

Entwicklung einer World-Wide-Web-Ausgabe einer Tageszeitung

Dirk Kuhlmann

Studienarbeit
6.6.1996

Technische Universität Berlin
Fachbereich Informatik
Fachgebiet Wissensbasierte Systeme

Aufgabensteller: Professor Dr. E. Konrad
Mitbetreuer: Josef Willenborg, Stephan Fruehauf

| | | |
|-------|--|----|
| 1.0 | Einleitung | 1 |
| 2.0 | Publizieren im Informationszeitalter: Ein neues Paradigma? | 3 |
| 2.1 | Strategische Faktoren | 6 |
| 2.2 | Traditionelles und elektronisches Publizieren | 11 |
| 2.2.1 | Charakteristika traditioneller und elektronischer Zeitungen | 11 |
| 2.2.2 | Typologie von Online-Angeboten für Zeitungen | 13 |
| 2.2.3 | Informationeller Mehrwert durch hypertextgestützte Zeitungsausgaben? | 16 |
| 3.0 | Problemkontext | 19 |
| 3.1 | Das World Wide Web | 19 |
| 3.1.1 | Das Dokumentenmodell des WWW | 20 |
| 3.1.2 | Das Client-Server Modell des WWW | 24 |
| 3.1.3 | Die Strukturbeschreibungssprache HTML | 27 |
| 3.2 | Organisatorische und technische Rahmenbedingungen | 29 |
| 3.3 | Ziele des Projekts | 30 |
| 4.0 | Systementwurf | 33 |
| 4.1 | Gestaltung des Informationsangebots | 33 |
| 4.1.1 | Strukturierung der Dokumente | 33 |
| 4.1.2 | Layout des Informationsangebots | 37 |
| 4.2 | Entwurf der technischen Komponenten | 41 |
| 4.2.1 | Dateitransfer und Spiegelung | 41 |
| 4.2.2 | Formatkonvertierung | 42 |
| 4.2.3 | Authentifizierung | 44 |
| 4.2.4 | Leserbriefe und Diskussionsforum | 44 |
| 4.2.5 | Stichwortsuche | 46 |
| 4.2.6 | Statistische Auswertung der Logdateien | 46 |
| 5.0 | Systemimplementation | 49 |
| 5.1 | Dateitransfer und Spiegelung | 49 |
| 5.2 | Formatkonvertierung | 50 |
| 5.2.1 | HTML-Konverter | 50 |
| 5.2.2 | Strukturgenerator | 53 |
| 5.3 | Authentifizierung | 56 |
| 5.4 | Leserbriefe und Diskussionsforum | 58 |
| 5.5 | Stichwortsuche | 59 |
| 6.0 | Prototyping und Test | 61 |
| 6.1 | Prototypingzyklen | 61 |
| 6.2 | Tests | 64 |
| 7.0 | Ergebnisse und Auswertung | 67 |
| 7.1 | Ergebnisse der Leserumfrage | 67 |
| 7.1.1 | Internetanbieter und Zugriffsort | 67 |
| 7.1.2 | Abonnementsbereitschaft | 68 |
| 7.1.3 | Finanzieller Wert | 68 |
| 7.1.4 | Vor- und Nachteile des Online-Angebots | 70 |
| 7.2 | Ermittlung des Bekanntheitsgrads | 72 |
| 7.3 | Auswertung | 73 |
| 7.3.1 | Auswertung der Leserumfrage | 74 |
| 7.3.2 | Auswertung der Prototypingzyklen und des laufenden Betriebs | 75 |
| 8.0 | Zusammenfassung und Ausblick | 77 |
| 9.0 | Quellen | 81 |
| | Anhang A: BNF-Grammatik des taz-Archivformats | 85 |
| | Anhang B: Liste der Systemmodule | 87 |
| | Anhang C: Wortlaut der Leserumfrage | 89 |

1.0 Einleitung

Der Markt der Online-Periodika ist in Deutschland in Bewegung geraten. Gaben noch 1994 die befragten Experten dem Experiment einer elektronischen Zeitung kaum eine Chance, so überschlugen sich seit dem Frühjahr 1995 entsprechende verlegerische Aktivitäten. Dies ist um so bemerkenswerter, als vielen Zeitungsverlegern in der BRD verlustreiche Abenteuer im Bereich BTX- und Videotext-gestützter, elektronischer Zusatzdienste noch in frischer Erinnerung sein dürften¹.

Waren im Frühjahr 1995 erst drei deutsche Tages- und Wochenzeitungen mit Ausgaben im Internet vertreten, boten im Dezember des gleichen Jahres nicht weniger als kommerzielle, deutschsprachige Periodika zumindest Teile ihrer Ausgaben in elektronischer Form an². Eine entsprechende Zählung vom März 1996 ergab 13 Tageszeitungen, 5 Wochenzeitungen bzw. -magazine, zwei vierzehntägig und zwei monatlich erscheinende Magazine³. Die Situation in Deutschland scheint sich damit rasch jener in den USA anzunähern, Die dortige Diskussion kreist inzwischen nicht mehr darum, "ob man im INTERNET publizieren soll, sondern nur noch darum, wie man dies realisieren kann"⁴.

Durch das explosionsartige Wachstum des Internet und die zunehmende Verbreitung des World Wide Web (WWW) werden für Verleger zwei grundsätzliche Fragen geklärt. Erstens beginnt sich abzuzeichnen, welche Infrastruktur verwendet werden muß, um einen möglichst breiten Marktzugang zu erhalten⁵. Weiterhin stehen mit der Dokumentenbeschreibungssprache HTML und grafikfähigen WWW-Browsern erstmalig ausreichende Mechanismen zur Verfügung, um plattformübergreifend strukturierte und gestaltete Informationen anzubieten. Die verfügbaren Layout-Optionen sind mittlerweile so ausgereift, daß über die graphische Aufmachung des Online-Angebots Markenzeichen der Anbieter transportiert werden können.

Der Ansturm der Informationsanbieter auf das World Wide Web sollte jedoch nicht darüber hinwegtäuschen, daß es hierbei eine Reihe von offenen Fragestellungen gibt. Unterscheidet sich die Leserschaft der elektronischen Ausgabe von jener der gedruckten Zeitung, und wenn ja, in welcher Hinsicht? Welche Informationen sollen angeboten werden? Wie müssen diese strukturiert sein? Welche zusätzlichen Dienste werden von den Lesern gewünscht? Wie könnten mögliche Finanzierungsmodelle und Abrechnungsverfahren aussehen? Welche Veränderungen ergeben sich für die Arbeitsweise der Redaktion, der Produktion und des Archivs?

Die genannten Fragestellungen lassen sich nur begrenzt durch theoretische Analysen klären. Sinnvoller erscheint es, einige von ihnen im Rahmen eines praktischen Projekts exemplarisch zu untersuchen. Dies ist Zielsetzung der vorliegenden Ausarbeitung. In ihr werden Konzeption, Entwurf, Implementation, Test und Evaluation des ersten öffentlich zugänglichen, täglich aktualisierten,

1. [Bre91], S.3

2. [Rie95b]

3. [Spi96]

4. [Rief95a], S.99

5. [Rie95a], S.99

WWW-basierten Tageszeitungsdienstes in der Bundesrepublik Deutschland beschrieben. Dieses Projekt wurde im Rahmen einer Forschungsk Kooperation der Prozeßrechnerverbundzentrale (PRZ) der TU Berlin mit der Berlin herausgegebenen *tageszeitung* realisiert.

Die vorliegende Arbeit gliedert sich wie folgt:

In Kapitel 2 dieser Arbeit werden zunächst eine Reihe von strategischen Faktoren und Mediencharakteristika und Gestaltungskriterien diskutiert, die sich während der Konzeption, Einführung und Evaluierung des Angebots als maßgeblich herausgestellt haben.

Kapitel 3 beschreibt den Problemkontext des entwickelten Systems. Hierzu werden die Grundlagen von Hypertext- und Hypermediasystemen sowie die Produktionsbedingungen und technischen Voraussetzungen bei der *tageszeitung* und an der PRZ der TU Berlin beschrieben.

Der Entwurf für die Komponenten des Online-Systems findet sich in Kapitel 4. Dieser umfaßt die Gestaltung der Informationsstruktur, das Layout der angebotenen Dokumente und die funktionale Beschreibung der integrierten Dienste und Hilfsprogramme. Die Implementation dieser Komponenten ist Thema von Kapitel 5.

Die Struktur des Online-Angebots wurde in einem evolutionären Prototypingprozeß unter Einbeziehung von Benutzervorschlägen erstellt. Mit der sukzessiven Erweiterung und dem Test des Systems befaßt sich Kapitel 6.

Die Ergebnisse der Arbeit befinden sich in Kapitel 7. Hierbei wird der Versuch unternommen, Rahmendaten für eine erste Beurteilung des Leserverhaltens, des Bekanntheitsgrads und der möglichen kommerziellen Verwertbarkeit des Angebots zu erheben und auszuwerten. Es handelt sich hierbei um die unseres Wissens erste veröffentlichte Untersuchung dieser Art im deutschsprachigen Raum.

Kapitel 8 beinhaltet die Zusammenfassung der Resultate und gibt einen Ausblick auf mögliche künftige Entwicklungen im Bereich WWW-gestützter Zeitungen.

Computer- und internetgestützte Medien befinden sich derzeit in einer explorativen Phase. Sollte diese Arbeit Anhaltspunkte für die Gestaltung künftiger Online-Periodika geben können, wäre ihre Zielsetzung erreicht.

An dieser Stelle möchte ich einer Reihe von Personen danken, die zum Zustandekommen dieser Arbeit beigetragen haben. Prof. Erhard Konrad und Dipl. Inform. Josef Willenborg bin ich für die inhaltliche Betreuung der vorliegenden Ausarbeitung verpflichtet, ebenso Dipl. Inform. Stephan Fruehauf, auf dessen Initiative das hier beschriebene Projekt zurückgeht. Thomas Bahls war mir bei der Erstellung des Fragebogens behilflich. Dipl. Ing. Peter Kim, Dipl. Inform. Gero Hoffmann und Andreas Stang bin ich für ihre fortwährende moralische Unterstützung ebenso verbunden wie für das Korrekturlesen der Rohversion des vorliegenden Textes. Mein ausdrücklicher Dank geht an die Mitarbeiter der Technik- und Archivabteilung der *tageszeitung*, insbesondere an Dipl. Inform. Ralf Klever; ohne dessen kontinuierliches Engagement dieses Projekt nicht möglich gewesen wäre.

2.0 Publizieren im Informationszeitalter: Ein neues Paradigma?

In seinem 1995 erschienenen Buch *Publishing in the Information Age* konstatiert Douglas Eisenhart, daß zumindest Teile der traditionellen Verlagsbranche vor einem Paradigmenwechsel stehen:

“The community of publishers is now in transition, each individual looking at the new publishing paradigm and deciding whether to embrace the new or to stay with the old. Every individual must decide on his or her own”.¹

Nun befaßt sich Thomas Kuhns Theorie², auf die Eisenhart hier Bezug nimmt, nicht mit der Struktur publizistischer, sondern mit derjenigen wissenschaftlicher Revolutionen. Der Versuch der Anwendung seiner Thesen auf soziale und ökonomische Entwicklungen ist jedoch nicht per se illegitim: eine zumindest partielle Übertragbarkeit wird von Kuhn selbst vermutet³. Nach dem akademischen Abdanken des dialektischen Materialismus bleibt seine Theorie eines der wenigen zitierfähigen Modelle zur Beschreibung revolutionärer Umgestaltung, qualitativer Sprünge also, für die in traditionelleren Theorien adäquate Erklärungsmuster fehlen. (Als direkte Folge davon sehen wir uns seit einigen Jahren mit einer Unzahl publizistisch verordneter Paradigmenwechsel konfrontiert.)

Die bevorstehende Umwertung verlegerischer Werte wird von Eisenhart zunächst durch eine Beschreibung von Branchensymptomen zu belegen versucht. Im entsprechenden Kapitel seines Buches führt er eine Reihe von Analogien zu der von Kuhn beschriebenen Phase einer wissenschaftlicher Krisensituation⁴ auf:

If we look at Kuhn’s model of paradigm change, all the earmarks are there for a fundamental shift in the world of publishing: philosophical analysis and the study of underlying issues; perplexing anomalies besetting the community of practitioners; different communities seeing the same thing differently; and the paradigm of future promise versus the one of past achievements⁵.

Als Beispiele konkreter Symptome zählt Eisenhart folgende Faktoren auf⁶:

- *Veränderung von Definitionen*: Elektronische Geräte und Produkte werden mit Metaphern aus dem traditionellen Verlagswesen belegt: “Sprechende Bücher” (Audiocassetten mit literarischem Inhalt), “Elektronisches Buch” (Sony Discman), “eZine” (elektronisches Magazin).
- *Fusionen und strategische Allianzen*: Unterhaltungselektronik-Konzerne übernehmen Filmstudios, Informationskonzerne fusionieren mit Kabelnetzanbietern, Kabelnetzbetreiber dehnen ihre Aktivitäten auf das Fernmeldewesen aus, etc.

1. [Eis94], S.40

2. [Kuhn96]

3. [Kuhn96], S. 219f

4. [Kuhn96], S.79ff

5. [Eis94] S.37

6. [Eis94], S.2ff

- *Neue elektronische Verfahren und Produkte*: Die meisten Autoren arbeiten inzwischen mit Textverarbeitungssoftware und übernehmen dadurch zunehmend die (typo-)graphische Gestaltung ihrer Arbeiten. Eine Vielzahl neuer Datenträger wird auf dem Markt lanciert.
- *Verändertes Nutzerverhalten*: Der Einsatz von Puffertechnologien (Videorecorder, Anrufbeantworter) dehnt sich aus. Gleichzeitig entstehen neue Kulturtechniken (zunehmende *computer literacy*).
- Es bestehen „*Furcht, Unsicherheiten und Zweifel*“ über die Zukunft traditioneller Verlagsprodukte angesichts einer als unkontrollierbar empfundenen technischen Entwicklung.

Die temporäre Unübersichtlichkeit sich reorganisierender Medienmärkte allein ist allerdings eine schwache argumentative Basis, um einen Paradigmenwechsel zu postulieren. Als weitere Belege seiner These führt Eisenhart deshalb zusätzlich folgende Entwicklungstendenzen an¹:

- Bislang sind publizistische Entscheidungen vornehmlich an der jeweiligen *Form* der Publikation ausgerichtet, beschränkten sich also in aller Regel auf *ein* Medium (z.B. Buch, TV, Zeitschrift). Dies wird zunehmend durch eine *inhaltsbezogene* Sichtweise abgelöst, die mehrere Medientypen integriert.
- Traditionelle Medien sind in erster Linie *performativ*, konfrontierten die Rezipienten also mit einem statischen Produkt. Bei vernetzten elektronischen Medien ist eine Verschiebung hin zu einer *partizipativen* Form zu beobachten, die durch vermehrte Interaktion zwischen Benutzern, Medien und Medienmachern gekennzeichnet ist.
- Bisherige verlegerische Aktivitäten waren primär *produktorientiert*. „Offene Kanäle“, d.h., bidirektionale elektronischen Transportmedien und Speicherinfrastruktur, führen tendenziell zu einer Ausweitung *serviceorientierter* Angebote.
- Über „offene Kanäle“ können gleichermaßen Massen- wie Mikromärkte bedient werden. Angesichts eines anschwellenden Informationsangebots und einer vergrößerten Autonomie der Rezipienten wird die zunehmende Ausrichtung an den *Bedürfnissen kleiner Kundengruppen* nötig.

Die primäre verlegerische Herausforderung besteht Eisenhart zufolge darin, sich von der Bindung an spezifische Medientypen zu befreien und zum Anbieter von Inhalten (*Content Providern*) zu werden. Geistiges Eigentum soll möglichst vielseitig vermarktet werden; der jeweilige konkrete Medientyp ist nurmehr eine von mehreren *packaging and delivery options*² des Verlegers.

Die publizistische Theoriebildung hat diese Perspektive bereits seit langem vorweggenommen. Schon 1907 sah Robert Brunhuber das *Drucken* von Zeitungen, also deren spezifische Form der Vervielfältigung, als vorübergehende Erscheinung an, die dem technologischen Wandel unterworfen und deshalb keine wesentliche Eigenschaft dieses Massenmediums ist.³ Diese Sichtweise deckt sich mit der Auffassung Otto Grohs, für den eine Zeitung (im zeitungswissenschaftlichen

1. [Eis94], S.38ff.

2. [Eis94], S.26

3. [Brö95], S.15

Sinne) “nicht aus den sinnlich wahrnehmbaren, aus Papier und Druckerschwärze oder sonstwie hergestellten ‘Exemplaren’”¹ besteht. Diese platonisch anmutende Trennung der “Idee” einer Zeitung von der Form ihrer Materialisierung und Übermittlung hat sich im Bereich der Zeitungswissenschaften jedoch nie vollständig durchsetzen können. Für Elisabeth Noelle–Neumann z.B. besteht ein wesentliches Merkmal von Zeitungen eben in deren Fixierung durch Schrift und Druck. Sie macht geltend, daß diese definatorische Festlegung zur Abgrenzung von Zeitungen zu anderen Formen und Techniken öffentlicher Kommunikation nötig ist².

Die Begriffsbildung Noelle–Neumanns dürfte für eine Beschreibung verlegerischer Aktivitäten im Bereich elektronischer Medien zu eng sein. Die nahezu vollständige Durchdringung des Produktionsprozesses mit Digitaltechnik legt diese Art publizistischen Engagements förmlich nahe, da einmal digitalisierte Informationen vergleichsweise einfach in weitere elektronische Angebote integriert werden können³. Die Fragestellung, ob eine Zeitung elektronisch sein kann, mag zwar im zeitungswissenschaftlichen Sinne interessant sein, für konkrete verlegerische Aktivitäten ist sie jedoch weitgehend belanglos. Den Zeitungsverlagen dürfte es primär darum gehen, ihre Position als Informationsanbieter in ihren jeweiligen Verbreitungsgebieten zu sichern. Zu einem Zeitpunkt, zu dem die Anzahl der Kommunikationskanäle und der Umfang der verfügbaren Speicherkapazität exponentiell zunimmt, verfügen sie über Inhalte, um Kanäle und Speicher zu füllen. Wo immer sich herausstellen sollte, daß bestimmte Informationsinteressen von Computermedien besser befriedigt werden können als etwa von gedruckten Zeitungen, werden sich Verleger finden, die ihre publizistischen Besitzstände als elektronische Produkte und Dienste vermarkten.

Es gibt allerdings bislang keine Anzeichen dafür, daß computerisierte Medien kurz- oder mittelfristig die gedruckte Zeitung ersetzen werden. Der Computer wird inzwischen weniger als Bedrohung denn als Ergänzung des Printproduktes angesehen⁴. Eisenhart stellt hierzu fest, daß elektronische und gedruckte Medien jeweils spezifische Vorteile aufweisen:

Print and electronic forms have continually been shown to coexist and complement each other, whereas competitive electronic formats in the same mode are more likely to displace one another, such as compact discs’ virtual annihilation of vinyl records.⁵

Erfahrungen der Vergangenheit deuten darauf hin, daß sich verschiedene Medientypen unter Ausnutzung ihrer jeweiligen Stärken komplementär ergänzen:

Significantly, there have been no displacements among the major categories of communications. All core underlying technologies – printing, photography, telecommunications, motion pictures, recording, broadcasting, and computing – continue to coexist, as does the demand of the products and services produced by them.⁶

1. [Gro60], S.621

2. [Noel89], S.287

3. [Rie95a], S.9

4. [Brö]95, S.39

5. Eis94], S.52

6. [Eis94], S.52]

2.1 Strategische Faktoren

Wenn, wie im letzten Kapitel dargelegt, mittelfristig von der Koexistenz und Komplementarität der verschiedenen Medientypen ausgegangen werden kann, existieren für Zeitungsverleger zwei prinzipielle strategische Optionen. Sie können sich entweder darauf konzentrieren, ihren traditionellen, d.h. gedruckten, Publikationen eine möglichst gute Position für den künftigen Wettbewerb mit elektronischen Medien zu verschaffen, indem sie gezielt auf die Vorteile gedruckter Veröffentlichungen setzen (vgl. Kap. 2.2.1). Andererseits können sie im Bereich elektronischer Medien selbst aktiv werden. In diesem Kapitel wird eine Anzahl von Faktoren diskutiert, die die Entscheidung für oder gegen ein solches Engagement direkt beeinflussen.

