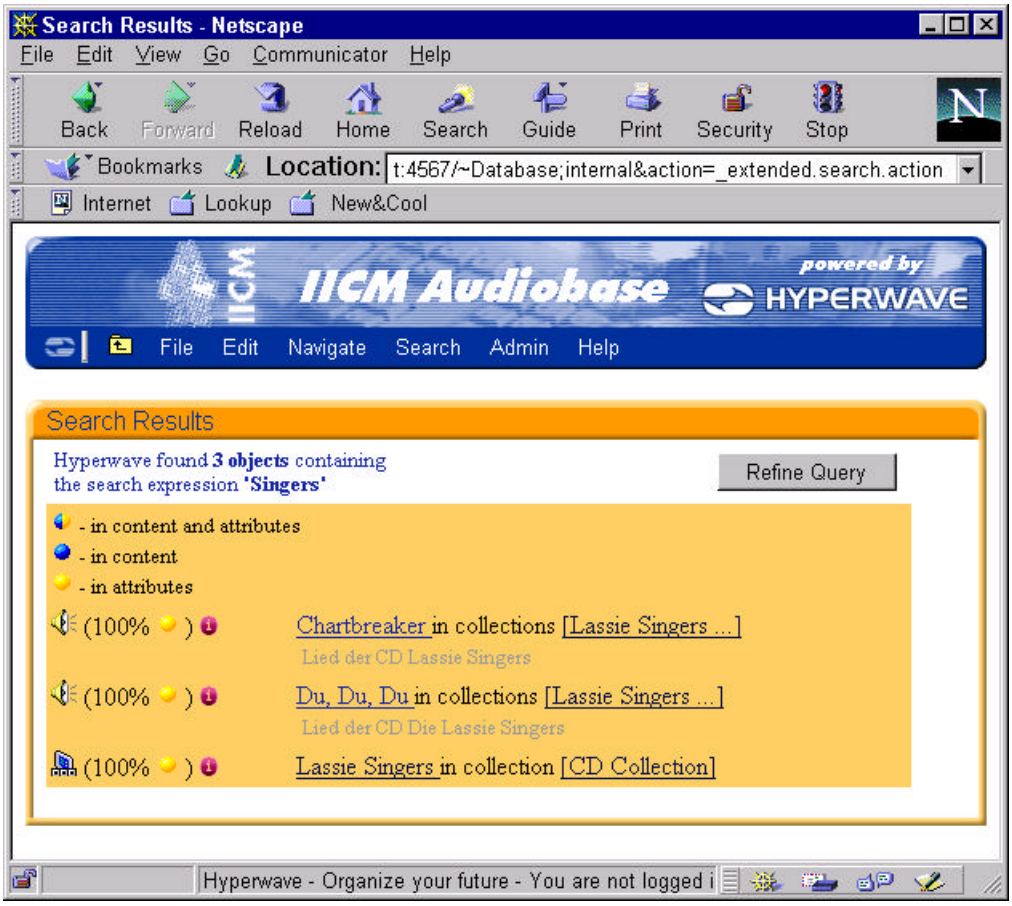

Abb. 7.1.2.b Screenshot des Audiosuchergebnisses.

7.1.3 Infofenster

Zur Anzeige der Metadaten einer gewählten Aufnahme wird ein neues Fenster - das Infofenster – geöffnet (siehe Abb. 7.1.3.1). Dadurch hat der Nutzer gleichzeitig sowohl eine Übersicht über den Datenbestand mit allen Aufnahmen als auch Detailinformation zur gewählten Aufnahme.

7.1.3.1 Infoansicht

Die Infoansicht zeigt alle Metadaten der gewählten Aufnahme, weiters ist das Vorhören der Aufnahme und das Abspeichern der Aufnahme im eigenen Filesystem in verschiedenen Formaten möglich.

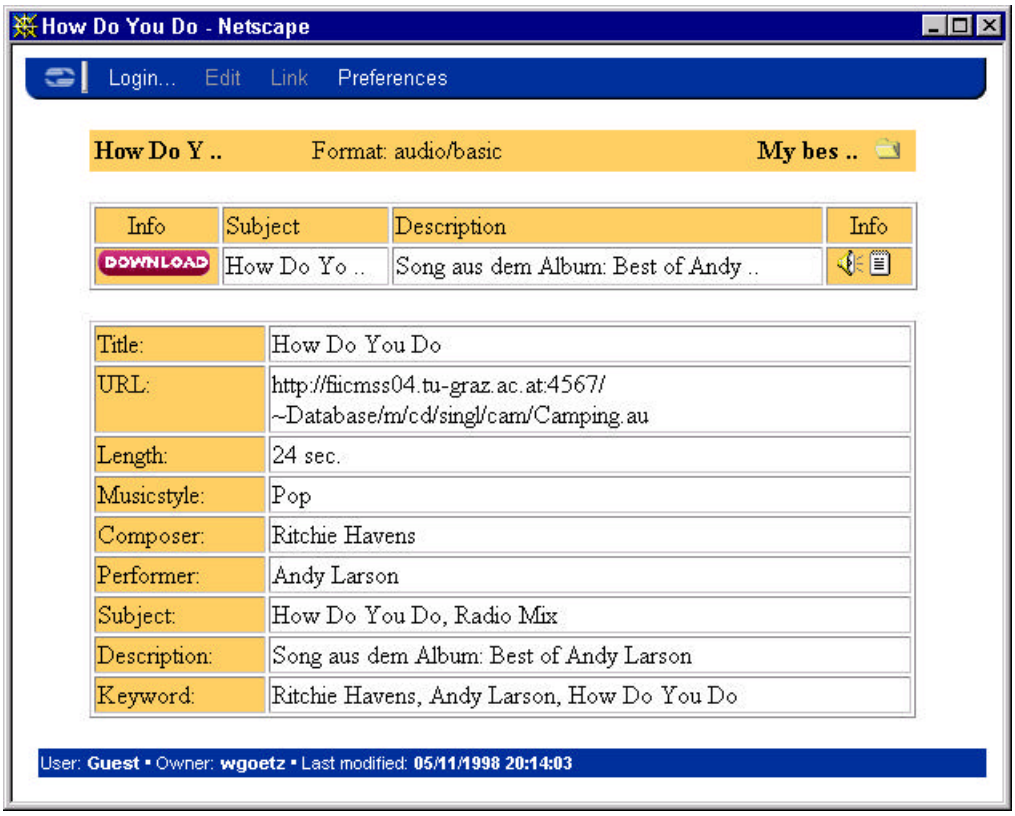

Abb. 7.1.3.1 Screenshot der Infoansicht.

Die Infoansicht zeigt im oberen Bereich das Audiouserinterface. Für nicht eingeloggte Nutzer ist nur der Menüpunkt „Preferences“ bzw. „Einstellungen“ aktiv, der Nutzer kann hier seine persönlich bevorzugten Einstellungen vornehmen.

Preferences:

- **Language**: Auswahl der Sprache des Userinterfaces und der Metaattribute. Mögliche Einstellungen sind Deutsch und Englisch.

- **Quality:** Auswahl des Downloadformates. Mögliche Einstellungen sind all jene Audioformate, in denen die Aufnahme vorliegt. In welchen Formaten die Aufnahme vorliegt, wird vom System geprüft.

- **Appearance:** Auswahl des Look & Feel des Hyperwave Userinterfaces. Mögliche Einstellungen sind die Standardlayouts Hyperwave und Windows.

Unter dem Audiouserinterface befindet sich der Objektstreifen. Er zeigt den Namen der gewählten Aufnahme (Audioobjekt), das aktuelle Downloadformat und den Namen der Collection, in der sich die gewählte Aufnahme befindet. Durch anklicken des Collectionsymbols wird das Infofenster geschlossen.