Die Leser: Unsicherheit über künftige Informationsgewohnheiten

Die Gruppe der regelmäßigen Tageszeitungsleser schrumpft seit Jahren kontinuierlich. Dieses internationale Phänomen ist in erster Linie auf die direkte Konkurrenz des Fernsehens als Informationsanbieter zurückzuführen. Oft übersehen werden hier die *indirekten* Wirkungen, also jene, die das Fernsehen durch die Veränderung der Wahrnehmungsgewohnheiten hervorbringt. Inzwischen gibt es Presseorgane, die diese Entwicklung aufzunehmen versuchen (relativ kurze Artikel, verstärkte graphische Aufbereitung von Information; typische Vertreter dieser publizistischen Gattung sind in Deutschland etwa *Focus* oder *Die Woche*). Ein spektakuläres Beispiel aus den Vereinigten Staaten stellt die Tageszeitung *USA Today* dar:

The radical design innovation of the national Newspaper USA Today was to view the newspaper as a visual medium as much as a textual one, creating colorful, graphic presentation of information throughout the paper. What text is typically provided in the paper is served up in bite-sized chunks for a visually oriented television audience accustomed to consuming information in smaller bits. Though highly controversial for its alleged debasement of the content of the medium, the intuition of the paper's creator, Allen Neuhardt, has proven correct.¹

Veränderungen von Wahrnehmungs- und von Informationsgewohnheiten bedingen sich gegenseitig, und sie geschehen derzeit vornehmlich auf Kosten der Tageszeitungen. Dies läßt sich besonders bei Kindern und Jugendlichen beobachten. *“Die jetzige Generation junger Menschen haben wir wahrscheinlich vollständig verloren. Sie sehen einfach nicht den Wert einer Zeitung. Die Kinder von heute leben in einer MTV-Welt, die schnellebig ist, farbig, interaktiv.”* So Arthur C. Neumann, der Ombudsmann beim *Sacramento Bee*, im Jahre 1994 auf einer Tagung der Bundeszentrale für politische Bildung in Königswinter². Seine in den USA gemachten Beobachtungen gelten offenbar ebenso für europäische Verhältnisse. Während der vom *European Journalism Centre* im Januar 1996 organisierten, international besetzten Gesprächsrunde *Starting a Digital Newspaper* wurden vom Herausgeber der Tageszeitung *Journal de Noticias*, Fernando Martins, nahezu gleichlautende Ausführungen zur Situation in Portugal gemacht³, die auf unge-

1. [Eis94], S.90

2. [Nau94] S.2

3. persönliche Mitteilung von Fernando Martins beim Roundtable *Starting a Digital Newspaper* des European Journalism Center vom 11.–12. Januar in Antwerpen

teilte Zustimmung stießen. Die Ausdehnung ihrer verlegerischen Aktivitäten auf vernetzte computerisierte Medien wird inzwischen von vielen Zeitungsmachern als effektives Instrument angesehen, um die Bindung zwischen der jungen Generation und Zeitungen neu zu beleben und sich ändernde Informations- und Kommunikationsmuster durch Benutzerforschung frühzeitig zu registrieren¹.

Auch bei Teilen der erwachsenen Generation sind Anzeichen einer veränderten Mediennutzung festzustellen. Besonders die Exponenten der vielumworbenen 'Info-Elite' werden inzwischen unter einer Lawine mehr oder weniger relevanter Informationen begraben. Als mögliche Lösung dieses Dilemmas bietet sich die rigorose Vorselektierung von Nachrichten an. Eine der wesentlichen strukturellen Stärken vernetzter computerisierter Medien liegt gerade darin, daß mit ihrer Hilfe individualisierte Informationspakete auf einfache Weise zusammengestellt werden können, sei dies in Form spezialisierter Diskussionsgruppen oder redigierter elektronischer Fachzeitschriften. Große Erwartungen werden auch an elektronische 'Gatekeeper' und 'intelligente Agenten' geknüpft, die benutzerrelevante Informationen künftig automatisch aus elektronischen Netzen einsammeln sollen.

Über die technische Machbarkeit automatischer Nachrichtenfilter läßt sich ebenso trefflich spekulieren wie über deren mögliche gesellschaftlichen Auswirkungen² (vgl. dazu Kap. 2.2.2). Das Problem der Nachrichtenüberflutung wird sich jedoch proportional zur exponentiell wachsenden Menge produzierter Informationen verschärfen und könnte zu einer Art ökologischen Frage des 21. Jahrhunderts werden. Der Einsatz fortgeschrittener Informations- und Kommunikationsmedien gilt zur Zeit noch als Statussymbol, doch dies könnte sich zukünftig ändern. Trendsetter wäre dann jener Personenkreis, der nicht gezwungen ist, derartige Medien zu nutzen.

Der Informationsmarkt: Machtkampf zwischen Informations- und Diensteanbietern

Bei der bereits erwähnten Gesprächsrunde *Setting up a Digital Newspaper* verglich Jan Lamers, Managing Director des belgischen Handelsblattes *Financeel Economische Tijd*, die derzeitige Situation der Informationsanbieter mit derjenigen von Nahrungsmittelproduzenten vor und nach dem Aufkommen von Supermarktketten. Traditionell besaßen diese Produzenten weitgehende Autonomie bei der Gestaltung ihres Produkts, ihrer Preise und Handelsbeziehungen. Diese Situation hat sich durch die Entwicklung zentral organisierter Distributionsnetze mittlerweile in ihr Gegenteil verkehrt: "Today, the power is held by the people who own the shelf space."³

Informationsanbietern droht nach Lamers' Ansicht heute eine ähnliche Situation. Wurden bislang verschiedene Medientypen über separate Distributionskanäle an die Kunden gebracht, so werden sich künftig die Anbieter digitaler Universaldienste zunehmend zwischen Kunden und Anbieter schieben. Die hiervon ausgehende Gefahr für Informationsanbieter im Allgemeinen und Zeitungsverlage im Speziellen kann Lamers zufolge kaum überschätzt werden: "Once we loose direct con-

1. [Brö95], S.53f. vgl. auch [Rie95a], S.31, 85ff, 115

2. [Brö95], S.72ff, 80ff

3. persönliche Mitteilung von Jan Lamers beim Roundtable *Starting a Digital Newspaper* des European Journalism Center vom 11.-12. Januar in Antwerpen.

*tact to our customers, we loose our customer franchise. And if we loose our customer franchise, we loose our business.”*¹

Gerade kleinere Zeitungen befinden sich gegenüber rapide wachsenden professionellen Anbietern elektronischer Dienste bereits jetzt in einer schwierigen Position: sie laufen Gefahr, sich die Konditionen diktieren lassen zu müssen. Inzwischen gibt es allerdings auch Beispiele konzertierter Gegenbewegungen seitens der Presseverlage. So vereint das verlagsübergreifende belgische Projekt *Central Station* die digitalen Ausgaben nahezu sämtlicher nationaler Periodika unter einem Dach. Das Projekt verfügt über eine eigene technische Infrastruktur und ist mittlerweile in der Lage, nahezu jedes elektronische Zielformat zu liefern. Statt sich einzeln – und unter dezenten Hinweisen auf die große Anzahl der Mitbewerber – die Vertragsbedingungen vorschreiben zu lassen, haben die belgischen Verlage hier eine Organisation geschaffen, die ihre Interessen gegenüber Anbietern elektronischer Dienste wahrnimmt. Diese Organisation tritt nach außen als alleiniger Vertragspartner für das Gesamtpaket belgischer Publikationen auf, was die Verhandlungsposition der Zeitungen beträchtlich stärkt. Investitionen und Gewinnaufteilung handeln die Verlagshäuser unter sich aus².

Der Verlust der Kundenbasis ist also eine mögliche, keineswegs jedoch zwangsläufige Entwicklung. Generell, so Lamers, sind Zeitungen gerade aufgrund ihrer Erfahrungen im Abonnementsbereich hervorragend gerüstet, um ihre Marktposition zu verteidigen. Katja Riefler nennt eine Anzahl weiterer Pluspunkte:

[Zeitungen] verfügen über eine Leserschaft, die den Informationen ‘ihrer’ Zeitung vertraut. Sie haben Expertenwissen in der Informationsverbreitung und können auf ein funktionierendes Vertriebssystem zurückgreifen. Das Verhältnis zu den Anzeigenkunden ist gewachsen und wurde jahrelang gepflegt. [Sie] verfügen über Inhalte, die mit mehr oder weniger Aufwand in elektronische Angebote einfließen können.³

Vieles wird davon abhängen, bis zu welchem Grad vernetzte computergestützte Medien den Markt durchdringen werden. Zeitungsverlage sollten diese Medien jedoch als möglichen Distributionskanal ernstnehmen. Obwohl mittelfristig von einer Koexistenz von Buch-/Zeitungsdruck und vernetzten computergestützten Medien ausgegangen werden kann, ist es möglich, daß sich diese Konstellation in fernerer Zukunft zugunsten digitaler Medien verschiebt. Ursachen hierfür könnten sowohl ökonomischer wie auch technischer Natur sein⁴, z.B. die immer perfekter werdenden Möglichkeiten des Home-Printing.

Konkurrierende Medientypen: Wettbewerb auf neuem Terrain

Obwohl sich hyper- und multimediale Medien aus Elementen traditioneller Medientypen zusammensetzen, besitzen sie eine eigene Logik und Ästhetik sowie spezifische Nutzergewohnheiten⁵.

1. persönliche Mitteilung von Jan Lamers beim Roundtable Starting a Digital Newspaper des European Journalism Center vom 11.–12. Januar in Antwerpen.

2. persönliche Mitteilung von Hans Cobben, General Manager von *Central Station*, beim Roundtable Starting a Digital Newspaper des European Journalism Center vom 11.–12. Januar in Antwerpen

3. [Rie95a], S.9f

4. [Brö95], S. 66f

Die traditionellen Medienanbieter (Zeitung, Zeitschrift, Rundfunk und Fernsehen) verfügen hinsichtlich des neuen Mediums jeweils über partielle Kompetenz. Sie alle stoßen derzeit im Bereich der computergestützter Online-Medien als Konkurrenten aufeinander. Erschwerend kommt hinzu, daß dies im diskursiven Universum des Internet mit seinen zumeist frei verfügbaren Informationsangeboten geschieht.

Direkte Konkurrenz mit Zeitschriften-, Rundfunk- und Fernsehredaktionen *im gleichen Medium* ist für Zeitungsverleger eine historisch neue Situation. Zwar gelten Tageszeitungen immer noch als traditionell wichtigster Informationsanbieter in ihren Verbreitungsgebieten¹, doch ihre Sonderstellung als einziges regional verankertes öffentliches Forum ist spätestens seit der Deregulierung des Rundfunkmarktes und dem damit verbundenen Aufkommen lokaler Radiosender gefährdet. Es ist durchaus vorstellbar, daß eine agile lokale Rundfunkredaktion, die ein zusätzliches Angebot im Bereich vernetzter computergestützter Medien zur Verfügung stellt, eine Tageszeitungsredaktion hinsichtlich Aktualität und Recherche überflügeln kann. Dies gilt ebenso für Regionalredaktionen von Fernsehanstalten. Werbekunden werden derartige Entwicklungen sorgfältig verfolgen.

Alle Anbieter traditioneller Medientypen befinden sich in einer ähnlichen Ausgangslage. Sie müssen jenseits ihrer angestammten Tätigkeitsfelder zusätzliche Kompetenz aufbauen, um attraktive Online-Angebote zu schaffen. Zeitungen und Zeitschriften besitzen dabei einen gewissen Startvorteil, da die Kommunikation im Internet nach wie vor hauptsächlich narrativ geprägt ist (vom Verschwinden der Schrift zugunsten eines mehr graphisch orientierten Angebots kann zur Zeit keine Rede sein – im Gegenteil: Computernetzwerke wie *The WELL* wurden bereits als eine moderne Form des klassischen literarischen Salons beschrieben²). Zudem ist die Übertragung audiovisueller Informationen zum Endkunden wegen immer noch stark begrenzten Netzbandbreiten zur Zeit wenig attraktiv. Andererseits zeigen vielfältige Aktivitäten gerade der Rundfunk- und Fernsehanbieter im Online-Bereich, daß die Erweiterung ihres redaktionellen Profils in Richtung textueller Angebote mit kalkulierbarem Aufwand machbar ist. Dies alles sollte für Zeitungsverleger Grund genug sein, ihren knappen Vorsprung zu verteidigen und – wenn möglich – auszubauen.

Die Produktion: Lange Lernkurve für Electronic Publishing

Das Engagement im Online-Bereich wird von Verlagen in der Regel als Zukunftsinvestition angesehen³. Dies gilt insbesondere für die im Internet abrufbaren Zeitungen. Nicht zuletzt wegen derzeit fehlender Abrechnungsmechanismen wird hierbei bewußt in Kauf genommen, daß Online-Zeitungsausgaben, ökonomisch gesehen, derzeit ein Verlustgeschäft sind.

Angesichts ungewisser Ertragsmöglichkeiten und der großen Konkurrenz sieht sich jeder Zeitungsverleger zunächst mit der Frage konfrontiert, ob ein Engagement in diesem Bereich derzeit überhaupt sinnvoll ist. Die Option "Abwarten und Beobachten"⁴ erscheint vielen Entscheidungsträgern besonders angesichts schwer kalkulierbarer Investitionskosten attraktiv.

5. [Rie95a], S.28

1. [Rie95a], S. 8

2. [Eis94], S. 105

3. [Brö95], S.45f

4. [Eis94], S.142 (Tabelle)

Die Option des Abwartens ist jedoch aus mehreren Gründen problematisch. Erstens bringt die Installation eines Online-Angebots neben technischen Anforderungen auch solche im gestalterischen, redaktionellen und organisatorischen Bereich mit sich. Dies gilt weniger für die nötige technische Ausstattung als für die zuweilen um Größenordnungen aufwendigere Integration in den Produktionsprozeß, die Konzeption und das Design des Angebots. Einige der wichtigsten Gründe dafür seien im folgenden genannt:

- Oft tauchen technische Schwierigkeiten beim Versuch auf, Texte und Layoutmerkmale aus dem Format des Redaktionssystems in jenes des Online-Mediums zu konvertieren. Dies ist auch und gerade beim Einsatz von fortgeschrittener Layout-Software (z.B. Quark XPress, PageMaker) der Fall¹.
- Die Präsentation hypermedialer Online-Angebote unterscheidet sich erheblich von derjenigen eines Printprodukts. Die Gestaltung eines solchen Angebots bedarf spezieller Qualifikationen in der Layoutabteilung. Diese sind in traditionellen Zeitungsverlagen zumeist nicht vorhanden.
- Multimediale Präsentationsformen bringen neue Arbeitsprofile für Redakteure mit sich². Hier sind Zeitungsredakteure i.d.R. schlechter vorbereitet als ihre Kollegen aus den Redaktionen der Hochglanzmagazine, der Rundfunk- und der Fernsehstationen. Zudem muß das Angebot in der Regel dem Interessenprofil der Online-Leserschaft angepaßt werden. Dies betrifft sowohl die redaktionelle Themenauswahl als auch die Bereitstellung technischer Dienste (z.B. Archiv-, Such- und Filterfunktionen).
- Die Kommunikation zwischen Rezipienten und Produzenten wird durch elektronische Netze radikal vereinfacht. Kommunikationsdienstleistungen für Leser müssen in der Regel ausgeweitet werden. Die in diesem Bereich arbeitenden Personen sollten sich zuvor mit der Nutzungs- und Diskussionskultur in elektronischen Netzen vertraut gemacht haben³.

Zweitens ist fraglich, ob die Investitionskosten für den Aufbau und den Betrieb eines Online-Angebots in Zukunft geringer sein werden. Wahrscheinlicher ist, daß sich die Nutzer an ständig verbesserte Dienste gewöhnen und diese schließlich als selbstverständliche Minimalvoraussetzung eines professionellen Angebots ansehen werden. Die Hürden für den Einstieg künftiger Anbieter werden dadurch laufend erhöht. Je mehr sich die Kluft zwischen entwickelten elektronischen Diensten und der traditionellen Printproduktion verbreitert, desto mehr schwindet die Option, einen solchen Dienst mit Verlagsmitarbeitern organisch aufzubauen. Das fehlende Know-How kann schließlich nur noch über externe Spezialisten oder Firmenakquisitionen beschafft werden.

Bedingt durch wachsende Konkurrenz wird es drittens immer schwieriger, das eigene Online-Produkt sichtbar am Markt zu platzieren. Der einfache Hinweis "auch wir sind jetzt online" wird in Zukunft immer weniger ausreichen, potentielle Nutzer von der Relevanz des eigenen Angebots zu überzeugen.

1. [Bou95]

2. [Brö95], S. 4f

3. [Kat94], S. 50ff

2.2 Traditionelles und elektronisches Publizieren

Worin bestehen die wesentlichen Merkmale traditioneller Zeitungen? Welche Vorzüge und Nachteile kann demgegenüber ein elektronischer Dienst aufweisen? Welche Ziele können sinnvollerweise mit einer elektronischen Zeitung verfolgt werden? Wir werden zunächst versuchen, uns diesen Fragen durch eine einleitende Darstellung der Funktion des Massenmediums 'Zeitung' zu nähern.

Zeitungen vermitteln der Öffentlichkeit portionsweise aktuelles oder aktualisiertes Wissen über gesellschaftliche Entwicklungen und Zusammenhänge. Dieses Wissen und der diskursive Prozeß, in dem es hervorgebracht wird, gewährleisten den kulturellen und ideologischen Zusammenhalt des selbstregulierenden, sozialen Systems einer Gesellschaft. Massenmedien institutionalisieren die Beschaffung und Verbreitung aller Arten von Nachrichten. Sie strukturieren Informationen in einer Form, die es dem Leser erleichtert, deren jeweiligen Kontext und Relevanz zu erkennen. Diese Beobachtungen mögen Otto Groh dazu bewogen haben, Zeitungen als ein *Kulturwerkzeug* zu beschreiben, das die Wesensmerkmale *Aktualität*, *Periodizität*, *Publizität* und *Universalität* aufweist¹. In demokratisch verfaßten Gemeinwesen haben Zeitungen zudem eine kommunikative Funktion – sowohl auf freiwilliger Basis (Veröffentlichen von Leserbriefen) wie auch rechtlich verankert (Verpflichtung zur Veröffentlichung von Gegendarstellungen).

Wenn wir – in idealtypischer Vereinfachung der Sachlage – davon ausgehen, daß der Leser der von ihm favorisierten Zeitung *vertraut*, bedeutet dies, daß er den Prozeß des Auswählens, Strukturierens und Darstellens von Information teilweise an die Redakteure 'seines' Publikationsorgans delegiert. Er erwartet von 'seinem' Medium in diesem Fall nicht nur Aufklärung über das, was ihn interessiert, sondern auch über das, was ihn (nach Meinung der Redaktion) interessieren *sollte*. Diese Feststellung gilt in demokratisch verfaßten Gemeinwesen mit verbrieft Pressefreiheit ungeachtet aller ideologiekritischen Vorbehalte. Die redaktionelle Auswahl und Bearbeitung von Nachrichten wird hier vom Leser subjektiv eben *nicht* als Zensur und Manipulation empfunden, sondern als notwendiger und geldwerter Informationsfilter.

2.2.1 Charakteristika traditioneller und elektronischer Zeitungen

Nach diesem kurzen Exkurs können wir uns nun den spezifischen Stärken und Schwächen des traditionellen und elektronischen Publizierens zuwenden. Dabei sind uns die positiven Eigenschaften traditioneller, d.h., gedruckter, Ausgaben so selbstverständlich, daß sie erst im Vergleich mit ihrem elektronischen Pendant wahrgenommen werden. Eine Zeitung ist an vielen Orten erhältlich, kann auf Wunsch aber auch ins Haus zugestellt werden. Sie ist erschwinglich und leicht transportierbar. Man kann sie sowohl am Frühstückstisch, in Verkehrsmitteln oder in der Badewanne benutzen. Eine Zeitung kann kostenlos tagelang aufbewahrt und abschnittsweise gelesen oder archiviert werden. Sie bietet dem Leser die Freiheit, Passagen oder ganze Artikel zu überspringen. Das vertikal ausgerichtete Format erlaubt, große Informationsmengen schnell und durch alleiniges Bewe-

1. [Gro60], S. 621

gen der Augen zu sichten. Schließlich ist eine Zeitung einfach zu bedienen und weder anfällig gegen Stromausfälle noch gegen Viren.

Zugunsten elektronischer Publikationen kann demgegenüber angeführt werden, daß praktisch unbegrenzte Mengen an (Speicher-)Platz zur Verfügung stehen. Hierbei ist das Angebot elektronischer Publikationen ist nicht auf Text und Standbilder beschränkt, sondern kann auch animierte Sequenzen, Audio- und Videomaterial enthalten. Eine dynamische Aktualisierung des Inhalts ist möglich; Teilinhalte können ohne nennenswerten technischen Aufwand zu individualisierten Informationspaketen zusammengefaßt und verschickt werden. Elektronische Publikationen erlauben eine Verbreitung aktueller Inhalte über die geographischen Vertriebsgrenzen der gedruckten Ausgabe hinaus. Der interaktive Zugang zu Archivmaterial und die Referenzierung externer Informationsquellen wird enorm vereinfacht, die direkte Kommunikation zwischen Leser und Anbieter unterstützt

Die Vorteile des traditionellen Mediums lesen sich wie eine Liste der Unzulänglichkeiten des neuen (und umgekehrt). Insbesondere hinsichtlich der wichtigen Kriterien Preis, Transportierbarkeit, Benutzerfreundlichkeit und Robustheit bleibt das elektronische Medium zur Zeit weit hinter der vertrauten gedruckten Form zurück. Zwar wird an der Entwicklung von handlichen Multimediageräten gearbeitet, die hinsichtlich ihres Gewicht, ihrer Abmessungen und Displaygröße gedruckten Zeitungen näher kommen sollen als derzeitige Notebook-Computer¹. Nicht zuletzt wegen der rein technischen Schwierigkeiten bei der Entwicklung und Produktion höchstauflösender, u.U. faltbarer Displays im Format einer Zeitungsseite für weniger als 500 US\$ steht ein Markteinführungstermin für derartige Geräte bislang aus.

Vorerst wird auf der Anbieterseite noch nach tragfähigen technischen und inhaltlichen Konzepten gesucht. Daß Zeitungen mit elektronischen Dienst wie Audiotex- und Faxdiensten Kunden gewinnen können, ist in den USA mehrfach erfolgreich demonstriert worden². Ähnliches gilt für Angebote über kommerzielle Online-Dienste. Im Bereich des Internet sind jedoch traditionelle *Information Provider*- oder *Publishing Partner*-Modelle zur Abrechnung in Anspruch genommener Dienste in aller Regel nicht tragfähig. Einerseits fehlen hier noch immer die nötigen Abrechnungsmechanismen, andererseits waren – bedingt durch die große Konkurrenz frei verfügbarer Informationsressourcen im Internet – Abonnementsgebühren bislang praktisch nicht durchsetzbar. Praktikabel bleibt z.Zt. nur das *Sponsorship*-Modell, d.h., die vollständige Finanzierung über Werbung³. Dieser Weg ist z.B. durch verschiedene Anbieter von Internet-Suchdiensten erfolgreich beschritten worden, doch der Markt ist hart umkämpft. - angesichts von derzeit 30 Millionen existierenden WWW - Seiten⁴ wenig erstaunlich ist. Ob sich das *Sponsorship*-Modell oder Abrechnungsverfahren auf breiter Ebene im Internet durchsetzen können werden, ist zur Zeit völlig offen.

Auch hinsichtlich der strukturellen Gestaltung, der anzubietenden Inhalte und der vom Kunden gewünschten Dienste gibt es zur Zeit mehr Fragen als Antworten. Dies ist beileibe kein neues Pro-

1. [Brös95], S. 63

2. [Rie95a], S.13–22

3. [Brö94], S. 49f

4. laut Alta Vista Index Mai 1996, URL: <http://altavista.digital.com/>

blem beim Vertrieb elektronischer Information. Bereits 1985 brachte Walter Baer das verlegerische Dilemma folgendermaßen auf den Punkt:

“We know that the electronic media differ from print and from each other. We know that successful products and services usually combine electronic and print elements. What we do not know is how these elements should be combined, and in which proportion: print, online retrieval, physical distribution of software, downloading, optical disc storage, and so on. The possible combinations are truly astounding. Determining which combinations provide the greatest value to our customers is perhaps the most interesting part of the electronic publishing story.”¹

Zunächst stellt sich allerdings die Frage, welche Nachteile der traditionellen Zeitungen eigentlich durch Online-Angebote ausgeglichen werden sollen und können. Natürlich gibt es naheliegende Einsatzgebiete (z.B. verbesserter Leserservice und Vertrieb durch die Möglichkeit, per email zu kommunizieren). Dazu ist allerdings nicht unbedingt erforderlich, daß auch die Zeitungsinhalte online angeboten werden. Der reichlich verfügbare Speicherplatz könnte für zusätzliche Berichterstattung genutzt werden. Allein: wie groß ist das Interesse der Leserschaft? Wie zuvor ausgeführt, wird eine Zeitung in aller Regel gerade deswegen gelesen, weil sie Informationen in aufbereiteter und konzentrierter Form anbietet. Gibt es überhaupt Bedarf für erweiterte Berichterstattung? In Jon Katz's Worten:

“It's an interesting effort to provide an additional news dimension, but it seems a pointless one. If this information wasn't important enough to be printed in the paper, why should we pay to retrieve it? That's the point of a newspaper, after all – to filter the worthwhile information, then print it.”²

Zeitungen könnten ihren Lesern ermöglichen, in einem öffentlichen Forum mit Autoren und anderen Lesern zu diskutieren. Angesichts der langjährigen Erfolgsgeschichte der USENET-News entbehrt ein solches Angebot allerdings der Originalität. Eine fortlaufende Aktualisierung von Artikeln wiederum hätte nur dann Sinn, wenn die Benutzer das Informationsangebot fortlaufend oder doch mehrfach pro Tag in Anspruch nehmen – ein eher unwahrscheinliches Szenario, zumal bereits Rundfunk, Fernsehen und nicht zuletzt die USENET-News genau diesen Dienstyp vertreten. Worin also bestehen mögliche Ansätze für eine elektronische Zeitung?

2.2.2 Typologie von Online-Angeboten für Zeitungen

Die folgende Absätze konzentrieren sich ausschließlich auf Szenarien für *Endkunden*. Mögliche Dienste für *Werbekunden* werden nicht erörtert. Hierzu sei lediglich angemerkt, daß die auf gedruckte Zeitungsausgaben ausgerichteten Strategien nur sehr eingeschränkt auf elektronische Ausgaben übertragbar sind. Die Autonomie der elektronischen Leserschaft ist um ein Vielfaches größer³, ihre Akzeptanz gegenüber unerwünschter Information hingegen geringer⁴. Daß im

1. [Bae85],S.123

2. [Kat94]

3. [Eis94], S.163

4. [Kat94]

Bereich der Online–Medien eigene Gesetzmäßigkeiten herrschen, nimmt auch die Werbebranche zunehmend zur Kenntnis¹.

“Sinnvoll erscheinen nur jene Dienste, die sich in Aktualität und Inhalt von der gedruckten Zeitung unterscheiden”, bemerkt Daniel Brössler in seiner 1995 erschienenen Untersuchung *Zeitung und Multimedia*². Legt man diese These zugrunde, bleibt nur eine vergleichsweise kleine Anzahl sinnvoller Dienste übrig:

- Bereitstellung des Angebots für Interessenten, die sich außerhalb des geographische Vertriebsgebiets befinden, jedoch über Computer gut zu erreichen sind.
- Zeitlich vorverlagerte Bereitstellung des Angebots.
- Kostengünstigerer Zugang zum aktuellen Angebot.
- Vertiefung des Angebots durch Querverbindungen zu internen und externen Ressourcen und zum eigenen Archiv (für Leser mit Bedarf an Hintergrundinformationen).
- Individualisierung durch Bereitstellen thematisch beschränkter Informationspakete für Nutzer mit besonderen Interessen.

Die drei ersten Punkte sind ebenso naheliegend wie einfach zu realisieren. Durch die weltweite Verbreitung des Internet kann der Zugriff auf ein Informationsangebot auch in abgelegenen Regionen erfolgen. Da die “elektronische Matrize” einer Zeitung oft mehrere Stunden vor der Auslieferung der gedruckten Ausgabe vorliegt, ist ein früherer Erscheinungszeitpunkt der elektronischen Version möglich. Für diese fallen keine Druck– und Transportkosten an, so daß die Angebotspreise deutlich gesenkt werden können. Ob ein solches Angebot zu Lasten der gedruckten Ausgabe geht und betriebswirtschaftlich vertretbar ist, wird jeder Verlag selbst kalkulieren.

Die optionale Vertiefung des Angebots ist für jene Personen interessant, die zu bestimmten Schwerpunkten einen genaueren Überblick benötigen (z.B. im Bereich der Ausbildung, Forschung, Lehre und Journalismus). Eine weitere Zielgruppe sind Einzelpersonen, die sich in ihrer Freizeit aus Interesse oder Hobby mit bestimmten Themengebieten beschäftigen. Dieser Dienstyp zielt nicht auf ein Massenpublikum, sondern auf kleinere Nutzergruppen mit spezialisierten Interessen. Generell werden Zeitungen hier inhaltlich nicht mit Fachmagazinen konkurrieren können. Ihre Aufgabe liegt eher darin, einführende Übersichten und kommentierte Verweise zu erstellen. Hier eröffnet sich ein neues Feld für Medienrezensionen³.

Traditionell wird die Option der Individualisierung von Nachrichten hauptsächlich für professionelle Zwecke in Anspruch genommen. Dabei sind wirklich individuell zugeschnittene Informationen ein Luxusgut, das derzeit ausschließlich Führungskräften vorbehalten sind. Sie verfügen über die notwendigen persönlichen Referenten und Stabsstellen, um sich ein Informationspaket “auf den Leib schneiden” zu lassen. Weniger exponierte Persönlichkeiten behelfen sich mit hochspe-

1. [Bla95]

2. [Brö95], S.93

3. Vgl. z.B. die wöchentliche Beilage OnLine der britischen Tageszeitung *The Guardian*

zialisierten Fachinformationsagenturen, etwa in Form von Branchenbriefe, Newsletters, Fax- und Clippingservices. Die Nutzer derartiger Dienste sind jedoch zumeist an einer zusammenfassenden Übersicht von Beiträgen verschiedener Medien interessiert. Für einzelne Zeitungen ist die Einführung eines solchen Angebots wahrscheinlich nur dann sinnvoll, wenn sie über außergewöhnliche Sachkompetenz, schwerpunktmäßige Berichterstattung und exklusive Quellen verfügen.

Die übermäßige Spezialisierung oder gar Individualisierung von Zeitungsangeboten entspricht derzeit nicht den Bedürfnissen des breiten Publikums¹. Daß sich dies zukünftig ändern könnte, ist Ausgangshypothese des von Jerome S. Rubin geleiteten Projekt *News in the Future (NiF)* am MIT Media Lab. Die geistigen Väter von *NiF*, Walter Bender und Nicolas Negroponte, vertreten die Auffassung, daß die heutigen Massenmedien in Zukunft von individualisierten Informationsdiensten zurückgedrängt werden. Bender und Negroponte arbeiten an der Entwicklung von automatischen Suchprogrammen, sogenannten *personal agents*, die auf der Grundlage benutzerdefinierter Informationsprofile operieren. Hierzu gibt der Benutzer eine Liste derjenigen Themen und Stichworte an, die ihn interessieren. Sein "digitaler Doppelgänger" durchsucht anhand dieser Liste kontinuierlich und systematisch alle im Netz verfügbaren Informationen und überprüft sie auf ihre Relevanz für ihren menschlichen Gegenpart. Die gefundenen Dokumente werden klassifiziert und in Form einer automatisch und individuell zugeschnittenen "Zeitung" – Arbeitstitel: *Digital Me* – präsentiert. Fortgeschrittene Versionen dieser digitalen Doppelgänger sollen Listen mit individuellen Benutzervorlieben sogar automatisch erstellen, und zwar anhand der Auswertung von bisher angeforderten Dokumenten.

Solche "persönliche Agenten" müssen auf der Grundlage von Heuristiken arbeiten, um Annahmen über die Inhalte von Nachrichten treffen zu können. Diesem Problem wollen die Entwickler durch einen tiefen Griff in die Werkzeugkiste der *Artificial Intelligence* zu Leibe rücken, z.B. mit formaler Wissensrepräsentation und adaptiver Benutzermodellierung². Ein solches automatisches generiertes, personalisiertes Medium wäre zwar aktuell und optional periodisch, jedoch nicht mehr universell (vgl. Einleitung von Kap. 2.2). Auch das Kriterium der Publizität wäre nur noch sehr eingeschränkt vorhanden: das personalisierte Informationspaket wird zwar öffentlich verfügbare Informationen enthalten, diese jedoch nur noch dekontextualisiert.

Legt man die in der Einleitung von Kap. 2.2 genannten Kriterien von Otto Groh zugrunde, kann das Ergebnis eines benutzerdefinierten Suchprozesses deshalb schwerlich als "Zeitung" bezeichnet werden. Die Frage, ob dieses Konzept tatsächlich die bekannte Form der Massenmedien ablösen kann und wird, ist von Daniel Brössler ausführlich diskutiert worden³, ebenso die Frage, ob dies wünschenswert wäre. Da dieser Autor detailliert auch auf Probleme der technischen Realisierbarkeit und mögliche gesellschaftliche Auswirkungen (z.B. die potentiellen Bedrohung der Privatsphäre) eingeht, soll hier auf eine genauere Darstellung verzichtet werden. Selbst die Entwickler der *Digital Me* scheinen hinsichtlich dieser Fragen gewisse Zweifel zu beschleichen.

1. [Brö95], S.69

2. [Mit01]

3. [Brö95], S.68–75, 80–82, 94

Noch bevor auch nur der Prototyp eines solchen Agenten präsentiert werden kann, macht man sich nämlich am MIT Media Lab bereits einige Sorgen um die Benutzerakzeptanz:

Independent of the application domain (advertising, news, etc.), a problem of distrust is likely to ensue if users do not have access to their own user model and do not understand how their user model works¹.

Wo aber Gefahr ist, wächst bekanntlich das Rettende auch – und zwar in Form eines neuen Systems namens *PeerGlass*:

The proposed solution is a system (...) that allows users to explore their own user models with varying degrees of depth. At the deepest level, PeerGlass tries to demystify how reasoning is applied to raw data².

Womit das Problem solange gelöst wäre, wie die Benutzer die demystifizierenden Erläuterungen von PeerGlass verstehen und ihnen Glauben schenken.

2.2.3 Informationeller Mehrwert durch hypertextgestützte Zeitungsausgaben?

Nach Kuhlen kann informationeller Mehrwert durch Rekonstruktion von Texten im Medium Hypertext unter anderem dann entstehen, wenn hypertextspezifische Möglichkeiten wie z.B. Verweise auf Ergänzungen und externe Daten genutzt werden³. Kuhlen geht dabei davon aus, daß durch unterschiedliche Aufteilung und Strukturierung von Inhalten (*chunking*) verschiedenartige Aufgabenstellungen besser unterstützt werden können⁴. Seine Aussagen beziehen sich allerdings in erster Linie auf die Hypertext-Repräsentation von *Wissensbasen* wie Lexika, Fach- und Lehrbücher. Seine Argumentation ist auf Zeitungen nur eingeschränkt übertragbar, da sich z.B. Artikel einer Zeitungsausgabe in aller Regel nicht aufeinander beziehen und deshalb Verweise zwischen Artikeln derselben Ausgabe zumeist nicht sinnvoll sind.

Informationeller Mehrwert kann zwar im Prinzip durch Verweise auf früher erschienene Artikel und externe Quellen zum gleichen Thema geschaffen werden. Dies ist allerdings mit zusätzlichem editorischen und archivarischem Aufwand für die Indexierung verbunden (der z.B. im vorliegenden Projekt nicht geleistet werden konnte). Zu vermuten ist, daß derartige Dienste z.Zt. nur für eng umrissene Nutzergruppen von Interesse sind. Diese haben jedoch bereits jetzt schon die Möglichkeit, Quellenrecherchen von Tageszeitungsredaktionen gegen entsprechendes Entgelt anzufordern.

Zeitungsartikel fallen nicht unter die von Kuhlen genannten, "gut konvertierbaren" Texte, sondern eher in die Kategorie der "problematischen", d.h., argumentativen, berichtenden und diskursiven Textsorten⁵. Interessanterweise stellt sich das kritische Problem der Segmentierung sinnvoller informationeller Einheiten bei Zeitungstext trotzdem nur in abgeschwächter Form dar. Jeder Arti-

1. [Mit01]

2. ebd

3. [Kuh91], S. 160

4. [Kuh91], S. 162

5. [Kuh91], S. 175f

kel stellt nämlich bereits eine solche stilistische Einheit dar. Die von Kuhlen geforderte *Kohäsivität* (vgl. Kap. 3.1.1) dieser Einheiten kann in aller Regel vorausgesetzt werden (d.h., zum Verständnis eines Zeitungsartikels sind, abgesehen von der Rubrikangabe, keine weiteren Kontextinformationen nötig). Die Segmentierung der Gesamtausgabe kann problemlos auf der Grundlage deskriptiver Textelemente und formaler Texteingenschaften (Struktur- und Layoutinformationen) erfolgen – jede neue Artikelüberschrift markiert den Beginn einer informationellen Einheit. Das Hauptproblem hypertextgestützter Zeitungen liegt demnach weniger in der Konstruktion informationeller Einheiten als vielmehr in deren Strukturierung und Darstellung. Zur Verbesserung der Übersichtlichkeit des Hypertextangebots müssen hierarchische Dokumentenstrukturen sowie zusätzliche Navigationselemente eingeführt werden. Diese Mechanismen sind der traditionellen Zeitung fremd – dort wird die jeweilige Relevanz von Artikeln durch Layoutmechanismen (Kästen, Überschriftengrößen, Fotos etc.) signalisiert. Ob ein hypertextgestütztes Angebot in dieser Hinsicht den informationellen Wert des Papierprodukts vermitteln kann, ist aus heutiger Sicht eher zweifelhaft.

Zusammenfassende Schlußfolgerungen

Obwohl Zeitungen nicht notwendigerweise an das Medium Papier gebunden sind, ist das ‘Benutzerinterface’ der traditionelle Druckausgabe für die meisten Anwender geeigneter als eine Darstellung auf dem Bildschirm. Informationeller Mehrwert eines vergleichbaren Hypertextangebots erfordert zusätzlichen intellektuellen Aufwand (vgl dazu auch Kap. 8.0).

Marketingargumente für ein Online-Angebot von Zeitungen sind aus Kundensicht der frühere Erscheinungszeitpunkt, die Möglichkeit des internationalen Vertriebs, leichtere Archivierbarkeit und geringere Angebotskosten. Die Vermarktungschancen derartiger Angebote sind jedoch ungewiß, zumal zur Zeit ist nur schwer einschätzbar ist, bis zu welchem Grade computergestützte vernetzte Medien den Markt durchdringen und zur Konkurrenz traditioneller Medien werden können. Hinsichtlich der künftigen Informationsgewohnheiten gibt es ähnliche Unsicherheiten. Vieles spricht für eine zumindest mittelfristige Koexistenz der bestehenden und neuen Medientypen.

Zeitungsverleger mögen angesichts unklarer ökonomischer Aussichten eine defensive Strategie gegenüber dem neuen Medientyp favorisieren. Sie sollten dabei jedoch beachten, daß die für die Gestaltung eines attraktiven Angebots unvermeidbare Lernkurve einen erfolgreichen kurzfristigen Einstieg erschwert. Zusätzliche Gefahren liegen in der drohenden Machtverschiebung zugunsten der Anbieter elektronischer Diensten und im möglichen Online-Engagement von Anbietern konkurrierender Medientypen im angestammten Verbreitungsgebiet. Die Option des “beobachtenden Abwartens” erscheint deshalb problematisch. Ein zumindest experimentelles Engagement im Bereich computergestützter elektronischer Medien ist derzeit als Option vorzuziehen – und sei es auch nur als Element einer überlegten Defensivstrategie.

3.0 Problemkontext

In diesem Kapitel werden die Randbedingungen der bearbeiteten Aufgabenstellung beschrieben. Eine Darstellung verschiedener Aspekte des World Wide Web (WWW) wurde der eigentlichen Problemanalyse vorangestellt, da spezifische Eigenschaften der WWW-typischen Client-Server-Architektur, des Dokumentenmodells und der Dokumentenbeschreibungssprache direkte Auswirkungen für den Entwurf und die Implementierung des im Rahmen dieser Arbeit entwickelten Systems haben. Es folgt eine Beschreibung der organisatorischen und technischen Bedingungen bei der *tageszeitung* und bei der Prozeßrechnerverbundzentrale der TU Berlin (PRZ). Das Kapitel schließt mit einer Erläuterung der zwischen der *tageszeitung* und dem PRZ vereinbarten Projektzielen und einer Anzahl von Hypothesen, die im Verlauf des Projekts überprüft werden sollen.

3.1 Das World Wide Web

Von einem kurzen historischen Abriss abgesehen, beschränkt sich die folgende Darstellung ausschließlich auf solche Aspekte des WWW, die für die vorliegende Aufgabenstellung von unmittelbarer Bedeutung sind. Es wird nicht versucht, einen umfassenden Überblick über das World Wide Web zu geben.

Das WWW ist ein verteiltes, multimodales (d.h., mehrere Medientypen integrierendes) Informationssystem mit Hypertextstruktur. Die Datenbasis kann hierbei über eine nahezu beliebige Anzahl von Servern verstreut sein. Sie befindet sich in der Regel an anderen geographischen Orten als die auf sie zugreifenden Benutzerapplikationen. Dokumente können verschiedene Medienformen umfassen (z.B. Graphiken, Texte, Audio) und über Verweise (*links*) miteinander verknüpft werden.

Das Konzept und die prototypische Realisierung des WWW als internetbasierter Hypermedia-Architektur stammt von Tim Berners-Lee¹. Sein ursprünglich für das europäische Kernforschungszentrum CERN entworfenes System integriert nahezu alle im Internet verwendeten Kommunikations- und Informationsmechanismen (u.a. *ftp*, *archie*, *telnet*, *mail*, *news*, *gopher*, *wais*) und vereinheitlicht deren Adressierung. Für die Übertragung und Codierung von WWW-Dokumenten wurden zwei plattformunabhängige Protokolle definiert: das *Hypertext Transfer Protocol* (HTTP) für die Kommunikation zwischen den Rechnern der Informationsanbieter (*Server*) und denen der Informationsnutzer (*Clients*) sowie die *Hypertext Markup Language* (HTML) als Strukturbeschreibungssprache für WWW-Dokumente. Diese Beschreibungssprache kann bereits von einfachen WWW-Benutzerinterfaces (*Browsers*) auf ASCII-Terminals interpretiert werden, ist aber auch in der Lage, leistungsfähige Browser auf grafikorientierten Personalcomputern und Workstations anzusteuern. Neben HTML-formatierten Dokumente können WWW-Server und Browser beliebige Dokumententypen definieren, übertragen und (ggf. über externe Applikationen) interpretieren. Dies erlaubt die Integration von Audio-, Video- oder Applikationsdaten.

1. [Cai95]

3.1.1 Das Dokumentenmodell des WWW

Die Dokumentenstruktur des WWW wird gemeinhin mit den Begriffen *Hypertext*– oder *Hypermedia*–*Informationsraum* umschrieben¹. Dabei bezeichnet *Hypermedia* die Erweiterung von Hypertext um Multimedia–Eigenschaften². Der Terminus *Hypertext* selbst wurde 1965 von Ted Nelson geprägt. Er bezeichnet “nicht–sequentiellen Text”, d.h. “Text, der nicht den Einschränkungen einer linearen Struktur unterliegt.”³.

Dies ist eine etwas unbefriedigende Begriffsbildung, da einerseits jeder Text qua definitionem linear ist, andererseits auch papiergebundene Dokumente unter die gegebene Definition fallen: “A magazine layout, with sequential text and inset illustrations and boxes, is thus hypertext. So is the front of a newspaper, and so are various programmed books...”⁴ Das World Wide Web Consortium definiert Hypertext als “Dokument, das Verbindungen zu anderen Dokumenten aufweist”⁵. Da der Begriff der “Verbindung” hierbei nicht präzisiert wird, trifft diese Definition allerdings auch auf traditionell gedruckte wissenschaftliche Veröffentlichungen mit referenzierten Quellen zu. Kuhlen versucht, Hypertext über die *Kohäsivität* der zur Verfügung gestellten informationellen Einheiten zu definieren, d.h., der angestrebten Kontextunabhängigkeit und Abgeschlossenheit der Informationseinheiten (*chunks*). Er kommt allerdings zu dem Schluß, daß eindeutige Abgrenzungen zwischen Text und Hypertext nicht sinnvoll sind⁶.