Unter dem Objektstreifen befindet sich die Kurzübersicht der gewählten Aufnahme. In den Auswahlfeldern der Kurzübersicht stehen folgende Buttons zur Verfügung:

- **Download Button**: DOWNLOAD
  Mit dem Download Button kann die Aufnahme im ausgewählten Format im eigenen Filesystem abgespeichert werden. Sobald die Maus über dem Button ist, wird die Dateigröße der Aufnahme im ausgewählten Format angezeigt. Durch anklicken des Buttons wird die Datei je nach Einstellung des Browsers gespeichert oder geöffnet.

- **Player Button:**
  Mit dem Player Button kann der Audioplayer in einem neuen Fenster gestartet werden (siehe Kapitel 7.1.4).

- **Text Button:**
  Mit dem Text Button kann zur Text Ansicht der Aufnahme gewechselt werden (siehe Kapitel 7.1.3.2).

Im unteren Bereich der Info Ansicht werden alle Metadaten der gewählten Aufnahme gezeigt.

7.1.3.2 Textansicht

Die Textansicht zeigt den zur Aufnahme gehörenden abgespeicherten Text (siehe Abb. 7.1.3.2). Dies kann sowohl der genaue Wortlaut der Aufnahme (z.B. Liedtext, Hörspieltext) als auch ein beschreibender Text sein.

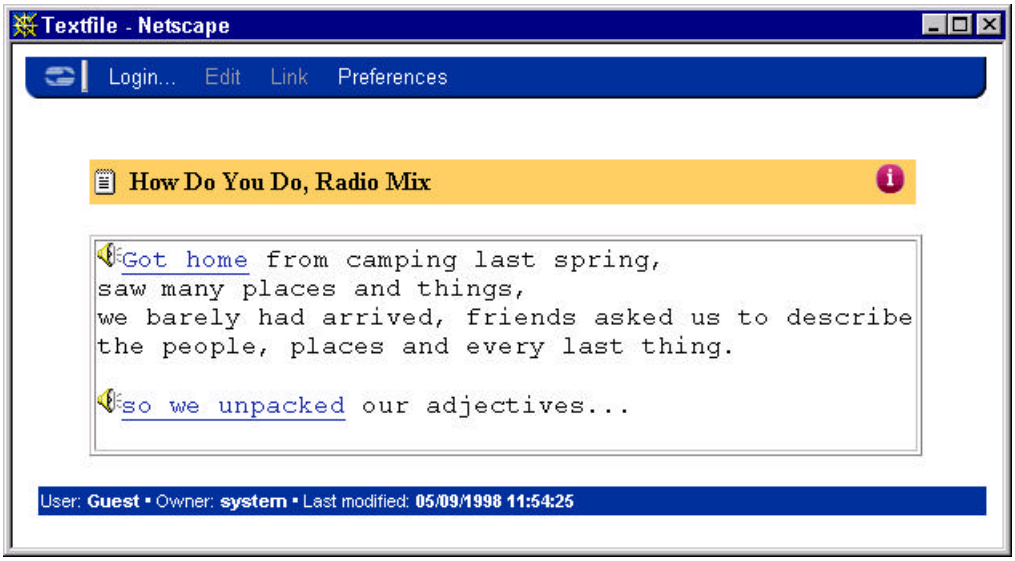

Abb. 7.1.3.2 Screenshot der Textansicht.

Die Textansicht zeigt im oberen Bereich das Audiouserinterface (siehe Kapitel 7.1.3.1). Darunter befindet sich der Objektstreifen. Er zeigt den Namen der gewählten Aufnahme und den Infoicon. Durch anklicken des Infoicons wird zur Infoansicht gewechselt.

Unter dem Objektstreifen befindet sich der zur Aufnahme gehörende Text. Der Text kann Links und Audiolinks enthalten. Durch anklicken eines Links gelangt man zu

weiterführender Zusatzinformation, durch anklicken eines Audiolinks wird der Audioplayer (siehe Kapitel 7.1.4) in einem neuen Fenster gestartet.

7.1.4 Audioplayer

Der Audioplayer beginnt sofort nach Aufruf aus der Infoansicht oder der Textansicht mit der Wiedergabe der Aufnahme (siehe Abb. 7.1.4). Im oberen Bereich des Fensters wird die aktuelle Zeitposition angegeben, darunter befinden sich die Buttons zur Steuerung der Wiedergabe.

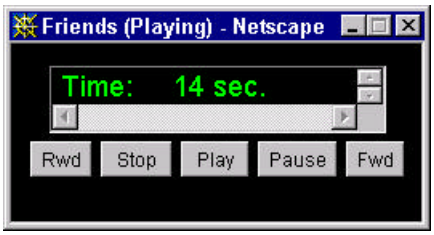

Abb. 7.1.4 Screenshot des Audioplayer.

Beim Start des Audioplayers aus der Textansicht erfolgt die Wiedergabe des Audiolinks vom vordefinierten Einstiegspunkt bis zum vordefinierten Ausstiegspunkt. Mit Hilfe der Steuerbuttons ist jedoch ein Vorhören der gesamten Aufnahme möglich.

7.2 Administration der Testimplementation

Die Testimplementation bietet dem eingeloggten Datenbankadministrator umfangreiche Möglichkeiten zum Erweitern, Warten und Löschen der Datenbestände. Der Aufbau der Datenbankhierarchie erfolgt mit der unter Hyperwave üblichen Funktionalität. Zum Einfügen von Aufnahmen in die Datenbank wurde das Hyperwave Userinterface um den Menüeintrag „New Audioobject“ erweitert. Im Audiouserinterface sind die Menüeinträge „Edit“ bzw. „Bearbeiten“ und „Link“ aktiv. Audiolinks können mit dem Audiolink Wizard eingefügt werden.

7.2.1 Aufbau der Datenbankhierarchie

Der Aufbau der Datenbankhierarchie erfolgt mit der Hyperwave Funktionalität, eine Unterteilung in mehrere Bereiche ist möglich. Bereiche, Gruppen und Untergruppen können je nach Anwendungsfall in dieser Testapplikation als Hyperwave Collection oder als Hyperwave Sequence angelegt werden, sinnvollerweise wird man Aufnahmen die in einer vorgegebenen Reihenfolge zueinander stehen in eine Sequence einfügen (z.B. die Tracks einer CD), Aufnahmen ohne vorgegebene Reihenfolge zueinander in eine Collection.

7.2.2. Einfügen einer Aufnahme

Eine Aufnahme wird in die Datenbank als Audioobjekt eingefügt. Ein Audioobjekt ist eine Collection die mindestens eine Audiodatei enthält. Da der Audioplayer (siehe Kapitel 7.1.4) nur bestimmte Audiodateien abspielen kann, muß eine Audiodatei in dieser Collection folgendes Format haben (siehe Kapitel 4.2.5.4):

- Format: SND
- MIME: audio/basic
- Samplerate: 8 kHz
- Kodierung: µ-law

Will man die Aufnahme in verschiedenen Audioformaten anbieten, muß die Collection zusätzlich die entsprechenden Audiodateien beinhalten. Die Metadaten der Aufnahme werden als zusätzliche Attribute in dieser Collection abgespeichert, das Attribut „Text“ wird - weil es auch Audiolinks enthalten kann - als Textfile in die Collection eingefügt.

Beim Einfügen einer Aufnahme kann man zwischen drei verschiedenen Metadatensets wählen. Diese Metadatensets sind für verschiedene Arten von Aufnahmen konzipiert und entsprechen den in Kapitel 6.2.1 vorgeschlagenen Metadatensets für Musik, Sprache und Geräusch.

Zum komfortablen Einfügen von Audioobjekten wurde das Hyperwave Userinterface um den Menüeintrag „File/New/Audioobject“ erweitert, in dem man die Wahl zwischen „Music“ (Metadatenset für Musikaufnahmen, siehe Abb. 7.2.2.a), „Speech“ (Metadatenset für Sprachaufnahmen) und „Sample“ (Metadatenset für Geräuschaufnahmen) hat. Nach anklicken des gewünschten Menüeintrags wird man zum eingeben der Metadaten aufgefordert.

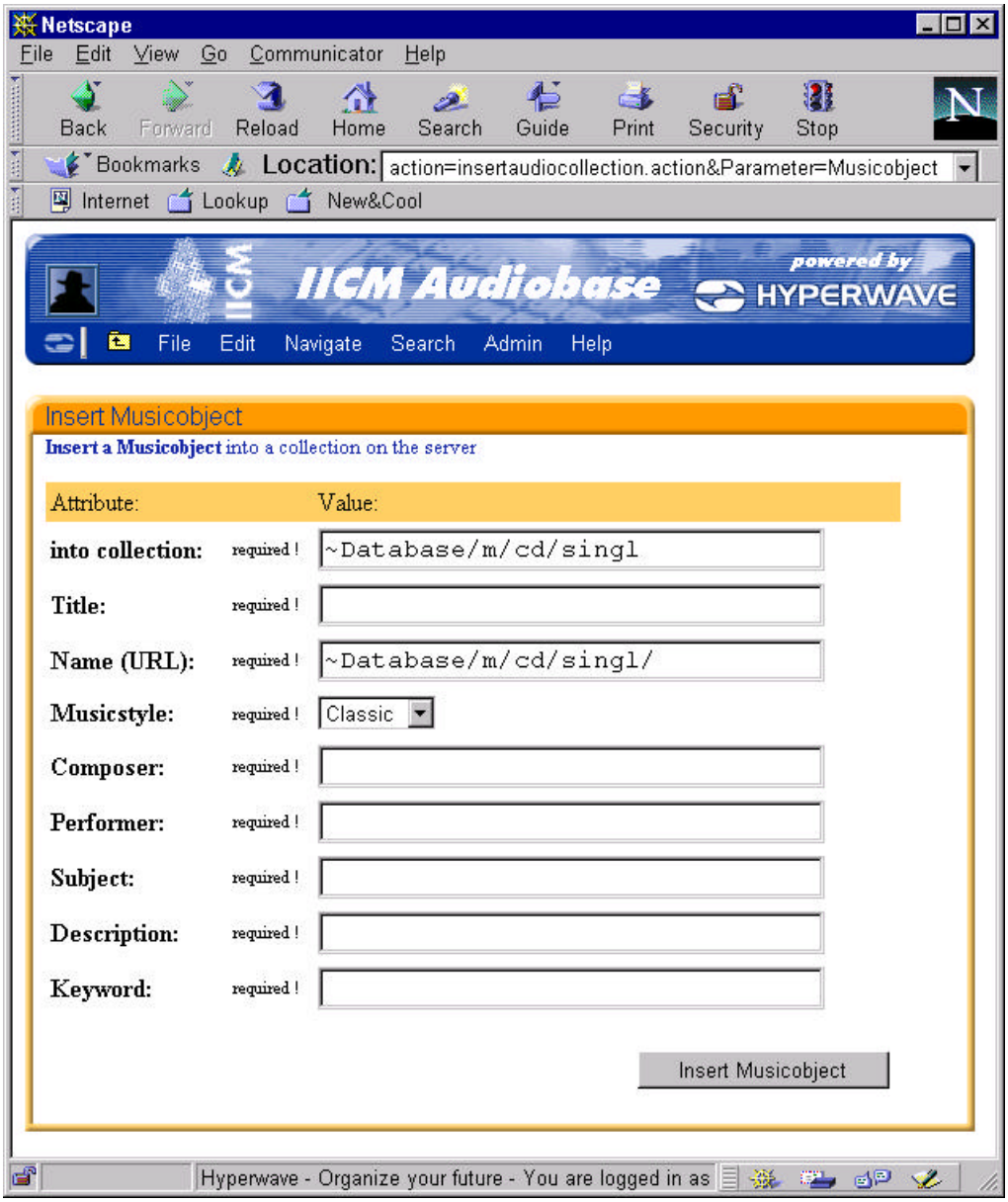

Abb. 7.2.2.a Screenshot: Einfügen eines Audioobjektes (Metadaten für Musikaufnahme).

Zum Einfügen des Metaattributes „Text“ wird man erst nach Einfügen der restlichen Metadaten aufgefordert (siehe Abb. 7.2.2.b).

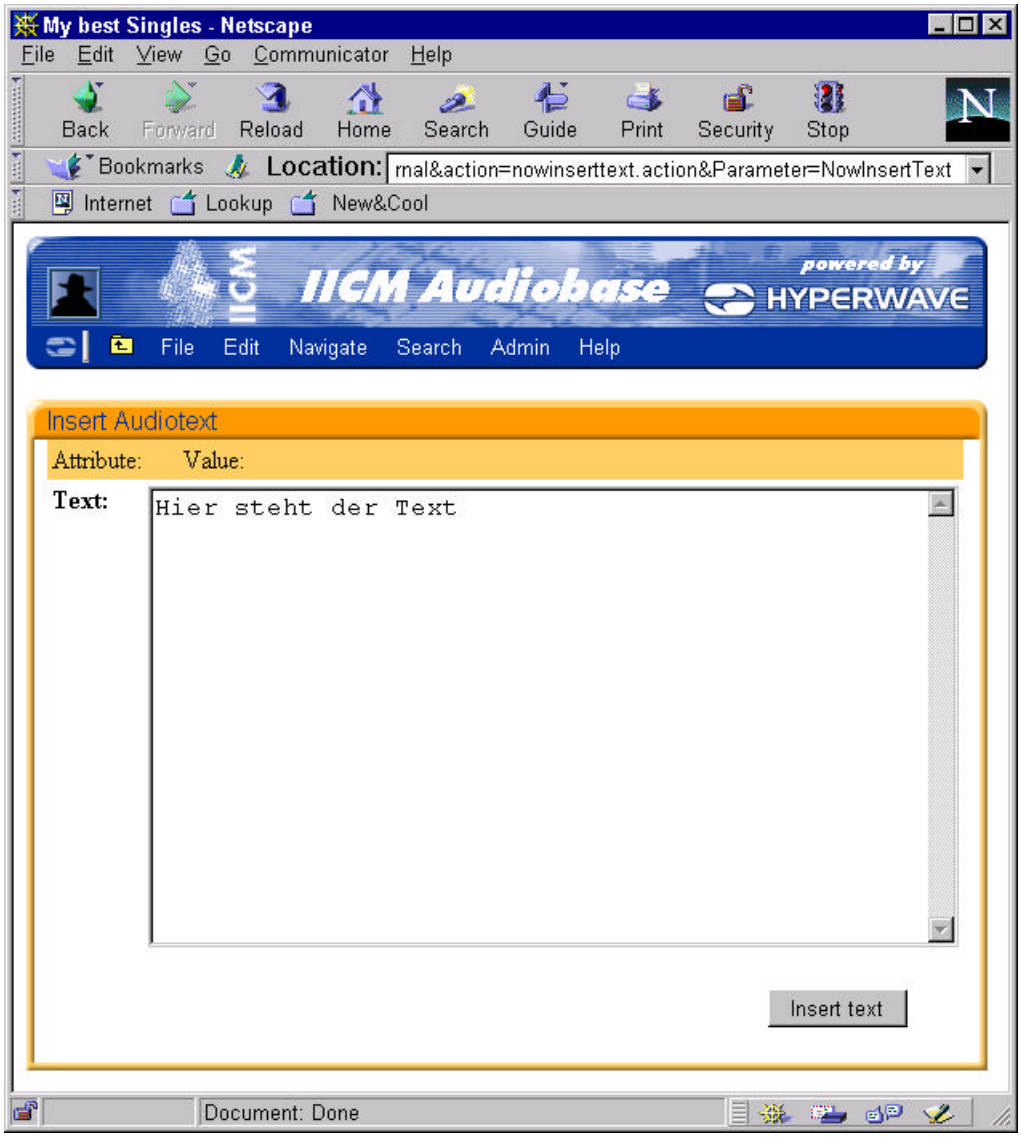

Abb. 7.2.2.b Screenshot: Einfügen eines Audioobjektes (Metaattribut „Text“).