Der Versuch, Hypertext über *Texteigenschaften* zu definieren, scheint demnach wenig erfolgversprechend zu sein. Aussichtsreicher ist es, die Definition auf die technischen Möglichkeiten zur Textmanipulation abzustellen. Das besondere Merkmal von Hypertext bestünde dann darin, daß ein beliebig strukturierter Korpus schriftlichen oder graphischen Materials⁷ mit technischen Hilfsmitteln für das assoziative Verknüpfen und Referenzieren der in ihm definierten informationellen Einheiten kombiniert wird. Dies erlaubt einen flexiblen Umgang mit den im Korpus enthaltenen informationellen Einheiten, d.h., deren technisch unterstützte *Manipulation* durch Autoren und Leser⁸.

Die Idee einer Kombination von Informations– oder Wissensbasen mit technischen Hilfsmitteln zur assoziativen Verknüpfung und Manipulation geht auf einen 1945 von Vannevar Bush veröffentlichten Artikel zurück⁹. Zwar ist die von Bush vorgeschlagene, mikrofilmorientierte *MEMEX*–Apparatur nie praktisch realisiert worden, doch sein bahnbrechendes Konzept wurde 1963 von Douglas Engelbart aufgegriffen und im Rahmen seines *AUGMENT*–Projekts durch einen computerbasierten Prototypen verwirklicht¹⁰. Bis zum Jahre 1968 wurde dieser Prototyp zu einem Werkzeug für *computer supported collaborative work* (CSCW) geographisch verteilter Personen

1. [Ber01]

2. [Wwc95]

3. ebd.

4. Nelson 1987, zitiert nach [Kuh91], S.37

5. [Wwc01]

6. [Kuh91], S. 37

7. [Bal94]

8. [Kuh91], S.13

9. [Bus45]

10. [Eng63]

ausgebaut.¹ Die von Engelbart eingesetzten Mechanismen waren richtungweisend für die weitere Entwicklung von Arbeitsplatzrechnern, Benutzerschnittstellen und Hypertextsystemen: als Eingabemedium wurde eine Computermaus verwendet, die Benutzeroberfläche arbeitete mit Mehrfenstertechnik und erlaubte die Integration graphischer Darstellung sowie die Verknüpfung heterogener Daten über Zeiger². Die Ideen Bushs, Eisenbarts und Nelsons wurden Anfang der 70er Jahre an den *Xerox Computer Science Laboratories* weiterentwickelt.

Elemente von Hypertextstrukturen

Auch beim Versuch, den Begriff der *Hypertextstruktur* zu definieren, sieht man sich mit Schwierigkeiten konfrontiert. In der Literatur herrscht weder über die Bezeichnung der die Struktur konstituierenden Elemente noch über deren Objektbereich Einigkeit. So läßt Kuhlen die abstrakte mathematische Beschreibung (relationierte Objekte, Graph) lediglich als Veranschaulichung gelten³, ebenso den aus der Wissensrepräsentation bekannten Begriff des (semantischen) Netzwerks⁴. Der von Kuhlen für Hypertextobjekte favorisierte Begriff der *informationellen* oder *Repräsentationseinheit*⁵ ist aus jedoch aus Informatikersicht problematisch: er ist abhängig vom jeweils eingesetzten Hypertextsystem⁶ und basiert auf Kriterien wie “kognitive Einheit” und “kohäsive Geschlossenheit”, die für eine intensionale Definition nicht weiterhelfen⁷.

Wir behalten den Begriff der *informationellen Einheit (IE)* bei, beschränken die Definition jedoch darauf, das eine *IE* ein vom Hypertextsystem referenzierbares Element sein muß. Eine *Hypertextstruktur* wird dann aus einer Anzahl *IEs* (referenzierbarer Elemente) gebildet, zwischen denen *Verweise* bestehen können. Eine derartige Struktur kann durch einen gerichteten Graphen mit einer endlichen Anzahl von *Knoten* modelliert werden. Ein *Knoten* ist jeweils einer informationellen Einheit zugeordnet. Die *Kanten* des Graphen werden durch die *Verweise (links)* zwischen informationellen Einheiten gebildet (vgl. Abb. 1a).

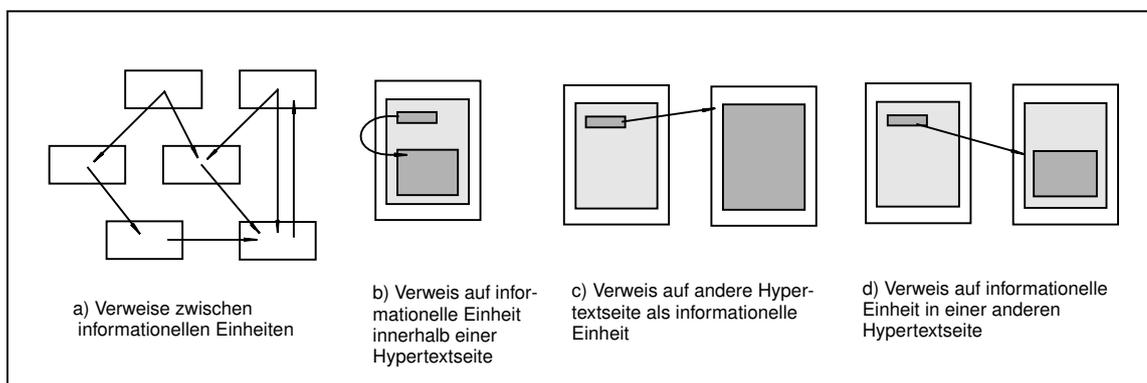


Abb. 1: Verweise zwischen informationellen Elementen in Hypertexten

1. [Bal94]
2. [Kuh91], S.68
3. [Kuh91], S.20f
4. ebd.
5. [Kuh91], S.79ff
6. [Kuh91], S. 84f
7. [Kuh91], S 86ff

Ausgangs- und Zielpunkte von *links* werden als *Anker* bezeichnet. Jeder ausgehenden Anker ist einem *aktiven Bereich* innerhalb einer informationellen Einheit zugeordnet, der in der Regel optisch gekennzeichnet ist. Eine oder mehrere informationelle Einheiten können in einer *Hypertextseite* zusammengefaßt sein. Eine solche Hypertextseite kann in ihrer Gesamtheit als informationelle Einheit referenziert werden, ebenso die informationellen Einheiten, die sie enthält (vgl. Abb. 1b,c,d). Wird der aktive Bereich eines ausgehenden Ankers vom Benutzer aktiviert, führt dies zur Realisierung des mit ihm verbundenen *links*, d.h., die durch den *link* referenzierte informationelle Einheit wird dargestellt. Im Falle des WWW kann ein *link* z.B. auf Text (ggf. mit eingebetteten Graphiken), graphische Informationen, Audio- oder Videodokumente zeigen.

Architektur von Hypertextsystemen

Kuhlen teilt Hypertextsysteme in die drei Komponenten *Autorensystem*, *Navigationssystem* sowie *Hypertext-Managementsystem* auf (vgl. Abb. 2). Das Autorensystem dient dabei der (manuellen oder automatischen) Erzeugung der informationellen Einheiten, der Hypertextseiten und -dokumente. Diese Elemente bilden zusammen mit den zwischen ihnen bestehenden Verknüpfungen die *Hypertext-Basis*. Dies ist die die durch das Hypertext-Managementsystem verwaltete inhaltliche Komponente des Hypertext-Systems. Die Navigationskomponente erlaubt dem Benutzer das Bewegen zwischen und Anzeigen von Elementen der Hypertext-Basis.

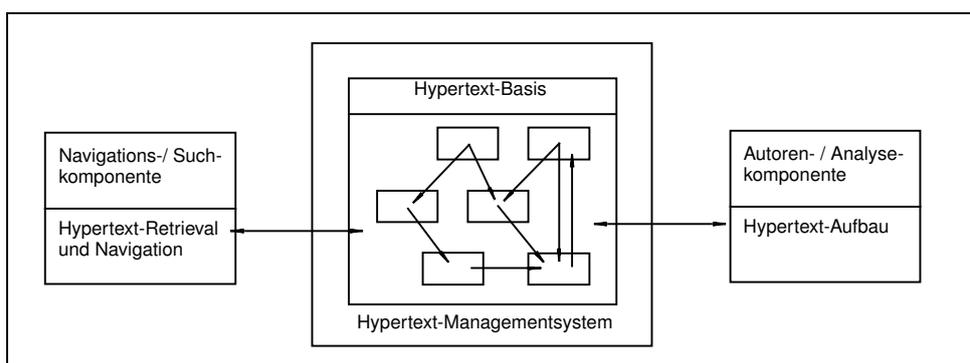


Abb. 2: Architektur eines Hypertextsystems (nach Kuhlen 1991, S. 21)

Kuhlens Darstellung ist für WWW nur eingeschränkt gültig. Erstens lassen sich WWW-Dokumente, -Seiten und -Informationseinheiten auch ohne den Einsatz spezieller Autorenkomponenten erzeugen – ein einfacher Texteditor reicht aus. Zweitens wird die WWW-Hypertext-Basis in der Regel nicht durch ein dediziertes Managementsystem (etwa eine relationale Datenbank) verwaltet, sondern erfolgt hauptsächlich auf der Grundlage von Dateisystem-Mechanismen. *Links* zwischen informationellen Einheiten verkörpern im WWW unidirektionale Relationen.¹ Sie werden nicht als Objekte der Hypertextbasis verwaltet, was dazu führt, daß für ein gegebenes informationelles Element nicht direkt ermittelt werden kann, ob es von anderen Elementen referenziert wird oder nicht.

1. [Kuh91], S.104ff

Der Zugriff auf WWW–Dokumente geschieht über einen Server–Prozeß, der als dynamisches Hypertext–Managementsystem agiert und dessen Hauptaufgabe darin besteht, Dokumenteninhalte an einen Client–Prozeß zu übertragen. Dokumente, auf die durch ein- und denselben Server direkt zugegriffen werden kann, bilden die (relativ zu diesem Server) *lokale Hypertextbasis*. *Interne links* verweisen auf informationelle Einheiten innerhalb, *externe links* informationelle Einheiten außerhalb der lokalen Hypertextbasis¹.

Die möglichen Verbindungsarten (etwa 1:1 oder 1:n) zwischen Knoten werden durch die Topologie des Hypertextsystems definiert. Die im WWW verwendete Strukturbeschreibungssprache HTML gestattet eine unbegrenzte Anzahl von Verbindungen zwischen beliebigen Knoten. Hierdurch können sämtliche Formen gerichteter Graphen erzeugt werden, insbesondere Listen und Bäume. Die spezifische Form eines solchen Graphen bildet die Dokumenten– oder Hypertextstruktur. Standardmäßig sind in WWW nur unspezifizierte, *referentielle* (assoziative) Verknüpfungen vorgesehen² – eine globale Typisierung von Verweisen ist in Ermangelung eines *link–Managements* (s.o.) nicht möglich. Werden mit Hilfe derartiger referentieller *links semantische* Verknüpfungen (z.B. Verweise zur Orientierung in einer hierarchischen Dokumentenstruktur) realisiert³, so muß der Autor dies durch ein sichtbar in das Dokument eingefügtes Symbol (etwa einen Pfeil) oder entsprechenden Text (z.B. “zum Inhaltsverzeichnis”) dafür sorgen, daß ein Benutzer die Funktion des *links* richtig interpretieren kann.

Die Navigation innerhalb von Dokumentenstrukturen kann durch verschiedene Mechanismen erleichtert werden, die teils von den verwendeten Hypertext–Browsern, teilweise durch die Struktur selbst bereitgestellt werden müssen. Hierzu gehören z.B. *History*–Mechanismen und Suchhilfen in Browsern oder Inhaltsverzeichnisse und Rückverweise auf zentralen Startpunkte in Dokumentensammlungen⁴.

Gestaltung von Hypertext–Informationsangeboten

Theoretisch ist mit den in WWW zur Verfügung stehenden Optionen jede durch einen Graphen modellierbare Strukturierung eines Informationsangebotes möglich. Das praktische Problem liegt in der Auswahl einer für das Angebot *geeigneten* Struktur. Diese ist stark abhängig vom konkreten Kontext. Im vorliegenden Fall einer Online–Tageszeitung stellt sich z.B. die Frage, ob und wie weit die Präsentation des Angebot an der Struktur der gedruckten Ausgabe orientiert sein soll. Werden mehrere Artikel pro Dokument angeboten oder nur einer? Wird nach Seiten oder Rubriken gegliedert? Sollen zusätzliche Verweise zwischen einzelnen Artikeln und externe *links* integriert werden?

Ob ein Benutzer eine Dokumentenstruktur als akzeptabel empfindet, hängt z.T. von rein technischen Faktoren ab. Dazu gehören die zur Verfügung stehende Netzbandbreite zwischen Client und

1. Alle Begriffsdefinitionen orientieren sich an dem vom *World Wide Web Consortium* veröffentlichten Glossar *Hypertext Terms*
URL: <http://www.w3.org/hypertext/WWW/Terms>

2. ebd.

3. [Kuh91], S.106

4. [Ber02]

Server oder die durch den Browser angebotenen Navigations- und Darstellungsmechanismen. Eine im Rahmen dieser Arbeit untersuchte Fragestellung ist, ob Inhalte parallel mehreren verschiedenen, spezifischen Benutzerbedürfnissen angepaßten Dokumentenstrukturen präsentiert werden sollten. Dies soll in einem Prototyping-Prozeß geklärt werden, der die Erfahrungen, Wünsche und Kritik der Online-Leser berücksichtigt.

3.1.2 Das Client-Server Modell des WWW

WWW ist ein verteiltes Informationssystem, in dem abfragende Instanzen (*Clients*) Informationen und Dienste von anbietenden Instanzen (*Servern*) anfordern können. Im einfachsten Fall wird ein vom Client angeforderter Dienst von einem einzigen Server erbracht (vgl Abb. 3a). Die im WWW

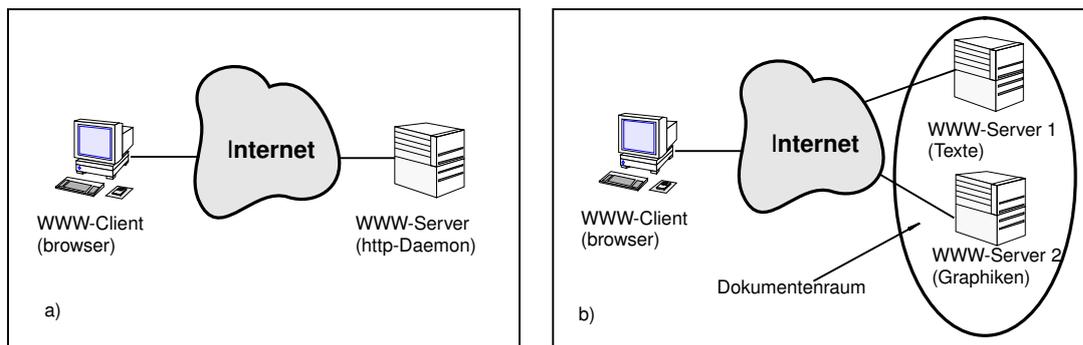


Abb. 3: a) Client-Server Modell des WWW b) verteilte Dokumentenstruktur

verwendete Strukturbeschreibungssprache HTML erlaubt jedoch auch die Konstruktion von Dokumentenstrukturen, deren informationelle Einheiten sich auf unterschiedlichen Servern befinden und einen gemeinsamen Dokumentenraum bilden. Dies ist selbst für Teile einzelner Hypertextseiten – etwa Texte und Grafiken – möglich (vgl Abb. 3b) Schließlich kann ein Server für die Bereitstellung eines Dienstes selbst weitere Server in Anspruch nehmen. Er agiert in diesem Fall gleichzeitig als Server und als Client (vgl Abb. 4). Server, die die Aufgabe haben, Anfragen von Clients an andere Server weiterzuleiten und das Resultat der Anfrage zurück zum Client zu übermitteln, werden als *Proxy* oder *Gateway* bezeichnet.

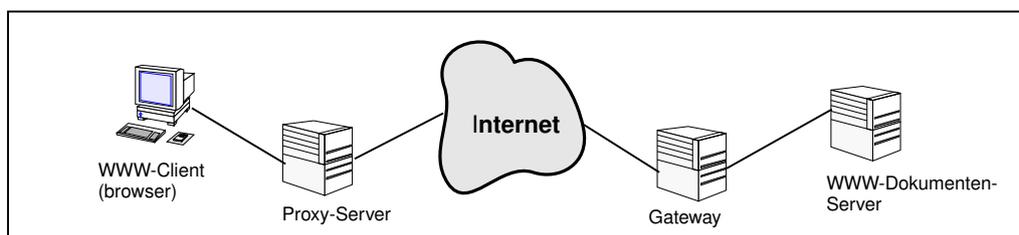


Abb. 4: Client-Server Modell mit vermittelndem Proxy und Gateway

Die Kommunikation zwischen Client und Server erfolgt über das plattformunabhängige Protokoll HTTP¹. Als Clients fungieren in erste Linie *Browser*, also Benutzerapplikationen für den Zugriff und die Darstellung von WWW-Dokumenten. Allerdings können auch andere Applikationen

1. [Ber95a]

Anfragen an WWW-Server stellen. Voraussetzung hierfür ist, daß sie über eine HTTP-Protokoll-schnittstelle verfügen. Diese Schnittstelle wird der Applikation in der Regel über eine Funktionsbibliothek wie z.B. *libwww*¹ zur Verfügung gestellt.

Server Requests

Die am häufigsten gestellte Anfrage an WWW-Server fordert zur Übertragung einer lokalen Dokumenten-Datei auf. Die gewünschte Datei wird dabei durch einen *Uniform Resource Locator* (URL) spezifiziert². Vor einer Übertragung der angeforderten Datei wird (unter UNIX) vom Server u.a. überprüft, ob er autorisiert ist, auf den entsprechenden Teil des Dateibaums zuzugreifen, ob die Attribute der angeforderten Datei den öffentlichen Zugriff gestatten und ob die Datei mit zusätzlichen Sicherheitsattributen (z.B. Passwortschutz) verknüpft ist.

Eine weiterer, häufig auftretender Anfragetyp schließt die Übertragung benutzerdefinierter Daten vom Client an den Server ein. Dies geschieht z.B. dann, wenn der Inhalt eines ausgefüllten HTML-Formulars übermittelt werden soll. Die übertragenen Daten werden hierbei nicht direkt vom HTTP-Serverprozeß ausgewertet, sondern über das sogenannte *Common Gateway Interface* (CGI) des Servers an eine für die Verarbeitung der Daten vorgesehene Applikation weitergeleitet. Der Serverprozeß extrahiert hierzu zunächst die in einem HTTP-Protokollelement gekapselten Daten und überführt diese in ein standardisiertes Format, das der verarbeitenden Applikation als Eingabe dient. Danach erfolgt die Verarbeitung durch die Applikation, deren Ausgabedaten vom Server zurück zum Client übertragen werden.

Proxies und Caching

Das in Abb. 3 dargestellte Schema illustriert das funktionale Client-Server Modell, abstrahiert aber von einer wichtigen Eigenschaft des Internet. In der Regel befindet sich eine mehr oder weniger große Anzahl von Routern zwischen dem Client- und dem Serverrechner. Die Datenübertragungsrates zwischen Client und Server ist dabei abhängig von der Auslastung jedes dazwischenliegenden Netzsegments. Da beim im Internet verwendeten Transportprotokoll TCP/IP derzeit keine *End-to-End* Ressourcenallokierung möglich ist, kommt es zu stark unterschiedlichen Ladezeiten für Dokumente. Sind Zwischensegmente stark belastet, kann auf weit entfernte oder schmalbandig an das Internet angeschlossene Server oft nur schwer oder überhaupt nicht zugegriffen werden.

Um diesem Problem zu begegnen, werden clientseitig in zunehmendem Maße Cache-Server eingesetzt, die einmal geladene Seiten lokal zwischenspeichern (vgl Abb. 4 auf S. 24). Derartige Server fungieren als WWW-Proxy und bedienen oft eine gesamte Internet-Domäne. Ein Cache-Server leitet nach außen gehende Anfragen nur dann weiter, wenn sich das angeforderte Dokument nicht bereits in seinem Speicher befindet. Derartige Server können hierarchisch gestaffelt werden³. Anfragen nach Dokumenten, die im lokalen Cache des direkt zugeordneten Servers nicht vorhanden sind, werden in diesem Fall zunächst an nahegelegene und mit hoher Bandbreite ver-

1. [Fry95]

2. [Ber94], vgl auch [Fie95]

3. Har96]

bundene Cache-Server weitergeleitet. Erst wenn auch diese die Anfrage nicht bedienen können, wird das Dokument an seinem Originalstandort angefordert.

Cache-Server finden auch serverseitig Verwendung (vgl. Abb. 4 auf S. 24). Hier werden sie in erster Linie eingesetzt, um den HTTP-Server zu entlasten und die Zugriffszeit auf Dokumente zu verkürzen. Ein weiterer Grund für den serverseitigen Einsatz von WWW-Proxies ist oft der Wunsch nach erhöhter Netzwerksicherheit. Proxies können in einer Domäne als "Schaltstelle" für alle HTTP-basierten Dienste fungieren. Eventuelle Angriffe, die Lücken im HTTP-Protokoll auszunutzen versuchen, sind leichter festzustellen, wenn es dafür nicht mehrere, sondern nur einen Angriffspunkt gibt. Diese Aussage gilt analog für Client-Proxies¹.