Nach Einfügen des Metaattributes „Text“ wird man zum Einfügen der Audiodateien aufgefordert (siehe Abb. 7.2.2.c). Es können beliebig viele Audiodateien und auch andere Daten eingefügt werden, allerdings muß eine Audiodatei den in Kapitel 7.2.2 genannten Anforderungen entsprechen.

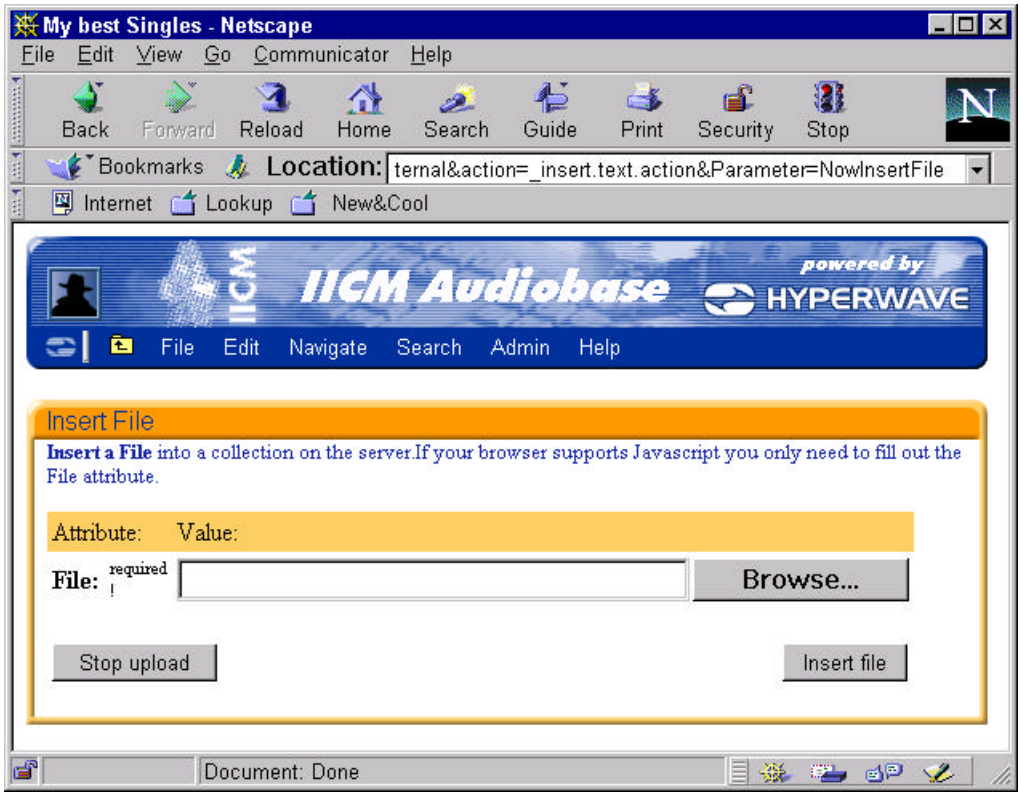

Abb. 7.2.2.c Screenshot: Einfügen eines Audioobjektes (Audiodatei).

Nach Einfügen der gewünschten Audiodateien ist die Aufnahme in der Audiodatenbank abgelegt.

7.2.3 Editieren des Metadatensets

Im Audiouserinterface der Info Ansicht ist der Menüeintrag „Edit/Attributes“ aktiv, mit ihm wird das Edit-Metadaten-Fenster geöffnet (siehe Abb. 7.2.3).

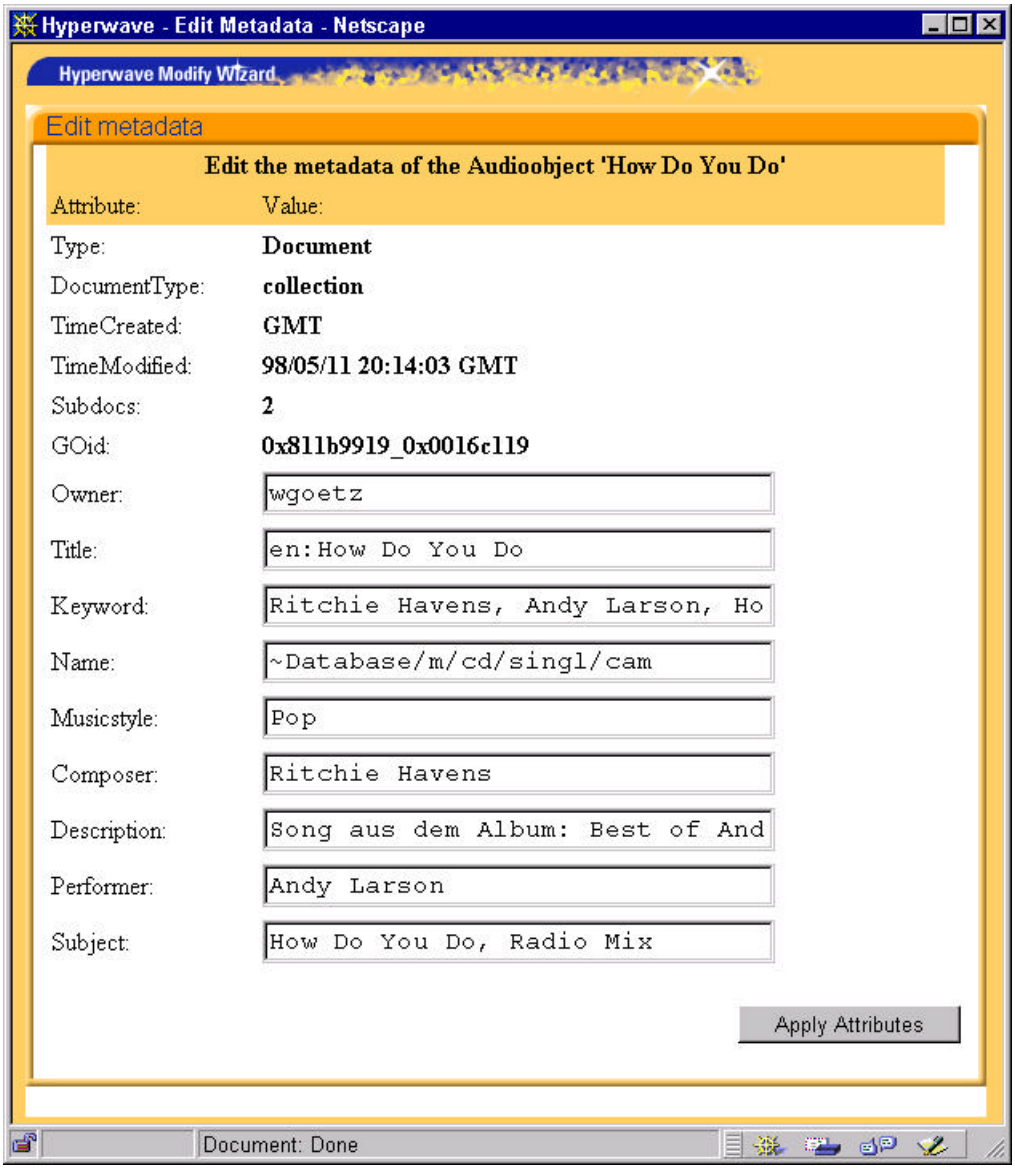

Abb. 7.2.3 Screenshot des Edit-Metadaten-Fensters.