Clientseitiges Caching läßt sich durch bestimmte Maßnahmen 'überlisten', z.B. durch Einbinden dynamischer Elemente oder Zufallsdaten in Dokumente oder geschickt gewählte Verfallsdaten im Dokumentenkopf. Diese Verfahren erhöhen jedoch die Auslastung von Dokumentenserver und Netzwerken.

Protokollierung von Zugriffen

Informationsanbieter und Betreiber von WWW-Servern sind naturgemäß daran interessiert, Aufschlüsse über die Nutzung ihres Angebots zu erhalten. Zu diesem Zweck wird jede Anfrage eines Clients in Protokolldateien des Servers festgehalten. Aus diesen Dateien geht z.B. hervor

- von welchem IP-Interface eine Anfrage gestellt wurde,
- zu welchem Zeitpunkt dies geschah,
- welcher Browsertyp verwendet wurde,
- auf welche Dokumente oder Dienste zugegriffen und
- aus welchem Dokument ein Link auf den Server aktiviert wurde.

Die Einträge sind anonym, d.h., es kann nicht ermittelt werden, welche Benutzer auf welche Dienste zugegriffen haben (obwohl einige Server das in RFC 1413² spezifizierte Benutzer-Identifikationsprotokoll unterstützen, wird diese Option clientseitig in aller Regel nicht eingesetzt)³.

Anonyme Log-Einträge erschweren den Versuch, Folgen von Aktionen zu rekonstruieren, die einzelne Nutzer während ihrer Navigation im Informationsangebot vollzogen haben. Der Einsatz von Cache-Servern, Proxies und Gateways macht eine individualisierte Rekonstruktion des Benutzerverhaltens in vielen Fällen unmöglich. Anfragen an den WWW-Server werden in diesem Fall nicht mehr direkt vom IP-Interface des clientseitigen Rechners, sondern vom dazwischengeschalteten Proxy gestellt. Wieviele Benutzer sich hinter einem solchen Proxy verbergen, kann auf direktem Wege nicht festgestellt werden. Der Betreiber eines WWW-Servers kann deshalb nicht ermitteln, wieviele Nutzer auf seine Dokumente zugegriffen haben. Es läßt sich hierbei lediglich

1. [Ber96], S. 34

2. [Joh93]

3. [Hed95]

eine untere Grenze angeben. Der Einsatz von Cache-Servern und Proxies bringt an dieser Stelle demnach Begleiterscheinungen mit sich, die für Informationsanbieter höchst unerwünscht sind.

Zusammenfassend ergeben sich aus dem Client-Server-Modell von WWW folgende Probleme:

- Designentscheidungen werden durch unbekannte Ladezeiten für Dokumente erschwert. Ein komplexes, graphisch orientiertes Angebot kann für Benutzer mit schmalbandigen Internet-Verbindungen unzumutbar sein.
- Der clientseitige Einsatz von Caching-Mechanismen erschwert das Nachvollziehen von Benutzeraktionen und die Bestimmung der Zugriffsanzahl auf der Seite des Servers.
- Maßnahmen, um clientseitiges Caching mittels dynamisch generierter Dokumente zu umgehen, senken die Performance des Servers und erhöht den Datentransfer im Netzwerk.

3.1.3 Die Strukturbeschreibungssprache HTML

Die *Hypertext Markup Language* (HTML) ist neben dem *Hypertext Transfer Protocol* (HTTP) die zweite konstituierende Komponente des WWW. HTML ist eine plattformübergreifende Sprache zur Beschreibung von Hypertextdokumenten. Sie erlaubt u.a. die Darstellung von News, Mail, Text- und Hypermediadokumenten, Ergebnissen von Datenbankabfragen und interaktiven Formularen.

Obwohl HTML seit 1990 eingesetzt wurde, existierte lange Zeit keine verbindliche Definition dieser Sprache. Mittlerweile existiert eine Spezifikation, die als ISO-Standard vorgeschlagen¹ sowie im RFC 1866² festgelegt wurde. Hiernach müssen HTML-Dokumente dem im ISO-Standard 8879:1986 festgeschriebenen SGML-Format³ entsprechen.

Standardisierte HTML Elemente

Die im RFC 1866 beschriebene, "offizielle" Struktur von HTML ist Ergebnis jahrelanger Diskussionen und Auseinandersetzungen, die hier nicht nachgezeichnet werden können⁴. Es sollen an dieser Stelle lediglich einige der wichtigsten Designentscheidungen aufgezählt werden:

- Eine wohldefinierte Konvertierung zwischen Stilelementen gängiger Textverarbeitungsprogramme und HTML-Elementen sollte möglich sein. HTML geht deshalb von einer Folge von Absätzen mit jeweils spezifischen Stilen aus, die durch Sprachelemente gekennzeichnet werden⁵.
- HTML beschreibt lediglich die *logische* Struktur eines Textes. Im Gegensatz zu einer Seitenbeschreibungssprache wie Adobe Postscript existieren in HTML z.B. keine Layoutdirektiven für die Einstellungen von Fonts, Schriftgrößen oder -farben⁶.

1. [Pri96]

2. [Ber95b]

3. [Iso86]

4. [Wwc02]

5. [Ber91]

6. [Ber91]

- HTML ist ein statisches Dokumentenformat. Im Gegensatz zu “programmierbaren” Formaten wie etwa *troff* oder *TeX* gibt es keine Sprachelemente, die als Anweisungen für die Generierung von Dokumententext verwendet werden können¹.

Ausgehend von diesen Designentscheidungen enthält HTML z.B. Direktiven zur Kennzeichnung von Überschriften, Listen, Zitaten, zum Einfügen graphischer Informationen und einiger Gestaltungselemente (z.B. horizontalen Linien). Die Kontrolle über die Darstellung von Fonts beschränkt sich auf die Angabe des Typs (etwa fett, kursiv, unterstrichen). Es können aktive Bereiche im Text oder in Grafiken definiert und mit Referenzen auf Dokumente verknüpft werden. Für die Gestaltung interaktiver Formulare stehen Standardelemente wie z.B. Felder zur Texteingabe, Auswahlmensüs und stilisierte Schaltknöpfe (*Buttons*) zur Verfügung. In einem explizit gekennzeichneten Vorspann kann ein HTML–Dokument eine Reihe von Struktur– und Metainformationen enthalten, etwa Angaben zu Titel und Autor, Indexangaben, Informationen zur Dereferenzierung von links oder ein Verfallsdatum. Dieser Vorspann kann automatisiert ausgewertet werden.

Nicht standardisierte HTML – Elemente

Bedingt durch die vergleichsweise späte Standardisierung von HTML erfüllt nur ein Bruchteil der im WWW verfügbaren Dokumente die im RFC 1866 festgelegten Vorgaben. Dies gilt insbesondere für die obligatorische Angabe des Dokumententyps. Viele Dokumente enthalten darüberhinaus eine Reihe “proprietärer” Sprachelemente, die nicht im RFC 1866 definiert sind und nicht von allen Browsern interpretiert werden können. Hierzu gehören z.B. Direktiven zur Definition des Hintergrundlayouts einer Dokumentenseite, zur spalten– und tabellenorientierten Textdarstellung oder zur – im HTML–Standard explizit ausgeklammerten – Definition von Fontgrößen und –farben. Der Einsatz dieser nicht standardisierten Elemente ist auf die große Verbreitung des *Netscape*–Browsers zurückzuführen, dessen Entwickler 1994 zusätzliche Layout–Direktiven einföhrteten, ohne den bis dahin üblichen Weg über das Standardisierungskonsortium zu gehen.

Obwohl diese einseitige Festlegung neuer Direktiven in Teilen der WWW–Gemeinde für einigen Unmut gesorgt hat, können viele der nicht im RFC festgelegten HTML–Elemente als Teil eines informellen WWW–“Industriestandards” angesehen werden. Dies ist in erster Linie auf das rapide Wachstum bei den nicht UNIX–basierten WWW–Browsern zurückzuführen (die diese Sprachelemente in aller Regel interpretieren können). Benutzten bei einer Ende 1993 vorgenommenen ersten Bestandsaufnahme noch 88% der Befragten ein UNIX–Derivat als Betriebssystem und *Mosaic* als Browser, sanken diese Zahlen bis zum Ende 1994 bereits auf 58% bzw. 44%. Bei einer weiteren Untersuchung im April 1995 waren die entsprechenden Anteile auf 3% bzw. 9% gefallen. Es dominieren derzeit der *Mozilla*–Browser von *Netscape* und *Microsoft Windows Explorer*².

Diese Zahlen erklären, warum inzwischen (Mai 1996) viele Informationsanbieter kein Problem darin sehen, in ihren Dokumenten nicht–standardisierte HTML–Elemente zu verwenden: es sind

1. [Con94]

2. [Ber95], S.33

genau diese Layoutdirektiven, die eine professionell wirkende Gestaltung von WWW-Seiten erlauben (z.B. textumflossene Bilder, mehrspaltigen Satz oder Tabellen).

3.2 Organisatorische und technische Rahmenbedingungen

Die *tageszeitung* wird mit Hilfe eines hybriden Redaktionssystem produziert. Die Erfassung der Artikel erfolgt in den einzelnen Redaktionen mittels DOS- oder WINDOWS-basierter Personalcomputer. Die Texte werden von Setzern in ein *taz*-eigenes Beschreibungsformat übertragen. Zu jedem Artikel existieren eine Reihe zusätzlicher Metainformationen, die z.B. Aufschluß über die Informationsquelle, die Zeitungsseite und -rubrik, die Autoren und die Länge des Textes geben. Die gesetzten Beiträge werden durch einen Konverter in ein Austauschformat (*Maker Interchange Format, MIF*) übertragen. Das Layout der Druckseite erfolgt mit dem Desktop Publishing Paket *FrameMaker*.

Die Produktionszentrale der *tageszeitung* liegt in Berlin, redaktionelle Außenstellen u.a. in Frankfurt, Hamburg und Bremen. Das Verfassen, Redigieren und Setzen des überregionalen Teils erfolgt dezentral, d.h., sowohl in der Produktionszentrale wie auch in den Außenstellen. Sämtliche Artikel der aktuellen Ausgabe werden am Ende eines Produktionstages elektronisch nach Berlin übertragen und dort in einer *provisorischen Archivdatei* zusammengefaßt. Das Format dieser Datei ist identisch mit dem Satzformat. Die digitale Archivierung der Artikel erfolgt nicht parallel zur Produktion der Ausgabe, sondern erst einige Tage nach deren Drucklegung. Hierbei wird jeder Beitrag mit einer Anzahl von Schlagworten versehen und um eventuell fehlende Metainformationen (etwa Seiten- oder Rubrikangaben) ergänzt. Mehrfach überredigierte Artikel befinden sich oft in mehr als einer Ausführung in der Archivdatei. Diese Duplikate werden bis auf den tatsächlich erschienenen Beitrag entfernt. Die provisorische Archivdatei enthält die Texte, jedoch nicht die Illustrationen der Druckausgabe, da die *taz* nur das Copyright an den Artikeln besitzt. Aus urheberrechtlichen Gründen können keine Agenturphotos in die Online-Ausgabe übernommen werden, sondern lediglich die *taz*-eigenen Karikaturen.

Die provisorische Archivdatei kann zum Zeitpunkt des Andrucks, d.h., etwa gegen 20 Uhr, zum PRZ TU Berlin übertragen werden. Dort bildet sie die Ausgangsbasis für die Erstellung der WWW-gestützten Ausgabe. Die Berliner Redaktion der *tageszeitung* verfügt über UNIX-basierte Server mit TCP-IP Verbindung zu einem Internet-Dienstanbieter. Diese war allerdings während des Projektzeitraums so schmalbandig ausgelegt, daß sie zwar für Dateiübertragung und *email*, nicht jedoch für das Aufsetzen eines WWW-Servers geeignet war. Im Unterschied dazu verfügt das PRZ der TU Berlin über einen 2-MBit-Anschluß an das Wissenschaftsnetz. Es betreibt einen *Apache* http-Server¹ auf einer SPARC20 Workstation unter SunOS 4.1.3 und bietet WWW als betreuten 24-Stunden-Dienst an. Die am PRZ vorhandene Server- und Netzinfrastruktur ist damit geeignet, um auch eine große Anzahl von Zugriffen zuverlässig und schnell bedienen zu können.

Das bei der *tageszeitung* verwendete Satzformat ist ASCII-basiert (Zeichensatz: ISO-Latin1) und enthält besonders gekennzeichnete, menschenlesbare Formatierungsangaben innerhalb des Fließ-

1. <http://www.apache.org/>

textes. Der größte Teil der Formatierungsangaben bezieht sich auf die *logische* Struktur des Textes, ein Umstand, der einer Konvertierung nach HTML entgegenkommt. Die dem Artikel zugeordneten Metainformationen erlauben eine Strukturierung des Ausgabe nach Rubriken, Seitenzahlen etc. Problematisch ist indes, daß diese Metainformationen unvollständig sein oder sogar komplett fehlen können.

Es ergaben sich folgende Randbedingungen für Entwurf und Implementation des Online-Angebots:

- Da die Layoutinformationen in den Fließtext eingestreut sind, können zur Konvertierung textbasierte Parser und Filter eingesetzt werden
- Es steht keine formale Syntax der beim Setzen der *tageszeitung* verwendeten Layout-Beschreibungssprache zur Verfügung. Diese muß empirisch, d.h., anhand von Beispieldateien, ermittelt werden
- Die von der *tageszeitung* erstellte provisorische Archivdatei enthält artikelbezogene Layoutinformationen, jedoch keine Angaben über die Gestaltung ganzer Druckseiten oder Rubriken.
- Es können Metainformationen fehlen und Artikel mehrfach auftreten.
- In Ermangelung von Schlagwortzuordnungen kann lediglich eine stichwortbezogene Volltextsuche angeboten werden

3.3 Ziele des Projekts

Das hier beschriebene Projekt ist Ergebnis einer Forschungskoooperation zwischen der berliner *tageszeitung* (*taz*) und der Prozeßrechnerverbundzentrale der Technischen Universität Berlin. Gemeinsames Ziel beider Partner war es, eine vollautomatisch erstellte, täglich aktualisierte Online-Ausgabe der *tageszeitung* für das World Wide Web zu konzipieren, praktisch zu realisieren und zu testen. Auf der Grundlage einer Leserbefragung und der während der Testphase anfallenden Logdateien sollen eine Anzahl von Fragestellungen untersucht werden, die sowohl aus verlegerischer Perspektive wie auch aus derjenigen von Informationssystementwicklern und –betreibern von Interesse sind und im Folgenden aufgeführt werden. Die Evaluierung der Projektziele und Hypothesen erfolgt in Kap. 8.0 dieser Arbeit.

Technische Machbarkeit und Aufwandsermittlung:

Vorgesehen ist ein verteilt arbeitendes System, bei dem die Generierung der Daten, die Konversion in das Zielformat HTML und das Anbieten der Ausgabe im WWW auf unterschiedlichen Servern erfolgt. Optional soll die Ausgabe auf mehreren geographisch verteilten Servern gespiegelt werden können. Hierbei sollen Mechanismen für die Authentifizierung der angebotenen Informationen entworfen werden, mit denen die Benutzer die Validität der ihnen von Spiegelservern angebotenen Daten überprüfen können. Schließlich sollen Leserbriefe, ein Leserbriefforum und eine Stichwortsuchfunktion als Dienste zur Verfügung gestellt werden.

Wie in Kap. 2.2 dargelegt wurde, muß bei einem Online-Informationsangebot mit einer Zunahme von Leseranfragen gerechnet werden. Andererseits besagen am PRZ gemachten Erfahrungen mit WWW-basierten Informationssystemen, daß der Aufwand für die laufende technische Wartung des Angebots vergleichsweise gering ist. Unsere Hypothese ist, daß der größte Teil des Arbeitsaufwands während des laufenden Betriebs anfällt, allerdings nicht für die technische Systemwartung, sondern für die Leserbetreuung.

Strukturierung und Gestaltung:

Es soll eine Dienstgestaltung gefunden werden, die intuitiv zu bedienen ist und eine möglichst einfache Navigation in der Dokumentenstruktur erlaubt. Das Angebot sollte dabei möglichst browserunabhängig sein. Von Seiten der *tageszeitung* wird gewünscht, daß das gewählte Layout in ausreichendem Umfang Corporate-Identity-Merkmale transportiert. Insbesondere soll untersucht werden, ob es nötig ist, bei der Strukturierung und Gestaltung des Angebots unterschiedliche Dokumentenladezeiten zu berücksichtigen. Über den Einsatz von nicht-standardisierter HTML-Direktiven kann pragmatisch entschieden werden, wenn sich herausstellen sollte, daß die überwiegende Mehrzahl der Benutzer mit Browsern arbeitet, die solche Direktiven interpretieren können.

Unsere Hypothese ist, daß sich wegen der stark unterschiedlichen technischen Voraussetzungen bei den Lesern (Browsertypen, Übertragungsbandbreite) keine gestalterische Gesamtlösung finden, die allen genannten Kriterien gleichermaßen gerecht wird. Wahrscheinlich ist vielmehr, daß verschiedene parallele Strukturierungen des Inhalts angeboten werden müssen, die jeweils spezifische Bedürfnisse der Leser erfüllen.

Nutzung:

Es soll untersucht werden, ob und in welchem Umfang das Online-Angebot überhaupt angenommen wird, von wo aus die Zugriffe erfolgen und zu welchen Zeitpunkten dies geschieht. Weiterhin soll ermittelt werden, ob und in wieweit von zusätzlichen Diensten des Online-Mediums Gebrauch gemacht wird. Hierzu gehören z.B. die Option, direkt per Leserbrief Stellung zu nehmen, in einem öffentlichen Forum zu diskutieren oder Such- und Filterfunktionen zu benutzen.

Unsere Hypothese ist, daß wegen des vereinfachten Feedbacks zwischen Produzenten und Leserschaft die Anzahl direkter Stellungnahmen und Anfragen deutlich über den entsprechenden Zahlen des gedruckten Mediums liegt. Wir vermuten, daß die Online-Leser ihre am Umgang mit gedruckten Tageszeitungen orientierten Informationsgewohnheiten beibehalten (vgl. Kap. 2.1). Such- und Filterfunktionen dürften nur von einem kleinen Teil der Leserschaft eingesetzt werden.

Vermarktung:

Anhand einer Umfrage und der Auswertung von Leserbriefen soll ermittelt werden, wie die Online-Leser die Qualität des Dienstes im Vergleich zur gedruckten Ausgabe bewerten und welche technischen Optionen gewünscht werden. Von besonderem Interesse ist die Frage, ob Leser zu einer Bezahlung des Angebots bereit wären und wie hoch sie den Wert des angebotenen Dienstes einschätzen.

Während einer mehrmonatigen Testphase kann beobachtet werden, ob das Online-Angebot die Verkaufszahlen der Druckausgabe beeinflusst, und ob das WWW-Angebot Auswirkungen auf die Nutzung ihres kommerziellen Angebots bei der über Bildschirmtext (BTX) zugreifbaren GENIOS-Datenbank hat. Schließlich sollte die Option der Werbung im Online-Medium und die der Vermarktung über kommerzielle Online-Dienste beurteilt werden.

Da die im Rahmen des WWW-Angebots zur Verfügung gestellten Such- und Filterfunktionen weit hinter denjenigen kommerzieller Datenbankanbieter zurückbleiben, ist unsere Hypothese, daß die Nutzung des kommerziellen Datenbankangebots nicht beeinflusst wird. Gleiches nehmen wir für die Verkaufs- und Abonnementszahlen an, da die Gestaltung und Benutzerfreundlichkeit der Online-Ausgabe mit derjenigen der gedruckten Ausgabe nicht zu vergleichen ist. Bezüglich der kommerziellen Vermarktbarkeit des WWW-Angebots vermuten wir, daß nur eine kleine Anzahl von Kunden bereit sein dürfte, das Angebot zu bezahlen.

4.0 Systementwurf

Dieses Kapitel beschreibt die *Gestaltung* des Online-Angebots sowie den *technischen Entwurf* der Systemkomponenten. Die *Gestaltung* umfaßt die Strukturierung des Informationsangebots, die Anbindung der in Kap. 3.3 genannten Dienste und das Layout der Dokumentenseiten. Der *technischen Entwurf* unterscheidet folgende sieben funktionale Komponenten:

1. Dateitransfer und Dateispiegelung,
2. Konvertierung des Satzformats nach HTML,
3. Generierung der Artikelstruktur,
4. Dienste (Leserbriefe, Leserumfrage und Diskussionsforum),
5. Stichwortsuche,
6. Authentifizierungsmechanismen,
7. statistische Auswertung der Logdateien.

Alle Erläuterungen beziehen sich auf die *Endversion* des im Verlauf mehrerer Prototyping- und Testzyklen erstellten Systems. Prototyping und Test werden in Kap. 6.0 beschrieben.

4.1 Gestaltung des Informationsangebots

In Kap. 4.1.1 wird die *logische* Struktur der Dokumente und Dienste beschrieben. Neben der Haupthierarchie werden verschiedene Alternativen für die Gestaltung des aktuellen Angebots vorgestellt. Kap. 4.1.2 befaßt sich mit dem *Layout* der Dokumente und Dienste.

4.1.1 Strukturierung der Dokumente

Den zentralen Ausgangspunkt des Online-Angebots bildet die Hauptseite (*Homepage*) der *tageszeitung*. Jedes Dokument der Online-Ausgabe ist von diesem Knoten aus erreichbar, entweder direkt oder über Zwischenschritte.

Das Angebot selbst gliedert sich in einen allgemeinen und einen aktuellen Teil. Der *allgemeine* Teil umfaßt das Impressum und Hintergrundinformationen, Verweise auf andere Medien und auf archivierte Artikel, ein Lesebriefformular für allgemeine Anfragen und den Fragebogen. Die Dokumente in dieser Sektion dienen hauptsächlich dazu, neuen Lesern die Orientierung zu erleichtern. Sie sind direkt von der Hauptseite aus aufrufbar.

Der *aktuelle* Teil des Angebots enthält die jeweils gültige Tagesausgabe sowie diejenige des Vortags. Abb. 5 auf S. 34 stellt Struktur der Haupthierarchie dar.

Für die Strukturierung der Tagesausgabe werden drei Alternativen vorgestellt. Da der Umfang einer Tagesausgabe relativ begrenzt ist (ca. 150 kB Text), können alle drei Alternativen parallel zur Verfügung gestellt werden. Die Wahl der für ihn optimalen Form der Darstellung wird dem Benutzer überlassen. In den folgenden drei Absätzen werden zuerst alle drei Alternativen erläutert. Danach wird der integrierte Entwurf präsentiert.

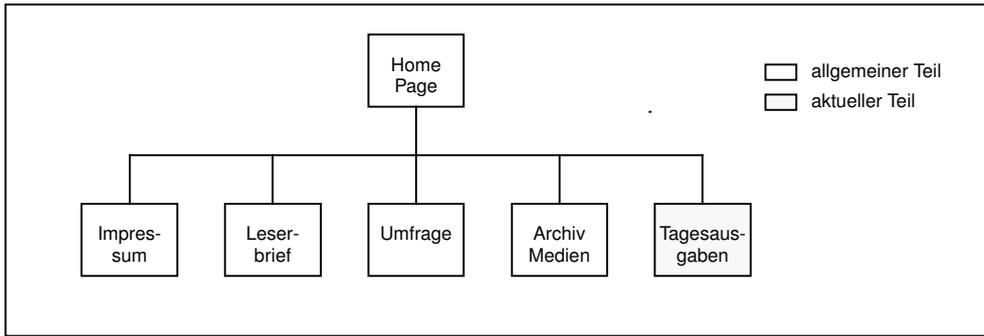


Abb. 5: Haupthierarchie des Online-Angebots

Seitenorientierte Strukturierung:

Die Anzahl der generierten Dokumente entspricht der Seitenzahl der Druckausgabe. Ein generiertes Seitendokument enthält sämtliche Artikel der entsprechenden Druckseite. Zusätzlich wird ein seitenbezogenes Gesamtverzeichnis erzeugt, das die Titelzeilen aller Artikel enthält. Jede Titelzeile in diesem Gesamtverzeichnis wird mit einem *link* auf den ihr zugeordneten Artikel innerhalb des Seitendokuments versehen (vgl. dazu Abb. 1 auf S. 21, Grafik d). Bei der Auswahl eines Artikels wird dem Benutzer immer der *gesamte* Inhalt der Seite übermittelt. Bei der Darstellung der Seite durch den Browser wird der gewählte Artikel in dessen aktiven Fensterbereich eingeblendet. Um das “Blättern” zwischen Seiten der gedruckten Ausgabe in der Online-Ausgabe nachzubilden, werden aufeinanderfolgende Seitendokumente durch speziell gekennzeichnete *links* miteinander verbunden (vgl. Abb. 6). Das “Überfliegen“ einer Seite ist durch Vor- und Zurückrollen (*scrollen*) innerhalb des Dokuments möglich. Diese Funktion wird von allen WWW-Browsern zur Verfügung gestellt.

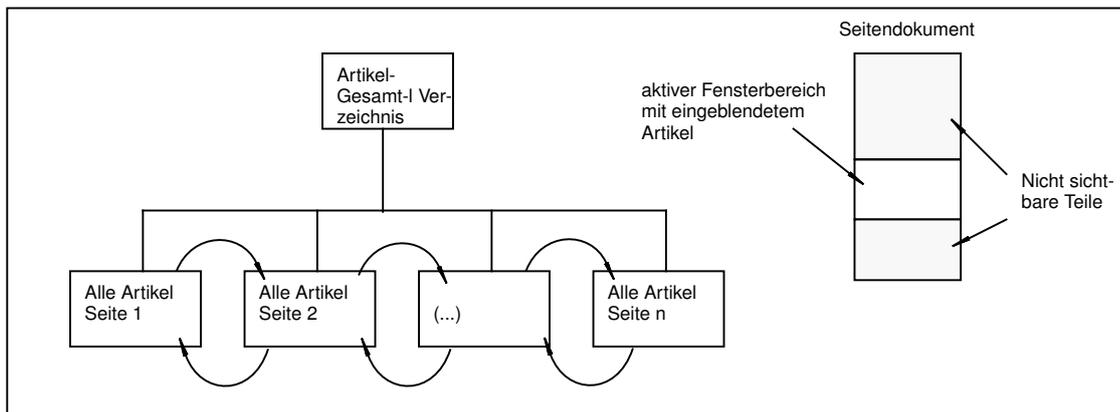


Abb. 6: Seitenorientierte Strukturierung mit Einblendung des gewählten Artikels im aktiven Fensterbereich

Artikelorientierte Strukturierung:

Die Anzahl der generierten Dokumente entspricht der Anzahl der Artikel in der Druckausgabe. Ein Dokument enthält also jeweils nur einen Artikel. Das zusätzlich generierte, nach Rubriken

geordnete Gesamtinhaltsverzeichnis enthält die Titelzeilen aller Artikel. Jede Titelzeile in diesem Gesamtinhaltsverzeichnis ist mit einem *link* verbunden, der auf das zugehörige Artikeldokument verweist. Um das *Browsen* zwischen einzelnen Artikeln einer Rubrik zu ermöglichen, werden aufeinanderfolgende Artikeldokumente durch speziell gekennzeichnete *links* miteinander verbunden (Abb. 7)

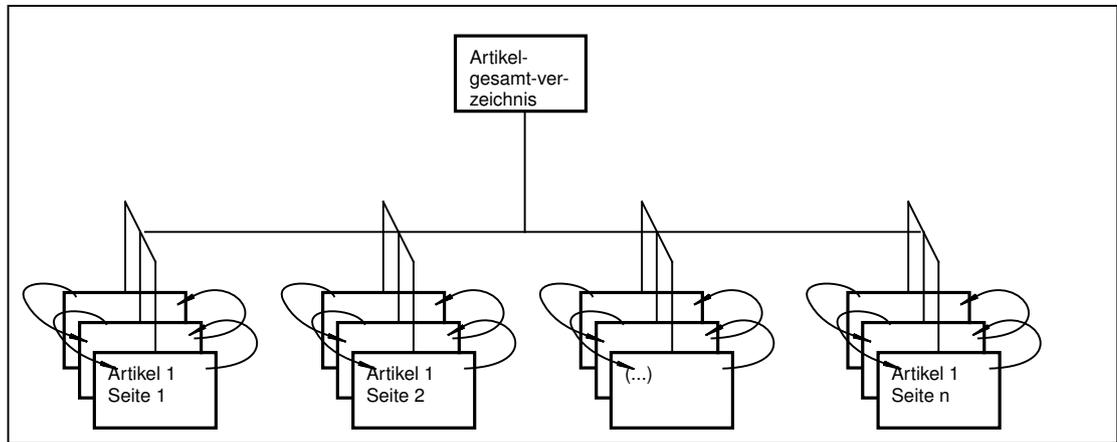


Abb. 7: Artikelorientierte Strukturierung

Rubrik- und artikelbezogene Strukturierung:

Als zusätzliches, interessenbezogenes Gliederungselement werden Dokumente generiert, die jeweils ein rubrikbezogenes Inhaltsverzeichnis enthalten.

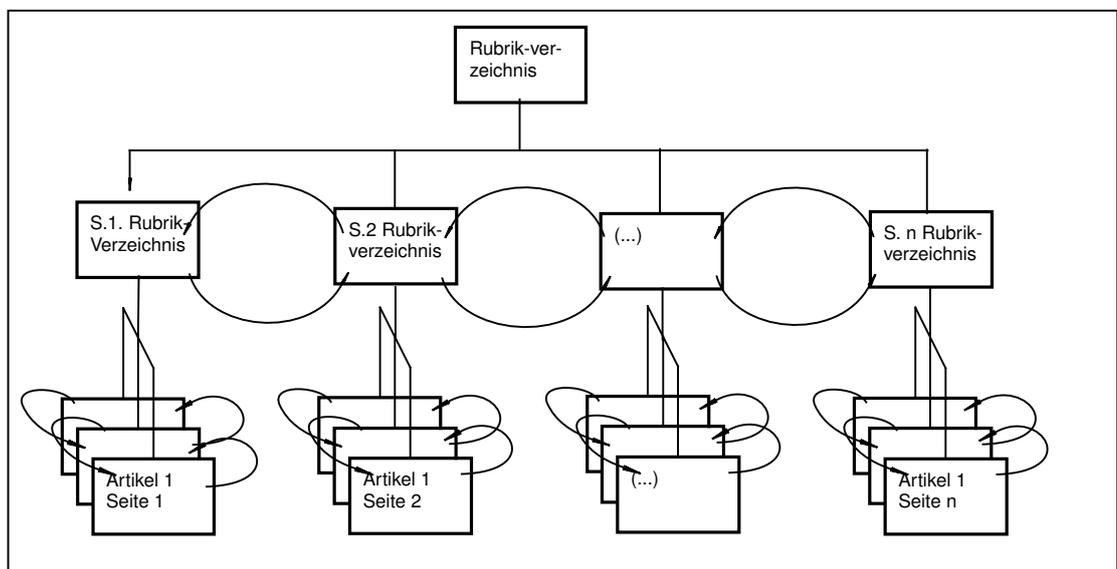


Abb. 8: Rubrik- und artikelorientierte Strukturierung

In der Regel entsprechen Rubriken (etwa "Ausland", "Meinung und Diskussion" oder "Aktuelles") einer gesamten Seite der gedruckten Ausgabe. Die Titelzeilen der Rubrikinhaltsverzeichnisse sind mit *links* ausgestattet. Diese verweisen auf Dokumente, die den jeweils zugeordneten Artikel

enthalten. Aufeinanderfolgende Artikel einer Rubrik werden durch speziell gekennzeichnete *links* miteinander verknüpft. Gleiches gilt für die Inhaltsverzeichnisse aufeinanderfolgende Rubriken. Alle rubrikbezogenen Inhaltsverzeichnisse können über ein Hauptinhaltsverzeichnis erreicht werden (vgl. Abb. 8)

Integration der Strukturalternativen

Wie aus der bisherigen Darstellung der drei Strukturen hervorgeht, werden die Artikel der Tagesausgabe in zwei Dokumentenformen angeboten: im *Seitenzusammenhang* und als *Einzeldokumente*. Die gewünschte Form der Darstellung kann in einem übergeordneten Hauptdokument gewählt werden. Die Option der *rubrikbezogenen* Darstellung von Teilinhaltsverzeichnissen ist über Verweise aus den Gesamtverzeichnissen der beiden Hauptalternativen (seitenorientierte bzw. artikelorientierte Darstellung) zu erreichen. Sie ist als experimentell zu betrachten, da die zur Generierung eines echten Rubrikenverzeichnisses nötigen Informationen in den von der *tageszeitung* übertragenen Archivdaten teilweise fehlen. Abb. 9 stellt die Struktur der drei integrierten Strukturalternativen dar.

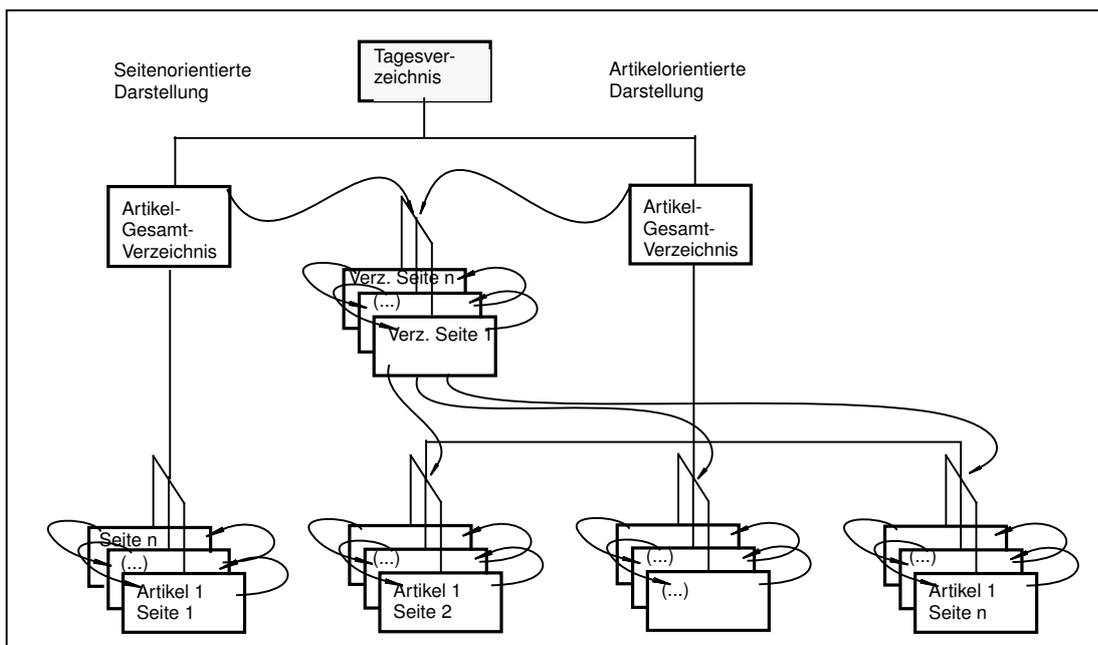


Abb. 9: Integrierte Struktur der Tagesausgabe mit drei Darstellungsalternativen

Anbindung der Dienste

Als Dienste werden Leserbriefformulare, ein Diskussionsforum und eine Volltextsuche angeboten. Alle diese Funktionen sind, unabhängig von der durch den Benutzer gewählten Darstellungsform, von jedem Inhaltsverzeichnis und Artikel aus zu erreichen. Durch diese Art der Anbindung ist es z.B. möglich, Überschrift und Datum eines Artikels als Parameter an das angebundene Leserbriefformular zu übergeben, so daß der Benutzer diese nicht noch einmal per Hand eintragen muß.

Für Volltextsuche und Leserbriefe werden HTML-Formulare eingesetzt, die serverseitig von Shell- oder Perl-Skripten ausgewertet werden. Für die Realisierung des Diskussionsforums wird eine geringfügig modifizierte Version des Programmpakets *hypermail*¹ eingesetzt, das in der Lage ist, *email*-Archivdateien (*mailfolder*) in HTML-Dokumente umzusetzen und hierfür datums-, autoren- und themenorientierte Indices zu erstellen.

4.1.2 Layout des Informationsangebots

Einen wichtigen Ausgangspunkt beim graphischen Entwurf des Layouts war der Einsatz möglichst browser-unabhängiger Mechanismen (vgl. Kap. 3.1.3). So wurden z.B. keine Direktiven für mehrspaltige Darstellung oder rechtsseitig begrenzte Textblöcke eingesetzt, da diese von verschiedenen Browser nicht korrekt interpretiert werden können. Die optische Gestaltung der Tagesausgabe sollte sich an derjenigen der gedruckten Version orientieren, ohne jedoch mit graphischen Elementen überladen zu sein. Gleichzeitig sollte dem Benutzer eine einheitliche Schnittstelle zur Navigation innerhalb des Angebots zur Verfügung gestellt werden.

Optische Elemente

Um die Ladezeiten für Dokumente möglichst gering zu halten, wurden nur wenige Grafiken eingesetzt. Ein großes Logo mit dem Namen der *tageszeitung*, das sich auch auf der Titelseite der gedruckten Ausgabe findet, bildet den Kopf jeder Dokumentenseite. Ein miniaturisiertes Logo

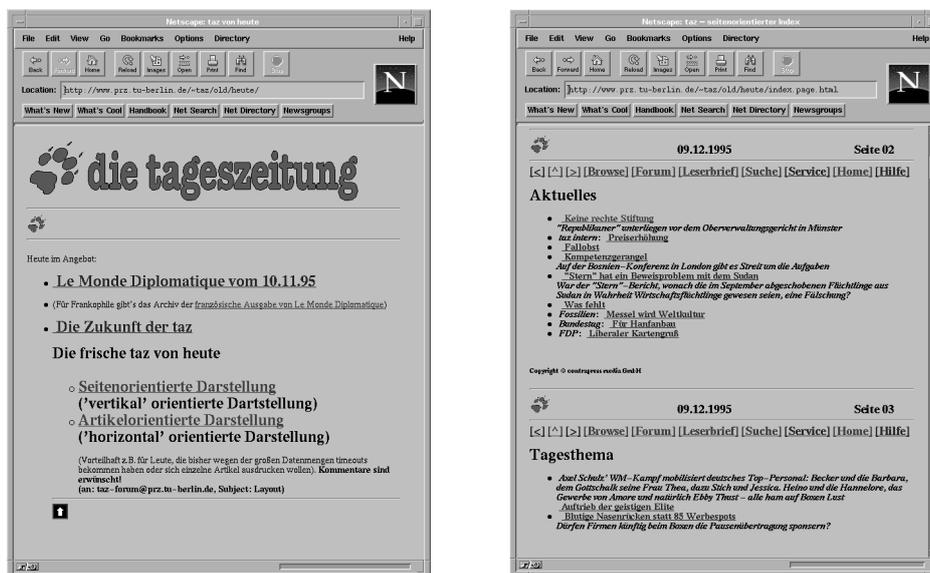


Abb. 10: Layout der Hauptseite der Tagesausgabe und des seitenorientierten Artikelverzeichnisses

ohne Text dient der Kennzeichnung von Artikelanfängen oder zusammengehörenden Passagen in Inhaltsverzeichnissen. Als weiteres Gliederungsmerkmal werden doppelte horizontale Linien eingesetzt, die das Datum und die Seitenzahl enthalten. Dies bildet eine gewisse optische Entsprechung zum oberen Seitenrand der gedruckten Ausgabe. Der linke Teil von Abb. 10 zeigt die

1. [EIT95]

Hauptseite der Online-Tagesausgabe mit den beiden *links* auf das seiten- bzw. artikelorientierte Angebot (unten) und Verweisen auf aktuelle Zusatzinformationen.

Der rechte Teil von Abb. 10 bildet den Ausschnitt eines Gesamtverzeichnisses ab. Die Titelzeilen verschiedener Seiten werden durch horizontale Balken voneinander abgesetzt. Unter diesen Balken befindet sich eine Navigationsleiste, die im nächsten Abschnitt beschrieben wird.

Navigationselemente

Um die Orientierung und Bewegung innerhalb des Angebots zu erleichtern, wird jedem Artikel und jedem seiten- bzw. rubrikbezogenen Teil eines Inhaltsverzeichnisses eine Reihe von Navigations-elementen zugeordnet. Von einer Ausnahme abgesehen, ist das Aussehen und die Funktion dieser Elemente überall identisch.

Abb. 11 gibt die Anordnung der Navigationselemente in Inhaltsverzeichnissen und Artikeln bei *seitenorientierter* Darstellung wieder. Um Ladezeiten zu minimieren und den Zugriff mit ASCII-orientierten Browsern zu erleichtern, wurde auf den Einsatz von Graphiken oder *clickmaps* verzichtet.

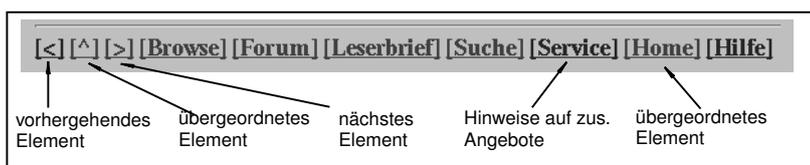


Abb. 11: Navigationsleiste bei seitenorientierter Darstellung

Die *links* [>] und [<] erlauben das Fortbewegen zwischen benachbarten Elementen derselben Hierarchiestufe (vgl. Abb. 9 auf S. 36). In Inhaltsverzeichnissen sind dies die Verzeichnisse der vorhergehenden bzw. folgenden Seite, bei Artikeldokumenten die jeweils benachbarten Artikeldokumente. Beim jeweils ersten Element einer Hierarchiestufe ist der *link* [<] deaktiviert; entsprechendes gilt für das jeweils letzte Element und den *link* [>]. Durch Aktivieren des *links* [^] gelangt der Benutzer in die jeweils nächsthöhere Hierarchiestufe, d.h. von einem Artikel zum jeweiligen Eintrag im Inhaltsverzeichnis und von einem Inhaltsverzeichnis zur Hauptseite der aktuellen Angebots. Das Element [**Browse**] erlaubt das Umschalten von seiten- zu artikelorientierter Navigation. Dies ist z.B. dann nützlich, wenn nur ein einzelner Artikel einer Seite abgespeichert oder ausgedruckt werden soll. Arbeitet ein Benutzer bereits mit artikelorientierter Navigation, so befindet sich an der Stelle des [**Browse**] - Elements das Navigationselement [**Paged**], mit dem in die seitenorientierte Darstellung gewechselt werden kann (vgl. Abb. 12).