Die Metadaten können verändert werden; nach dem Anwenden der Änderungen schließt sich das Fenster wieder.

7.2.4 Editieren des Metaattributes „Text“

Im Audiouserinterface der Text Ansicht ist der Menüeintrag „Edit/Text“ aktiv, mit ihm wird das Edit-Text-Fenster geöffnet (siehe Abb. 7.2.4).

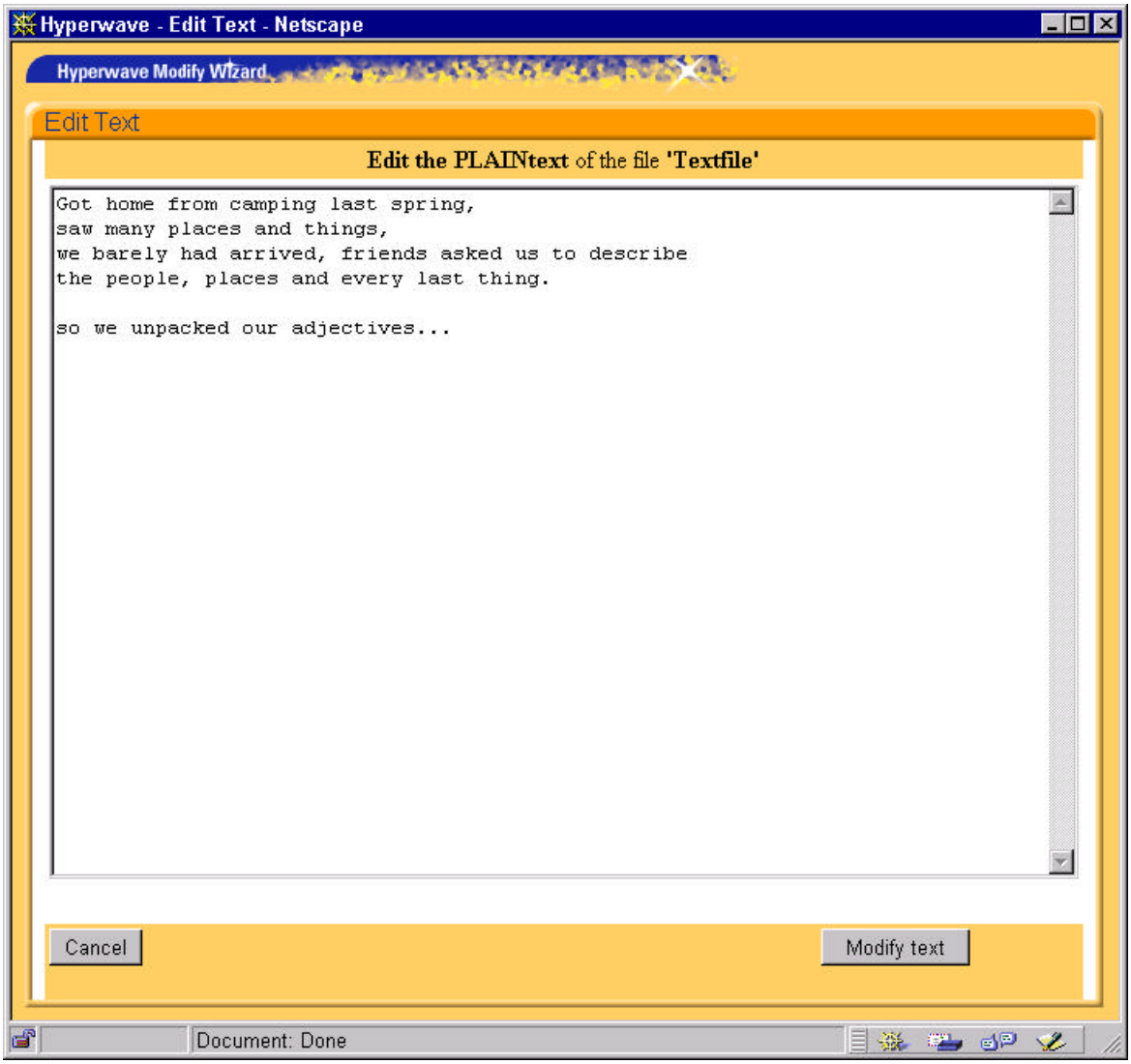

Abb. 7.2.4 Screenshot des Edit-Text-Fensters.

Der Text kann verändert werden; nach dem Anwenden der Änderungen wird das Fenster wieder geschlossen.

7.2.5 Einfügen eines Audiolinks

Im Audiouserinterface der Textansicht ist der Menüeintrag „Link/New Audiolink“ aktiv, mit ihm wird der Audiolink-Wizard geöffnet (siehe Abb. 7.2.5).

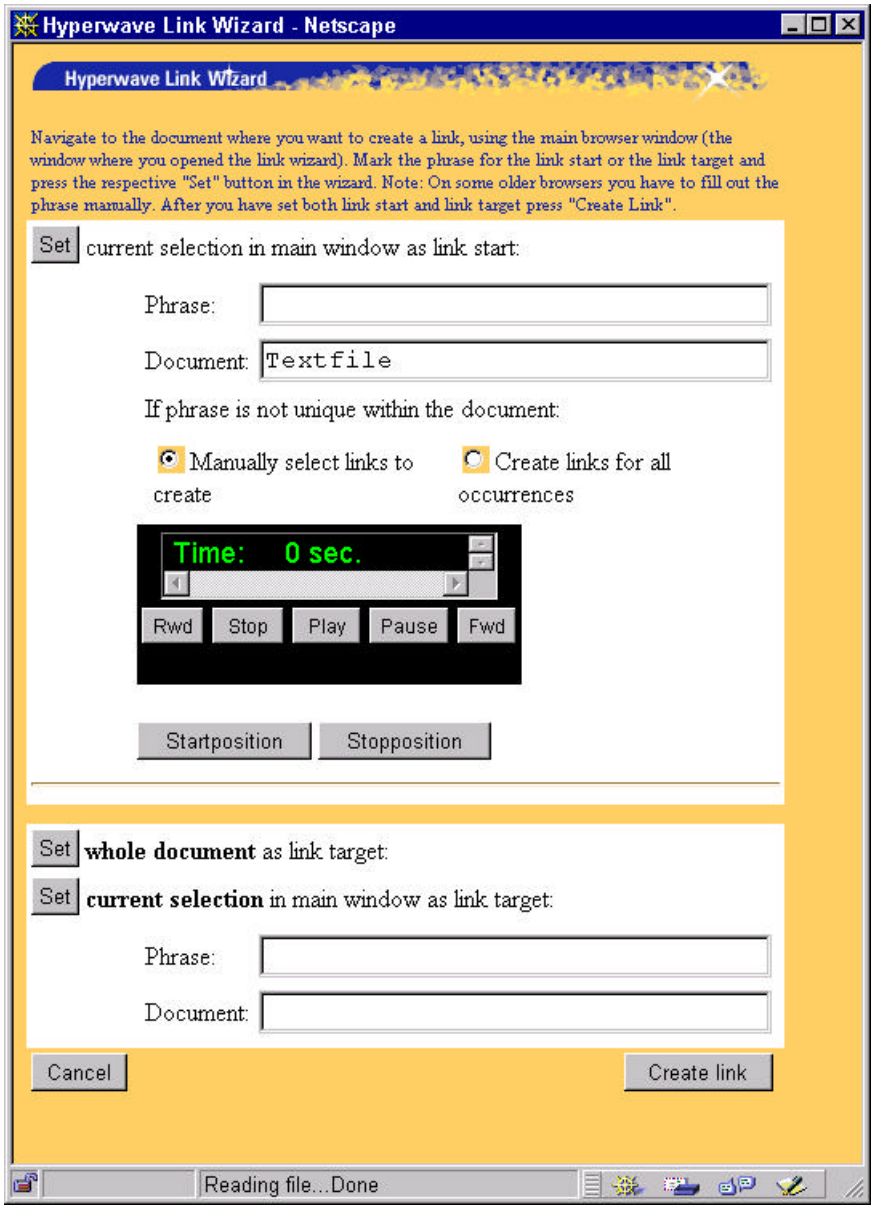

Abb. 7.2.5 Screenshot des Audiolink-Wizards.