Abb. 12: Navigationsleiste bei artikelorientierter Darstellung

Leserbriefe und Diskussionsforum

Das hier beschriebene Eingabeformular dient mehreren Zwecken. Erstens können mit ihm zur *Veröffentlichung* bestimmten Nachrichten an die Leserbriefredaktion der *taz* und an das Online-Diskussionsforum erstellt und versandt werden. Es kann jedoch auch zum Erstellen und Versenden *nicht zu veröffentlichender* Mitteilungen an die Autoren benutzt werden. Entsprechende

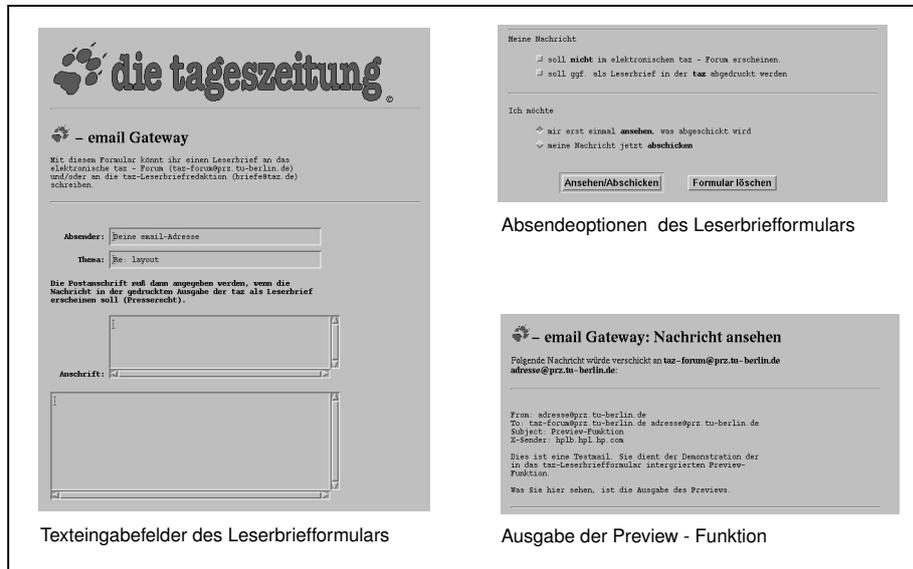
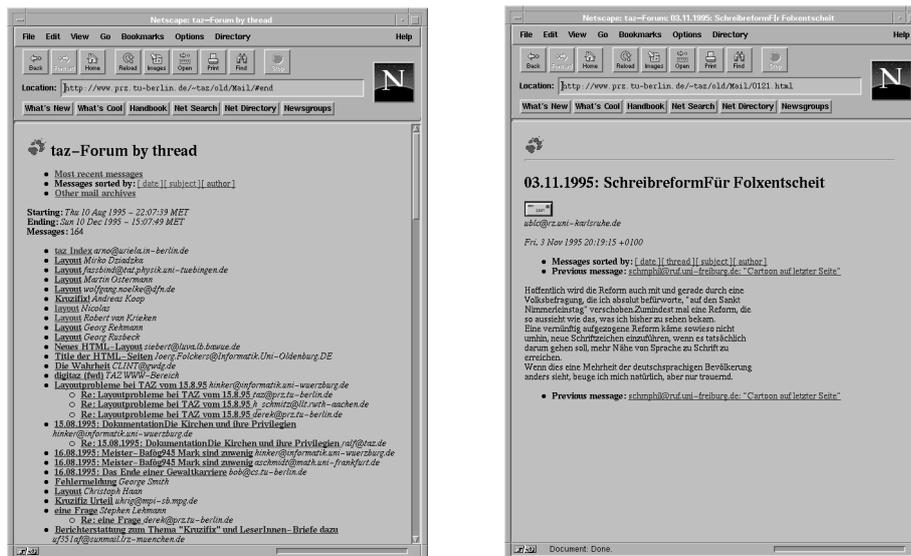


Abb. 13: Leserbrief und Preview

Optionen können innerhalb des Formulars gewählt werden. Zusätzlich gibt es eine *Preview*-Option, mit der sich ein Benutzer die von ihm verfaßte Nachricht vor dem eigentlichen Versenden in ihrer Originalform ansehen kann. Der linke Teil von Abb. 13 zeigt die für die Texteingabe



Gliederung der Diskussionsforums

Nachricht im Diskussionsforum

Abb. 14: Layout des Diskussionsforums

vorgesehenen Felder, der rechte obere Teil die durch *buttons* repräsentierten Optionen zum Absenden der Nachricht. Rechts unten ist die Ausgabe der Preview-Funktion dargestellt.

Die Gestaltung des Diskussionsforums ist durch das verwendete Paket *hypermailer* weitgehend vorgegeben. Dieses Programm wird so angepaßt, daß in allen Indices und Artikel das Logo der *tageszeitung* eingefügt wird. Der Benutzer kann wählen, ob er auf das Forum über einen themen-, autoren- oder datumsorientierten Index zugreifen möchte.

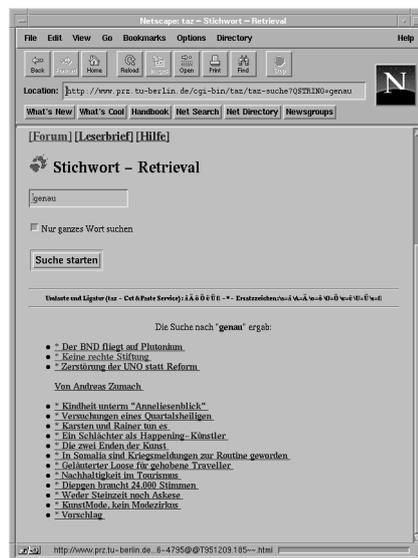
Die linke Hälfte von Abb. 14 auf S. 39 zeigt das Layout des Diskussionsforums. Im abgebildeten Fall sind Nachrichten themenorientiert angeordnet. Einrückungen deuten an, daß an dieser Stelle zu vorausgehenden Nachrichten Stellung genommen wurde (*followup*). Die rechte Hälfte von Abb. 14 gibt einen Artikel innerhalb des Diskussionsforums wieder. Durch Aktivierung des symbolisierten Briefumschlags wird das Leserbriefformular aufgerufen. Hierdurch können Benutzer Kommentare zu bereits vorhandenen Nachrichten hinzufügen.

Suchfunktion

Es wird nur eine einfache Stichwortsuche in der aktuellen Tagesausgabe zur Verfügung gestellt, da ein ausgefeilterer Dienst mit einem über kommerzielle Datenbanken zugänglichen Angebot der *tageszeitung* kollidieren würde. Die einzige wählbare Option besteht in der Einschränkung der Suche auf ganze Worte. Abb. 15 zeigt das Layout des Suchformulars und die formatierte Ausgabe einer Suchanfrage.



Formular für Stichwortsuche



Resultat einer Stichwortsuche

Abb. 15: Layout der Stichwortsuche

4.2 Entwurf der technischen Komponenten

Die technischen Elemente des Systems können in zwei Funktionsgruppen unterteilt werden. Die erste Gruppe umfaßt jene Komponenten, die – normalerweise einmal pro Tag – im Batchbetrieb eingesetzt werden. Hierzu gehören die Dateiübertragung, –konvertierung und –spiegelung sowie das Generieren von Authentifizierungsmerkmalen. Die zweite Gruppe wird von interaktiven Komponenten gebildet, die Bestandteile der Online–Angebots sind, d.h., von den Funktionen für Leserbrief, Diskussionsforum und Stichwortsuche. Einen Sonderfall bildet die statistische Auswertung der Logdateien, die lediglich bei Bedarf erfolgt.

4.2.1 Dateitransfer und Spiegelung

Diese Systemkomponente ist für die Übertragung der Archivdatei von der *tageszeitung* zum PRZ, für das Anstoßen der Formatkonvertierung, für die Sicherung der Archivdatei und der Ausgabe vom Vortag sowie für die Übertragung der HTML–formatierten Ausgabe auf weitere Server zuständig. Dieser Prozeß wird im Folgenden kurz *Transfer* oder *Dateitransfer* genannt.

Der Transfer der Archivdatei von der *taz* zum PRZ erfolgt über die UNIX–Mechanismen *remote copy* und *remote shell*. Als Voraussetzung hierfür ist die *tageszeitung* mit einem eigenen Benutzerbereich am PRZ ausgestattet, der so konfiguriert ist, daß die genannten Operationen ausführbar sind. Sobald die provisorische Archivdatei mit den Artikeln der Tagesausgabe auf dem Server der *tageszeitung* vorliegt (vgl Kap. 3.2), wird sie mit den zur aktuellen Ausgabe gehörenden Grafiken zur aktuellen Transferdatei zusammengefaßt und komprimiert. Danach wird versucht, eine Verbindung zum Server des PRZ aufzubauen und den Transfer auszulösen. Schlägt die Verbindungsaufnahme, die Dateiübertragung oder die Konvertierung fehl, wird der Systemadministrator der *taz* und der WWW–Administrator am PRZ per *email* benachrichtigt. Der Transferversuch wird in regelmäßigen zeitlichen Abständen wiederholt. Sollte er nach einem oder mehreren vergeblichen Versuchen gelingen, wird der WWW–Administrator am PRZ durch eine zweite *email* davon in Kenntnis gesetzt, daß sich die Situation normalisiert hat.

Dieses Transferverfahren läßt sich bei der Spiegelung der Ausgabe auf anderen Servern nicht immer einsetzen, da Serveradministratoren im Interesse der Systemsicherheit oft keine *remote shell* Kommandos für externe Benutzer zulassen. In diesem Fall kann eine alternative Transfermethode verwendet werden. Hierbei werden die Daten zunächst kryptographisch signiert, um die Authentizität des Absenders zweifelsfrei nachzuweisen. Danach erfolgt die eigentliche Datenübertragung an den Zielservers per *email* oder *ftp*. Auf Seiten des Zielservers wird nach abgeschlossener Übertragung die Signatur überprüft und im Erfolgsfall ein Prozeß zur Weiterverarbeitung der Datei angestoßen.

Die Weiterverarbeitung der von der *tageszeitung* übertragenen Transferdatei durch einen Batchjob am PRZ geschieht in folgenden Schritten:

1. Die Daten werden dekomprimiert und in Einzeldateien aufgespalten.
2. Aus der Archivdatei werden die einzelnen Artikel extrahiert. Dabei wird das Satzformat der *tageszeitung* in HTML-formatierte Artikeldateien umgesetzt. Die Archivdatei wird danach in einem dafür vorgesehenen Verzeichnis abgelegt.
3. Die HTML-formatierten Artikeldateien einer Seite werden zu jeweils einem Seitendokument zusammengestellt. Parallel dazu werden die Inhaltsverzeichnisse generiert.
4. Das Verzeichnis mit den Artikeln der Vortages wird archiviert und durch dasjenige mit der aktuellen Ausgabe ersetzt.
5. Die Hauptseite für die Tagesausgabe wird aktualisiert.
6. Die HTML-formatierten Dokumente und Grafiken der aktuellen Ausgabe werden ggf. zusammengefaßt, komprimiert und an Spiegelserver übertragen.

Nach Beendigung von Schritt 5 ist die aktualisierte Ausgabe für Online-Leser verfügbar. Da zu Spiegelservern die bereits konvertierte Ausgabe übertragen wird, sind hier außer der bereits erwähnten Überprüfung der Signatur nur die Schritte 1 bis 3 erforderlich.

4.2.2 Formatkonvertierung

In der provisorischen Archivdatei der *tageszeitung* sind Anfang und Ende jedes Artikels durch eingestreute Direktiven markiert, ebenso die zum Artikel gehörenden Metainformationen (z.B. Autor, Seitennummer und Rubrik), Titel und Untertitel sowie der eigentliche Text des Beitrags.

Die Formatkonvertierung besteht in der Umsetzung des Archivformats der *tageszeitung* nach HTML und der Generierung der in Kap. 4.1.1 beschriebenen Struktur. Der HTML-Konverter erzeugt hierbei eine Anzahl von Zwischendateien, die in einem darauffolgenden Schritt vom Strukturgenerator in die eigentlichen WWW-Seiten umgesetzt werden. Abb. 16 veranschaulicht diesen Prozeß.

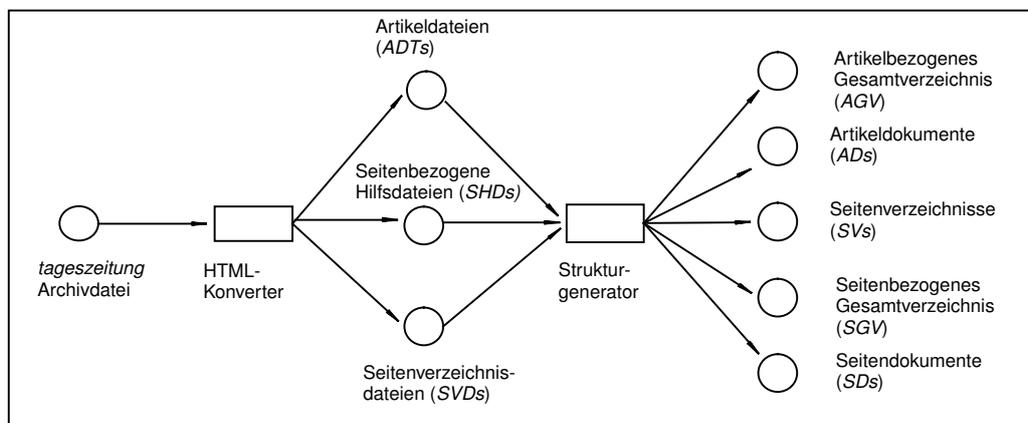


Abb. 16: Arbeitsschritte und Dateien bei der Formatkonvertierung

HTML-Konverter

Das bei der *tageszeitung* verwendete Satzformat ähnelt der Sprache TeX: logische Einheiten werden durch Klammern zusammengefaßt, wobei der öffnenden Klammer die Bezeichnung des Makros folgt. Die Verschachtelung von Makros ist möglich. Von der *tageszeitung* konnte leider keine konsistente Grammatik ihres Satzformats geliefert werden. Dies machte eine empirische Ermittlung der Grammatik nötig.

Der notwendige Parser wurde im Trial-and-Error Verfahren entwickelt und langwierigen Tests mit verschiedenen Archivdateien unterzogen, um sicherzustellen, daß er der real existierenden Grammatik des Satzformats genügt. Die ermittelte Grammatik ist in Anhang A wiedergegeben. Zu beachten ist, daß die Interpretation von Layoutdirektive innerhalb des Fließtextes teilweise kontextabhängig ist, d.h., in Bezug auf den übergeordneten logischen Block vorgenommen werden muß.

Der HTML-Konverter besteht aus einem *Parser*, der die Layoutdirektiven des *taz*-Satzformats im Fließtext lokalisiert, und einem *Generator*, der die vom Parser gefundenen Direktiven in HTML-Formatanweisungen umsetzt. Die Ausgabe des Formatkonverters besteht aus einer Anzahl von Dateien, die einzelne Artikel im HTML-Format enthalten, und eine weiteren Anzahl artikelbezogener Dateien, die HTML-formatierte Titelzeilen und Unterüberschriften enthalten, aus denen Inhaltsverzeichnissen aufgebaut werden. Diese Zwischendateien bilden die Eingabe des im nächsten Abschnitt beschriebenen Strukturgenerator.

Strukturgenerator

Der Strukturgenerator kombiniert die vom Formatkonverter gelieferten Artikeldateien (im folgenden *ADTs*) und Seitenverzeichnissedateien (*SVDs*) zu WWW-Dokumentenseiten. Die generierten Verweis-Strukturen entsprechen dabei der in Kap. 4.1.1 gegebenen Beschreibung. Sie werden in folgenden Schritten erzeugt:

Artikelbezogene Strukturierung

1. Alle *SVDs* werden um Navigationselemente und graphische Elemente (vgl Kap. 4.1.2) zu Seitenverzeichnissen (*SDs*) erweitert. Durch Aneinanderfügen der einzelnen *SVDs* wird das artikelbezogene Gesamtverzeichnis (*AGV*) erstellt.
2. Alle Artikeldateien (*ADTs*) werden um Navigationselemente erweitert.

Seitenbezogene Strukturierung

3. Die *ADTs* werden nach Seitenzahl geordnet. Aus den einer Seite zugeordneten *ADTs* wird jeweils ein Seitendokument (*SD*) generiert. Die auf Artikeldokumente (*ADs*) weisenden *links* der Navigationselemente werden durch *links* auf Artikel innerhalb der entsprechenden Seitendokumente ersetzt.
4. Es wird eine Kopie des artikelbezogenen Gesamtverzeichnisses (*AGV*) erzeugt. Die in diesem Verzeichnis enthaltenen *links* auf *ADs* werden durch *links* auf Artikel in Seitendokumenten (*SDs*) ersetzt und so damit das seitenbezogene Gesamtverzeichnis (*SGV*) erstellt.

Rubrik- und artikelbezogene Strukturierung

Diese ist implizit bereits nach Schritt 3 vorhanden, da zu diesem Zeitpunkt bereits die nötigen Rubrik/Seitenverzeichnisse und Artikeldokumente vorliegen. Es müssen lediglich die auf Elemente in Gesamtinhaltsverzeichnissen verweisenden *links* in der Navigationszeile durch Verweise auf die entsprechenden Seitenverzeichnisse ersetzt werden.

4.2.3 Authentifizierung

Zur Erstellung von Authentifizierungsmerkmalen wird das Programm *PGP* von Phil Zimmermann verwendet, das eine Art informellen Standard im Internet darstellt. Mit diesem Programm lassen sich zu beliebigen Nachrichten sogenannte Signaturen bilden. Diese bestehen aus einem über die zu signierende Nachricht gebildeten Hashwert, der mit dem privaten RSA-Schlüssel des Signierenden verschlüsselt ist. Die Signatur wird in der Regel der Nachricht hinzugefügt. Kennt der Empfänger der Nachricht den öffentlichen Schlüssel des Signierenden, so kann er mit dem Programm PGP überprüfen, ob der Inhalt der Nachricht verfälscht wurde.

Für die Authentifizierung von Artikeln sind mehrere Mechanismen vorstellbar. Die einfachste Lösung besteht darin, jedem Artikel eine Signatur anzufügen und die Verifikation dem Benutzer zu überlassen. Eine serviceorientierte Alternative hierzu könnte darin bestehen, Authentifizierungsmerkmale von Artikeln lediglich auf explizite Benutzeranforderung hin zu übertragen. Eine weitere derartige Möglichkeit wäre, den zu überprüfenden Artikel durch den Benutzer an einen beim PRZ stationierten Authentifizierungsservice schicken zu lassen, ihn dort zu verifizieren und dem Benutzer das Resultat mitzuteilen.

Da Authentifizierung ein Mehrwertdienst ist, für den in Zukunft u.U. Gebühren erhoben werden können, sind serviceorientierte Lösungen mittelfristig zu favorisieren. Da in dieser Arbeit jedoch lediglich die prinzipielle technische Machbarkeit der Integration von Authentifizierungsmerkmalen in HTML-Dokumente demonstriert werden soll, wird die einfachste Option gewählt: die Signatur wird im Dokument mitübertragen.

Das Signieren von Nachrichten ist ein aufwendiges Verfahren, das selbst auf leistungsfähigen Rechnern mehrere zehn Sekunden in Anspruch nehmen kann (vorbereitende Tests ergaben ca. 15 Sekunden pro Datei auf einer unbelasteten HP-720 Workstation). Bei einem durchschnittlichem Umfang von etwa 200 Artikeln pro Ausgabe bedeutet dies, daß eine vorherige Signierung *aller* Artikel den Zeitpunkt der Veröffentlichung um etwa eine Stunde verzögern würde. Aus diesem Grund wird die Signierung von Artikeln lediglich experimentell erprobt. Die prototypische Realisierung des Verfahrens wird in Kap. 5.3 erläutert.

4.2.4 Leserbriefe und Diskussionsforum

Leserbriefe können sich auf den Inhalt von Artikeln ebenso beziehen wie auf allgemeine Aspekte des WWW-Dienstes. Bisherige Erfahrungen der *tageszeitung* mit *electronic mail* haben gezeigt, daß dieses Medium auch zur Benachrichtigung der Autoren oder der Vertriebsabteilung genutzt wird. Für die Zuordnung von Leserbriefen zu bestimmten Artikeln oder Abteilungen ist es daher

essentiell wichtig, Informationen über den Kontext zu übermitteln, auf den sich eine Nachricht bezieht.

Entsprechende Kontextinformationen können beim Aufruf von Formularen automatisch mitgeliefert werden. Im Falle von Leserbriefen bietet sich an, die Subject-Zeile vorzuinitialisieren. Dies schließt allerdings die Verwendung der in Browser integrierten *mailto*-Funktion aus. Stattdessen wird das in Kap. 4.1.1 beschriebene Formular verwendet. Bei einer dynamischen Erzeugung des Formulars durch ein serverseitiges Modul können Initialisierungsvariablen ausgewertet werden. Die Werte dieser Variablen werden bereits bei der Formatkonvertierung gesetzt, und zwar so, daß beim Aufruf des Leserbriefformulars aus einem Artikel heraus dessen Datum und die Überschrift in der Subject-Zeile an den Server übertragen werden.

Zur weiteren Vorsortierung von Nachrichten werden, abhängig von den im Leserbriefformular gesetzten Optionen (gedruckter Leserbrief, Privatnachricht, Forum) unterschiedliche Empfängeradressen verwendet. Die nötige Zuordnung des Empfängers geschieht serverseitig durch ein eigenständiges Modul (vgl Abb. 17).

Da bei Nachrichten, die über WWW-Formulare versandt werden, serverseitig keine Überprüfung des tatsächlichen Absenders möglich ist, wird an die Absenderadresse eine Kopie der an die *tageszeitung* übermittelten Nachricht verschickt. Mißbräuchliche Verwendung von Absenderadressen kann so zwar nicht unterbunden, jedoch vom Inhaber der entsprechenden Adresse immerhin bemerkt werden.

Da für das Diskussionsforum ein bereits existierendes Programmpaket (*hypermail*) verwendet wird, entfällt der Entwurf dieser Komponente. Diese muß lediglich so angepaßt werden, daß sie das entworfene Formular für Leserbriefe überträgt.

Abb. 17 stellt das Zusammenwirken der technischen Komponenten für Leserbriefe bei Client und Server im zeitlichen Verlauf dar.