Nach setzen der Linkposition im Text - wie bei Hyperwave üblich - wird die Start- und Stopposition des zugeordneten Teilbereichs der Aufnahme festgelegt. Dazu wird mit dem in den Wizard integrierten Audioplayer an die gewünschte Stelle navigiert und dann der „Startposition“- bzw. „Stopposition“-Button gedrückt. Durch klicken auf den „Create Link“-Button wird der Audiolink eingefügt.

7.2.6 Löschen eines Audiolinks

Im Audiouserinterface der Text Ansicht ist der Menüeintrag „Link/Delete Audiolink“ aktiv, mit ihm wird der Modify-Wizard geöffnet. Mit dem Modify-Wizard können Audiolinks und normale Hyperwave Links gelöscht werden (siehe Abb 7.2.6).

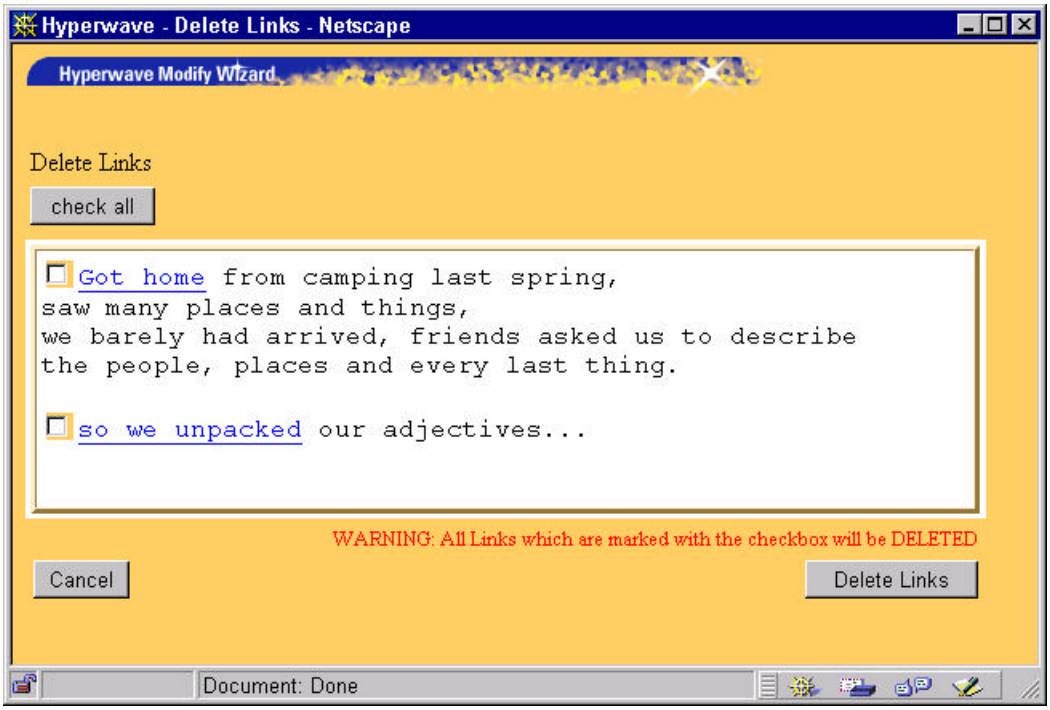

Abb. 7.2.6 Screenshot des Modify-Wizard.

7.3. Funktionsweise der Testimplementation

Veränderungen bezüglich des Aussehens und der Funktionalität des Hyperwave Userinterfaces wurden durch Ändern der Templates vorgenommen. Diese Änderungen betreffen unter anderem das Hinzufügen neuer Menüeinträge in das Hyperwave Userinterface (siehe Kapitel 7.2.2), die geänderte Darstellung von Audioobjekten in der Übersichtsdarstellung (siehe Abb. 7.1.1) und in der Darstellung eines Audiosuchergebnisses (siehe Abb. 7.1.2.b). Außerdem betreffen einige Änderungen in den Templates die Verzweigung zu neuen Templates. Tabelle 7.3.a zeigt die adaptierten Templates:

Template	Beschreibung
`hwinclude.html`	Includiert alle Templates, neue Templates wurden eingefügt
`master.html`	Detektiert die Actions, neue Actions wurden eingefügt
`header.html`	Baut den neuen Kopfbereich auf
`body.html`	Verzweigt zu Template `audiobody.html`
`header_menu.html`	Verzweigt zu Template `audiomenu.html`
`treeview.html`	Stellt Baumstruktur dar, Icondarstellung wurde geändert
`exsearchresults.html`	Stellt Suchergebnis dar, neue Darstellung wurde hinzugefügt

Tab. 7.3.a Adaptierte Templates

Umfangreichere Veränderungen wurden durch Einbindung neuer Templates realisiert (siehe Tab. 7.3.b). Mit Hilfe dieser neuen Templates werden unter anderem die Infoansicht (siehe Abb. 7.1.3.1), die Textansicht (siehe Abb. 7.1.3.2), die Einfügemasken (siehe Abb. 7.2.2.a, Abb. 7.2.2.b und Abb. 7.2.2.c) und die Edit-Fenster (siehe Abb. 7.2.3 und Abb. 7.2.4) dargestellt. Tabelle 7.3.b zeigt die zusätzlichen Templates:

Template	Beschreibung
`audiomenu.html`	Definiert die Funktionen des Audiouserinterfaces
`audiobody.html`	Stellt die Infoansicht und die Textansicht dar
`editaudioattr.html`	Stellt das Edit-Metadaten-Fenster dar
`insertaudiocollection.html`	Stellt das Einfügen-Audioobjekt-Fenster dar
`insertaudiotext.html`	Stellt das Einfügen-Audiotext-Fenster dar
`insertaudiofile.html`	Stellt das Einfügen-Audiolink-Fenster dar
`audiosearchform.html`	Stellt das Audiosuche-Fenster dar
`mkaudiolinks.html`	Stellt das Audiolink-Fenster dar

Tab. 7.3.b Zusätzliche Templates

Der Audioplayer wurde als Java-Applet implementiert. Um eine korrekte Funktion des Audioplayers zu gewährleisten, müssen die compilierten Classes `DJ.class` `Bais.class` und `Counter.class` auf dem Hyperwave Server vorhanden sein. Die aktuelle Zeitposition des Audioplayers wird in den Templates mit der Javascript-Variable `document.DJ.sendposition()` abgefragt. Die Start- und Stopposition einer Aufnahme wird beim Starten des Applets als Parameter Start bzw. Stop übertragen.

Die Verzweigung zu neuen Templates erfolgt in Abhängigkeit eines Attributes oder in Abhängigkeit einer definierten Action. Sobald eine Collection ein Attribut „Subject“ hat, wird sie als Audioobjekt behandelt (siehe Kapitel 7.2.2) und mit Hilfe der neuen Templates dargestellt. Die Verzweigung erfolgt im Template `body.html` und wurde folgendermaßen implementiert:

```
%%if object.get_attrib(Subject)%%
     %%audiobody%%
%%else%%
     %%show_document%%
%%endif%%
```

das Prinzip von Attributabfragen und Detektion von Actions erlaubt den Eingriff in die Hyperwave Standardfunktionalität. Sobald eine Action aufgerufen wird, erfolgt eine Verzweigung zum entsprechenden Template, z.B. wurde die Verzweigung zum Template `mkaudiolinks` aus dem Template `master.html` folgendermaßen implementiert:

```
%%if action==‘audiolink.action‘%%
     %%mkaudiolinks%%
%%endif%%
```

Die Verzweigungen zu anderen Templates erfolgen in ähnlicher Weise (siehe Abb. 7.3). Die mit einem Request an den Hyperwave Server übermittelte Action wird detektiert und der Server reagiert entsprechend der implementierten Funktionalität.

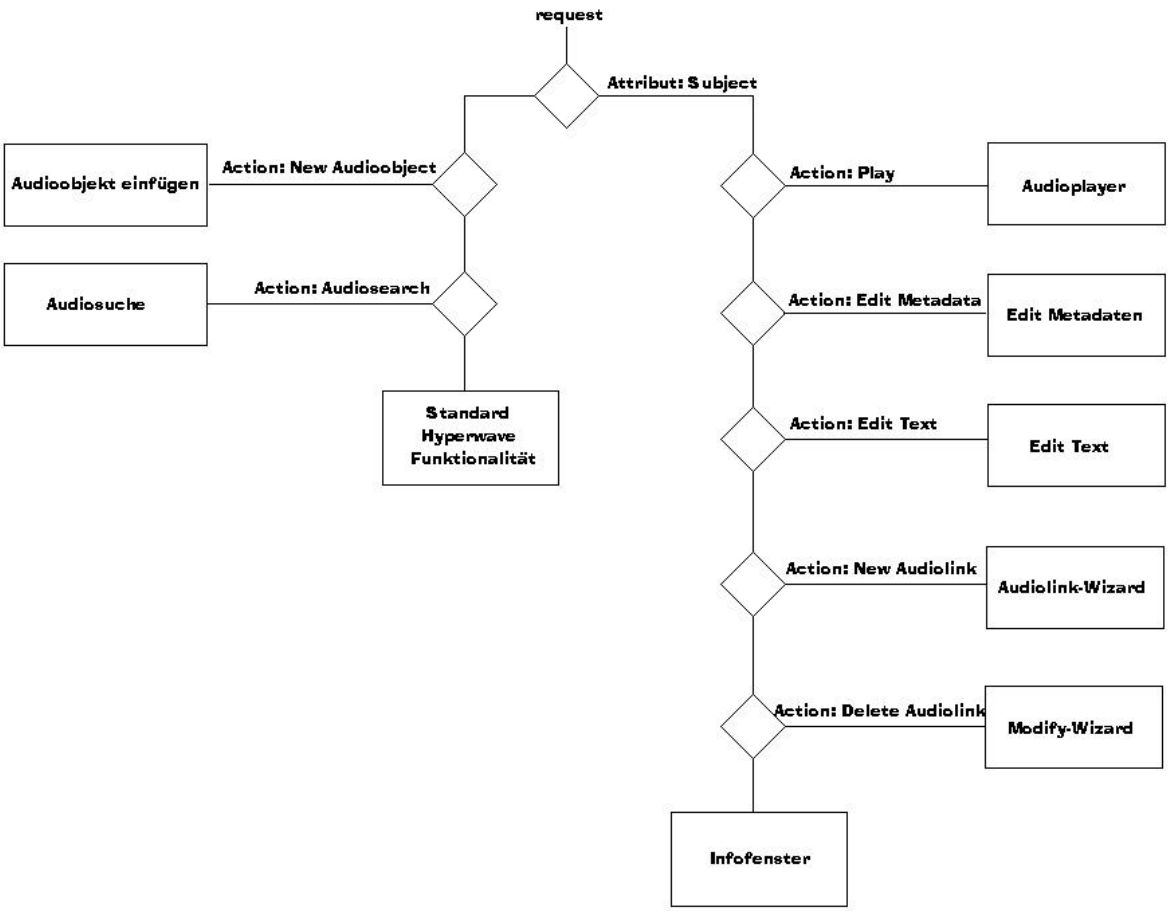

Abb. 7.3 Verzweigungen zu den Templates

8. Zusammenfassung und Ausblick

8.1 Zusammenfassung

Im Rahmen dieser Diplomarbeit wurde eine Audiodatenbank auf dem Hyperwave Information Server implementiert. Dabei wurde die Funktionalität des Hyperwave Servers um Funktionen zur komfortablen Verwaltung von Audiodaten erweitert. Der Anwender hat dadurch vielfältige Möglichkeiten bei der Nutzung der Audiodatenbank. Er kann mit Hilfe von Metadatensets und Audiolinks einen raschen Überblick über den Datenbestand gewinnen, die für ihn interessanten Audiodaten in verschiedenen Formaten in das lokale Filesystem laden und - falls er entsprechende Rechte hat - den Datenbestand komfortabel erweitern und warten. Navigieren durch den Datenbestand ist ebenso möglich wie Suchen nach bestimmten Datensätzen. Der Einsatz von Audiolinks eröffnet eine Vielzahl von Anwendungen, eine Verwendung von Audiolinks ist nicht nur im Bereich einer Audiodatenbank sondern auch in vielen anderen Einsatzgebieten (z.B. Medizintechnik, Gerichtsakten) vorstellbar. Das mögliche Einspielen von Zusatzinformation (Text, Bild, Video) ermöglicht ebenfalls eine Verwendung in anderen Einsatzgebieten, und so ist die im Rahmen dieser Arbeit implementierte Audiodatenbank eine für viele Fachbereiche interessante Implementation. Der theoretische Teil der Arbeit informiert den Leser über den gegenwärtigen Stand zur Verwendung von Audiodaten im Internet, den Einsatz von Metadaten, Audiodatenreduktionstechniken und Audiostreamingtechniken.

8.2 Ausblick

Die Implementation der Audiodatenbank erfolgte mit einem im Rahmen der vorliegenden Arbeit vertretbaren Aufwand. Darüber hinaus existieren jedoch eine Reihe interessanter Aspekte, die im Rahmen einer weiteren Ausbaustufe enthalten sein könnten:

- **Automatische Formatkonvertierung:** Will man eine Aufnahme in verschiedenen Audioformaten anbieten, müssen alle dazu nötigen Audiodateien in die Audiodatenbank eingefügt werden. Eine interessante Erweiterung wäre eine automatische Formatkonvertierung, bei der nach Einfügen der Audiodatei eine automatische Konvertierung dieser Datei in andere Audioformate erfolgt. Das manuelle Einfügen der Audiodatei in verschiedenen Audioformaten wäre dadurch überflüssig. Möglichkeiten zur Implementation dieser Formatkonvertierung werden zur Zeit im Rahmen eines EDV-Projektes am IICM geprüft.

- **Automatische Texterkennung:** Das Einfügen des Metaattributes „Text“ erfolgt bei dieser Implementation manuell. Es existieren jedoch bereits Algorithmen, die eine automatische Speech-to-Text-Konvertierung einer Audiodatei durchführen. Eine interessante Erweiterung wäre die Implementation eines solchen Algorithmus. Nach Einfügen einer Audiodatei könnte dann eine automatische Konvertierung in eine

Textdatei erfolgen, das manuelle Einfügen des Metaattributes „Text“ wäre dann nicht mehr nötig und eine inhaltliche Volltextsuche wäre damit möglich.

- **Automatische Erstellung von Audiolinks:** Die automatische Speech-to-Text-Konvertierung der Audiodatei würde auch eine automatische Erstellung von Audiolinks ermöglichen. Bei der Konvertierung müßte die Information über Audiobereich und den diesem Audiobereich zugeordneten Textbereich als Linkobjekt in den Hyperwave Server eingefügt werden, das manuelle Einfügen von Audiolinks wäre dann nicht mehr nötig.

- **Audioannotationen:** In der vorliegenden Arbeit ist das Metaattribut „Text“ eine Zusatzinformation zu einer Aufnahme in der Audiodatenbank; die Aufnahme ist sozusagen die Hauptinformation, das Metaattribut „Text“ mit den Audiolinks ist nur Zusatzinformation. Eine umgekehrte Sichtweise (d.h. die Aufnahme ist eine Zusatzinformation zum Text) bietet den Zugang zu vielen neuen Anwendungsmöglichkeiten. Ein Textdokument könnte z.B. Audioannotationen enthalten, die man jederzeit mit Hilfe eines Audiowizards aufnehmen und in den vorhandenen Text einfügen kann; z.B. für Anmerkungen zu einem medizinischen Befund, Gerichtsakt, etc.

Literaturverzeichnis

[aeiou.at]	http://www.aeiou.at/
[arts.gla]	http://www.arts.gla.ac.uk/PADS/projects/
[asia.microsoft]	http://www.asia.microsoft.com/asf/
[bbc.co]	http://www.bbc.co.uk/
[echo.lu]	http://www2.echo.lu/oii/en/
[ircam.fr]	http://www.ircam.fr/
[intertrust.com]	http://www.intertrust.com/
[ix.de]	http://www.ix.de/newsticker/data/
[liquid]	http://www.liquidaudio.com/
[microsoft.com]	http://www.microsoft.com/
[mpib-berlin]	http://www.mpib-berlin.mpg.de/dok/metadata/
[musicblvd.com]	http://www.musicblvd.com/
[nw.com]	http://www.nw.com/zone/WWW/
[purl.org]	http://purl.org/metadata/dublin_core/
[real.com]	http://www.real.com/
[sil.org]	http://www.sil.org/sgml/
[subaudio.net]	http://www.subaudio.net/
[uu.net]	http://www.uu.net/lang.de/products/
[vivo.com]	http://www.vivo.com/
[wired.com]	http://www.wired.com/wired/3.06/features/Xanadu/
[wwwots.let.ruu]	http://www-uilots.let.ruu.nl/Audiodatabase/

[CASSELBERRY97]	Casselberry, R.: Das perfekte Intranet. Markt und Technik, München 1997
[COMPWO96]	Computerwoche, Nr. 22, München 1996
[GÜTL98]	Gütl, C.: Future Information Harvesting and Processing on the Web, Conference "European Telematics: advancing the information society" Barcelona, 4-7 February 1998
[DEMPSEY97]	Dempsey, L.: Metadata: An Overview of Current Resource Description Practice, Journal of Documentation 1997
[DUFFY97]	Duffy C.: Resource Discovery Workshops: Sound Resources Report from the Performing Arts Data Service - Final draft, Glasgow 1997
[HABERFEL92]	Haberfellner, R.: Systems Engineering. Verlag Industrielle Organisation, Zürich 1992
[HYPERWAVE98]	Hyperwave Programmer‘s Guide, Hyperwave Information Management GmbH, München 1998
[KROL97]	Krol, E.: The whole Internet. O’Reilly & Associates, Inc., Sebastopol 1997
[KYAS96]	Kyas, O.: Internet professionell. International Thomson Publishing, Bonn 1996

[MAURER96] Maurer, H.: Hyper-G now Hyperwave. Addison-Wesley Publishing Company, Harlow 1996

[MEYER91] Meyer-Wegener, K.: Multimedia-Datenbanken. Teubner, Stuttgart 1991

[RICHTER94] Richter, U.: WWW-Begriffe. Friedrich-Schiller-Universität Jena 1994

[RUSCH-FEJA97] Rusch-Feja, D.: Dublin Core Metadata. Bibliotheksdienst 31, Berlin 1997

[SCHÖNFELDER88] Schönfelder, H. (Hrsg.): Digitale Filter in der Videotechnik. Drei-R-Verlag, Berlin 1988

[SOCHATS96] Sochats, K.: Metadata Requirements for Evidence, Automating 21st Century Science - The Legal, Regulatory, Technical and Social Aspects of Electronic Laboratory Notebooks and Collaborative Computing, TeamScience, 1996

[WEIBEL95] Weibel, S.: Metadata: The Foundations of Resource Description. D-Lib Magazine, July 1995 http://www.dlib.org/dlib/July95/

[ZWICKER82] Zwicker, E.: Psychoakustik. Springer-Verlag, Berlin 1982

Abbildungsverzeichnis

Abb. 2.2 Weltweite Entwicklung von Hosts [nw.com]
Abb. 2.3.5 Hypertext [MAURER96]
Abb. 4.2 Digitalisierung eines analogen Audiosignals, Wolfgang Götzinger
Abb. 4.3.1 Kurven gleicher Lautstärkenempfindung [ZWICKLER82]
Abb. 4.3.1.1.a Mithörschwellen [ZWICKLER82]
Abb. 4.3.1.1.b Mithörschwellen [ZWICKLER82]
Abb. 4.3.1.1.c Temporäre Verdeckung [ZWICKLER82]
Abb. 4.3.3.1 Datenrahmen MPEG II [SCHÖNFELDER88]
Abb. 5.1. Screenshot Music Boulevard [musicblvd.com]
Abb. 5.2. Screenshot IRCAM [ircam.fr]
Abb. 5.3. Screenshot Subaudio [subaudio.net]
Abb. 5.4. Screenshot Universität Utrecht [wwwots.let.ruu]
Abb. 6.2.3 Textversion mit Links, Wolfgang Götzinger
Abb. 6.2.4.1.a Textversion mit Audiolinks, Wolfgang Götzinger
Abb. 6.2.4.1.b Textversion mit Audiolinks, Wolfgang Götzinger
Abb. 7.1.1 Screenshot Übersichtsdarstellung, Wolfgang Götzinger
Abb. 7.1.2.a Screenshot Sucheingabe, Wolfgang Götzinger
Abb. 7.1.2.b Screenshot Suchergebnis, Wolfgang Götzinger
Abb. 7.1.3.1 Screenshot Infoansicht, Wolfgang Götzinger
Abb. 7.1.3.2 Screenshot Textansicht, Wolfgang Götzinger
Abb. 7.1.4 Screenshot Audioplayer, Wolfgang Götzinger
Abb. 7.2.2.a Screenshot Einfügen Audioobjekt, Wolfgang Götzinger
Abb. 7.2.2.b Screenshot Einfügen Audioobjekt, Wolfgang Götzinger
Abb. 7.2.2.c Screenshot Einfügen Audioobjekt, Wolfgang Götzinger
Abb. 7.2.3 Screenshot Edit-Metadaten, Wolfgang Götzinger
Abb. 7.2.4 Screenshot Edit-Text, Wolfgang Götzinger
Abb. 7.2.5 Screenshot Audiolink-Wizard, Wolfgang Götzinger
Abb. 7.2.6 Screenshot Modify-Wizard, Wolfgang Götzinger
Abb. 7.3 Verzweigungen zu den Templates, Wolfgang Götzinger

Tabellenverzeichnis

Tab. 2.3.8 Speichermöglichkeiten im Überblick
Tab. 3.3.5 Beispiel Dublin Core Elemente-Set
Tab. 4.2.4 Datenraten digitaler Audiosysteme
Tab. 4.3.2 Datenraten datenreduzierter Audiosysteme
Tab. 6.2.2.1 Nicht konfigurierbares Metadatenset
Tab. 6.2.2.2 Teilweise konfigurierbares Metadatenset
Tab. 6.2.2.3 Voll konfigurierbares Metadatenset
Tab. 7.3.a Adaptierte Templates
Tab. 7.3.b Zusätzliche Templates