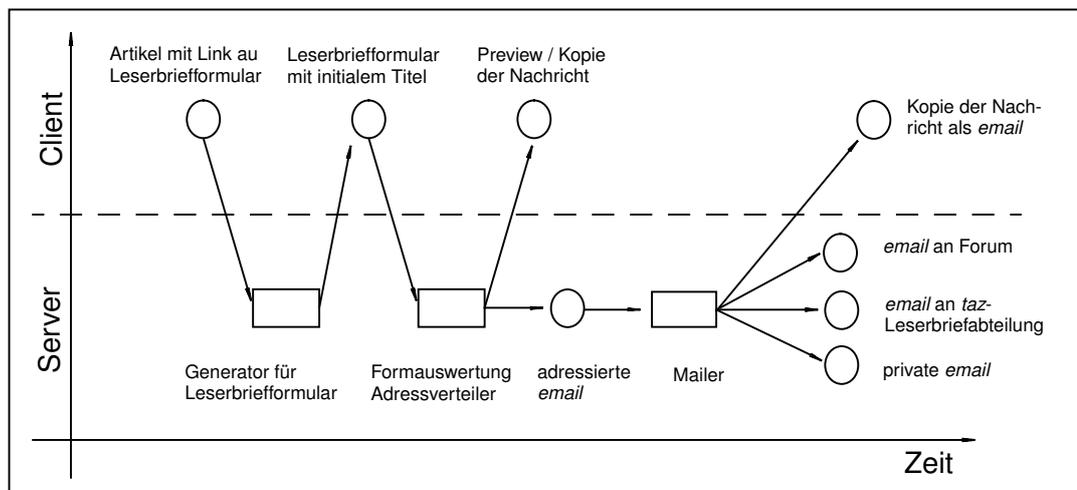


Abb. 17: Zusammenwirken der Komponenten für Leserbriefe

4.2.5 Stichwortsuche

Da eine Stichwortsuche nur in der jeweils aktuellen Ausgabe möglich sein soll und keine komplexen Suchfunktionen zur Verfügung gestellt werden (vgl. Kap. 4.1.2), beschränken sich die Aufgaben des entsprechenden Moduls auf eine einfache Volltextsuche.

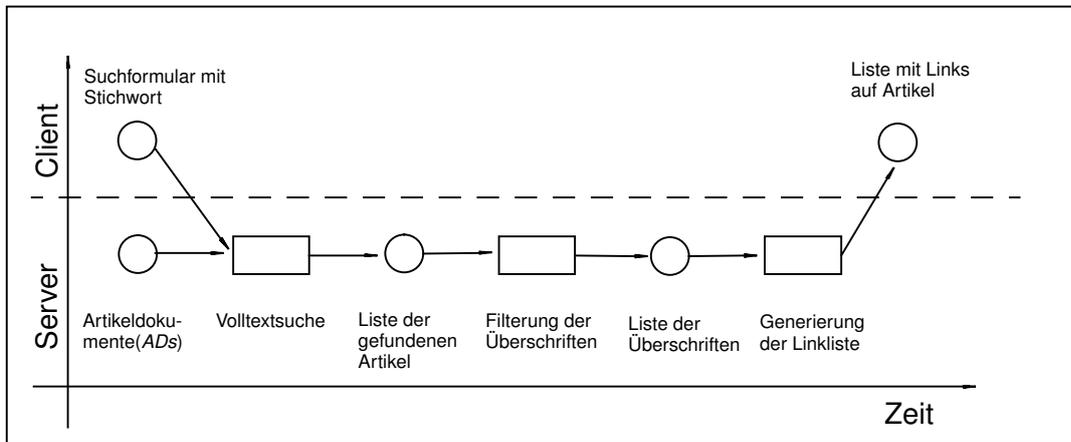


Abb. 18: Ablauf eines Suchprozesses

Das hierzu nötige Durchsuchen der Artikeldokumente (ADs) geschieht durch Hilfsprogramme, die standardmäßig unter UNIX verfügbar sind.

Der Ablauf des Suchprozesses wurde in Abb. 18 dargestellt. Als Resultat der Suche wird eine Liste von Titelzeilen mit *links* auf jene Dokumente zurückgeliefert, in denen der gesuchte Terminus aufgetreten ist. Diese Liste wird durch Filterung der Überschriften aus den Artikeldokumenten generiert.

4.2.6 Statistische Auswertung der Logdateien

Die Online-Ausgabe der *tageszeitung* wird durch einen Server verwaltet, der neben diesem Angebot weitere Informationsdienste zur Verfügung stellt. Aus den Server-Logdateien müssen deshalb zunächst jene Einträge herausgefiltert werden, die Zugriffe auf Dokumente der *tageszeitung*

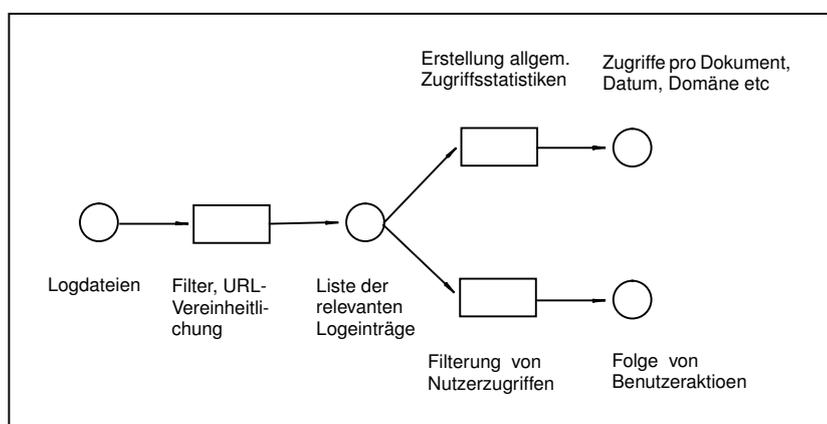


Abb. 19: Auswertung der Logdateien

dokumentieren. Aus den gefilterten Einträgen können danach mit Auswertungsprogrammen (wwwstat¹, analog²) Statistiken generiert werden, die Aussagen z.B. zur Zugriffshäufigkeit nach Tageszeit, Tag, Woche oder Monat, zu Browsertypen oder zu den am häufigsten angeforderten Dokumenten erlauben.

Die genannten Statistikwerkzeuge produzieren z.T. graphische Ausgaben und sind damit gut geeignet, um sich einen generellen Überblick zu verschaffen. Für das Nachvollziehen einzelner Folgen von Leseraktionen sind sie allerdings nicht ausreichend. Hierzu müssen direkt aufeinanderfolgende Zugriffe einzelner Clients ausgefiltert werden (vgl. Abb. 19). Dies ist z.B. mit UNIX-Filterprogrammen möglich.

1. [Ics 94]

2. [Tur96]

5.0 Systemimplementation

Zur Abbildung der in Kap. 4.1.1 dargestellten Dokumentenstruktur wurde ein Datei- und Verzeichnisbaum angelegt. Ausgehend vom Wurzelverzeichnis `~taz` wurden folgenden Unterverzeichnisse eingerichtet:

| | |
|---|---|
| <code>~/taz/.public_html/heute</code> | aktuelle Tagesausgabe |
| <code>~/taz/.public_html/gestern</code> | Ausgabe vom Vortag |
| <code>~/taz/.public_html/LeMonde</code> | monatliche Beilage <i>Le Monde Diplomatique</i> |
| <code>~/taz/.public_html/Mail</code> | Leserbriefforum |
| <code>~/taz/.public_html/archiv</code> | alte Ausgaben, nicht öffentlich zugänglich |
| <code>~/taz/.public_html/hilfe</code> | Hilfeseiten |
| <code>~/taz/.public_html/gifs</code> | Layouthilfen (Grafiken und Logos) |

Für die den Diensten zugeordneten Programme ein eigenes Verzeichnis kreiert:

```
/www/www/cgi-bin/taz
```

Die Konfiguration des *httpd* Servers am PRZ wurde so modifiziert, daß alle in diesem Verzeichnis vorhandenen Programme vom CGI-Gateway des Servers aufgerufen werden können.

Eine Liste aller erstellten bzw. modifizierten Softwarekomponenten des Systems befindet sich in Anhang B.

5.1 Dateitransfer und Spiegelung

Das Zusammenfassen und Übertragen der Archivdateien erfolgt bei der *tageszeitung* durch ein UNIX Shell-Skript, das nach Redaktionsschluß automatisch gestartet wird. Zum Packen der Transferdatei (vgl. Kap. 4.2.1) werden die Standardverfahren *tar* und *gzip* eingesetzt. Die so komprimierte Datei wird mittels eines *remote copy* Aufrufs (*rcp*) zum PRZ zu kopieren versucht. Ob der Kopiervorgang erfolgreich war, wird über den Return-Wert des *rcp*-Kommandos bestimmt. Schlägt das Kopieren fehl, so versetzt sich das Skript in den *sleep*-Zustand und versucht nach 30 Minuten erneut, die Archivdaten zu kopieren. Dieser Prozeß wird solange wiederholt, bis die Übertragung erfolgreich durchgeführt werden konnte. Die Mechanismen, die zur Benachrichtigung der Administratoren im Fehlerfall eingesetzt werden, sind in Kap. 4.1.2 erläutert worden. Die entsprechenden *email*-Nachrichten werden vom Skript automatisch generiert. Nach erfolgreichem Kopieren wird die Dateikonvertierung durch einen *rsh*-Aufruf des am PRZ vorhandenen Shell-Skripts `tazmake` angestoßen.

Am PRZ sind zum Zeitpunkt des Konvertierungsvorgangs die zwei Tagesausgaben vorhanden: die bis dahin aktuelle im Unterverzeichnis

```
~/taz/www/heute
```

und diejenige des Vortages im Unterverzeichnis

```
~/taz/www/gestern.
```

Nach erfolgreicher Konvertierung der Transferdatei in eine neue Ausgabe wird das Verzeichnis

`~taz/www/gestern` gelöscht. Das Verzeichnis `~/taz/www/heute` wird umbenannt in `~taz/www/gestern`. Schließlich wird die zuvor konvertierte tagesaktuelle Ausgabe im neu kreierte Verzeichnis `~/taz/www/heute` abgelegt. Die Verwendung eines gleichbleibenden Verzeichnisnamens (`~/taz/www/heute`) hat für Online-Leser den Vorteil, das sie einen Verweis auf dieses regelmäßig aktualisierte Verzeichnis in die *Hotlist* ihres Browsers aufnehmen können.

Für die optionale Spiegelung der Ausgabe werden die konvertierten Dateien in `~taz/www/heute` mittels *tar*, *gzip* und *uuencode* zu einer neuen Transferdatei zusammengefaßt und durch das Programm PGP kryptographisch signiert. Danach wird diese Transferdatei per *email* zu einem oder mehreren Spiegelservern übertragen.

Die Dekompression und Installation dieser Dateien erfolgt auf den Spiegelservern durch einen automatischen *email-processor* (*gnu-procmail*). Hierbei werden zunächst der Absender der Nachricht und die kryptographische Signatur überprüft. Können diese verifiziert werden, so erfolgt die Installation der Ausgabe auf die bereits für das PRZ beschriebene Art und Weise (vgl. Kap. 4.2.1). Ergibt die Überprüfung einen Fehler, wird der WWW-Administrator des PRZ per *email* benachrichtigt. Da lediglich die prinzipielle Machbarkeit der Spiegelung auf weiteren Servern demonstriert werden sollte, wurde auf weitergehende Mechanismen (z.B. mehrfache Übertragungsversuche im Fehlerfall) verzichtet.

5.2 Formatkonvertierung

Für die Formatkonvertierung war der mit Abstand größten Implementationsaufwand aller Komponenten erforderlich. Da keine konsistente Beschreibung der Grammatik der *taz*-Satzformats vorlag, wurde auf den Einsatz von Parsergeneratoren wie *lex* und *yacc* verzichtet. Bei Wahl der Implementationssprache waren die Erfahrungen des Verfassers ausschlaggebendes Kriterium. Der HTML-Konverter wurde deshalb in der Sprache C implementiert, der Strukturgenerator mit Hilfe von Bourne-Shell-Skripten realisiert

5.2.1 HTML-Konverter

Der HTML-Konverter setzt sich aus den zwei Hauptmodulen Parser (*parser.c*) und Generator (*generator.c*) zusammen. Der Parser lokalisiert struktur- und layoutbezogene Makros im Satzformat der *tageszeitung*. Die vom Parser aufgerufenen Generatorfunktionen ersetzen diese Makros durch geeignete HTML-Direktiven. Diese werden zusammen mit dem Fließtext in Artikeldateien (*ADs*) und Seitenverzeichnisdateien (*SVDs*) abgelegt, die das Zwischenformat für den Strukturgenerator bilden (vgl. Kap. 4.2.2). Parser und Generator umfassen zusammen etwa 2000 Zeilen C-Code.

Parser

Für einen ersten Überblick über die Struktur des Ausgangsformats wurden zwei Satzdateien (provisorische Archivdateien) untersucht. Hierdurch konnten zunächst 15 Hauptmakros ermittelt werden, von denen eines eine verschachtelte Struktur aufwies. Nun wurden durch Textfilterung alle

Makros des verschachtelten Typs isoliert und auf das Skelett der in ihm enthaltenen Submakros und der Klammerstruktur reduziert. Hieraus ergab sich eine – zunächst unvollständige – Liste zulässiger Makroverschachtelungen. Auf der Grundlage dieser Liste wurde der erste Prototyp des Satzformatparsers entwickelt. Dieser wurde mit zehn weiteren Satzdateien getestet und dabei um bis dahin nicht aufgetretene Makros erweitert. Der so überarbeitete Parser war während eines mehrmonatigen Tests in der Lage, alle von der *tageszeitung* gelieferten Satzdateien korrekt zu verarbeiten. Die aus dem funktionsfähigen Parser abgeleitete Grammatik ist in Anhang A aufgeführt.

Hinsichtlich der Syntax des Satzformats ist für die folgenden Erläuterungen nur von Belang, daß jedes Makro mit einer öffnenden geschweiften Klammer und einer direkt danach folgenden Makrobezeichnung beginnt und durch eine schließende geschweifte Klammer beendet wird. Zwischen den Klammern befindet sich Fließtext, ggf. mit weiteren eingestreuten Submakros.

Die Verschachtelungstiefe von Makros war zu Beginn der prototypischen Implementation nicht bekannt. Aus diesem Grunde wurde der Parser nicht als endlicher Automat, sondern rekursiv implementiert. Ausgehend von einem Wurzelement (Artikel) wurde dazu jedem Makro eine Funktion zugeordnet. Diese enthält eine Fallunterscheidung für alle innerhalb des Makros zulässigen Submakros. Wird beim Abarbeiten eines Makros ein Element gefunden, das nicht in der Liste zulässiger Submakros enthalten ist, bricht der Parser mit einer Fehlermeldung ab, die die Zeilennummer der Satzdatei und den Namen der Funktion enthält, in der der Fehler aufgetreten ist.

Die Struktur aller Parserfunktionen, denen Makros zugeordnet sind, ist weitgehend identisch. Die Funktionen enthalten ggf. einen vorbereitenden `pre-`, `set-` oder `check-` Aufruf des HTML-Generators, um diesem das Auftreten eines neuen Elements zu signalisieren oder das Setzen von Parametern layoutbezogener Variablen auszulösen (derartige Aufrufe erfolgen z.B. zu Beginn eines neuen Artikels, einer Überschriftenzeile oder einer Ressortangabe). Danach werden ggf. enthaltene Submakros geparkt. Falls das Makro einem HTML-Formatierungselement entspricht, wird eine Generatorfunktion aufgerufen. Vor Verlassen einer Parserfunktion werden ggf. `post-` oder `unset-` Funktionen aufgerufen, z.B. um Variablen im Generator zurückzusetzen.

| | Fließtext als HTML-Text darstellen | Fließtext nicht als HTML-Text darstellen |
|------------------------|---|--|
| komplexes Makro | Nach öffnenden und schließenden Klammern suchen, Text ausgeben (<code>DisplayPT</code>) | Nach öffnenden und schließenden Klammern suchen, Text nicht ausgeben (<code>NDisplayPT</code>) |
| einfaches Makro | Nur nach schließenden Klammern suchen, Text ausgeben (<code>DisplayT</code>) | Nur nach schließenden Klammern suchen, Text nicht ausgeben (<code>NDisplayT</code>) |

TABELLE 1. Display-Funktionen

Das Lokalisieren von Submakros und Makro-Endpunkten (d.h., von schließenden geschweiften Klammern) wird durch eine `Display`-Funktion gesteuert, die den innerhalb von Makros enthaltenen Fließtext abarbeitet. Ob und wie dieser Text verarbeitet wird, hängt von des Art des ihn umschließenden Makros ab.

Oftmals enthält der Fließtext Metainformationen, die nicht auf den produzierten HTML-Seiten abgebildet und deshalb nicht vom Generator ausgegeben werden sollen. Makros können sowohl einfach als auch komplex sein, wobei komplexe Makros öffnende geschweifte Klammern enthalten können. Die vier zugeordneten `Display`-Modi wurden in Tab. 1 dargestellt, wobei die angegebenen Bezeichnungen der Modi den Namen der Definitionen im C-Quellcode entsprechen.

Der Fließtext muß zusätzlich auf enthaltene Sonderzeichen geprüft und z.T. parallel auf mehrere Ausgabedateien verteilt werden. Diese geschieht in der weiter unten beschriebenen `inText()` Funktion des HTML-Generators.

HTML-Generator

Der HTML-Generator stellt dem Parser als Schnittstelle eine Anzahl von Funktionsaufrufen zur Verfügung. Die Semantik einer Schnittstellenfunktion wird im C-Quellcode durch einen Präfix ihres Funktionsnamens signalisiert. Folgende Gruppen können unterschieden werden:

- **check-**: Diese Funktionen überprüfen und setzen Parameter, die Metainformationen enthalten. Einige dieser Informationen (Autor, Datum, Ressort, Seite, Ausgabennummer) werden in den generierten HTML-Dokumenten dargestellt, andere (die Nummer der Ausgabe und der Artikel) zur Generierung von Dateinamen verwendet.
- **set-/unset-**: Diesen Funktionen sind zuständig für das Einfügen von HTML-Formatanweisungen, für die Auswahl von Schrifttypen (`setFormat()`), für Absatzumbrüche (`setParagraph()`) und Überschriften (`setTitle()`). Die Funktionen `setTitle()` und `unsetTitle()` kontrollieren zusätzlich den Ausgabestrom für Seitenverzeichnisdateien (SVDs).
- **pre-/post-** bzw. **init-/exit-**: In diesen Funktionen werden Initialisierungen vorgenommen oder HTML-Formatanweisungen eingefügt.

Eine Sonderrolle spielt die bereits erwähnte Funktion `inText()`. In ihr wird der Fließtext zeichenweise auf Umlaute und Ligatur geprüft und ggf. in die entsprechenden HTML-Korrelate umgesetzt.

Optische Elemente für die graphische Gestaltung, die Navigation und das Ansprechen der Dienste (Suche, Forum und Leserbriefe, vgl. Kap. 4.1.2) werden vom HTML-Generator eingefügt. Das miniaturisierte *taz*-Logo wird dabei in der Funktion `newPage()` eingebunden. Die Verweise auf Leserbrief, Forum und Volltextsuche werden durch die in der Funktion `addGimmicks()` zusammengefaßten Aufrufe generiert, die Navigationselemente von der Funktion `addPreInfo()`. Alle mit Navigationselementen verbundenen *links* beziehen sich generisch auf Dokumente in der artikel- und rubrikorientierten Struktur (vgl. Kap. 4.1.1). Für die seitenorientierte Struktur und die artikelorientierte Struktur *mit Gesamtverzeichnis* (ebenfalls Kap. 4.1.1) müssen diese *links* auf die entsprechenden informationellen Einheiten dieser Strukturen umgesetzt werden. Die zu modifizierenden Dokumentenpassagen sind durch vorangestellte HTML-Kommentare gekennzeichnet. Die nötigen textuellen Modifikationen werden durch den Strukturgenerator vorgenommen, dessen Implementation im nächsten Kapitel erläutert ist.

Der Konvertierungsvorgang erzeugt aus der Satzdatei für jeden enthaltenen Artikel eine HTML-formatierte *Artikeldatei* (*ADT*). Die erste Zeile jeder Artikeldatei bildet ein Anker der Form

```
<Jahr><Monat><Tag> . <ARTIKEL-ID>
```

Dies erlaubt eine Referenzierung des Artikels innerhalb von Dokumenten der seitenorientierten Ausgabe (Seitendokumente (*SDs*) werden aus mehreren *ADTs* zusammengefügt, vgl. Kap. 4.2.2). Die Definition der Namen von Artikeldateien wurden so gewählt, daß sie eine einfache Ermittlung von Metainformationen des jeweiligen Artikels zuläßt. Ein solcher Dateiname hat das Format

```
<Seite>-<Ausgabe>@@<Jahr><Monat><Tag> . <ARTIKEL-ID>~~ .html
```

Hierbei ist *<Seite>* eine zweistellige Dezimalzahl zwischen 01 und 99, *<Ausgabe>* eine vierstellige Dezimalzahl zwischen 0001 und 9999, *<Monat>* eine zweistellige Dezimalzahl zwischen 01 und 12, *<Tag>* eine zweistellige Dezimalzahl zwischen 01 und 31 und *<ARTIKEL-ID>* eine dreistellige Dezimalzahl zwischen 001 und 999.

Beispiel: der erste Artikel der Ausgabe 4747 vom 14.10.1995 auf Seite 1 trägt den Dateinamen

```
01-4747@@951014.001~~.html
```

Die Namen der Seitenverzeichnisdateien (*SVDs*, vgl. Kap. 4.2.2) haben die Form

```
<Seite>-<Ausgabe>.dx.
```

Für jede Seite werden zwischen horizontale Linien eingeschlossene, HTML-formatierte Datums- und Seiteninformationen in eine seitenbezogene Hilfsdatei (*SHD*) geschrieben. Der zugehörige Dateiname hat die Form:

```
<Seite>-<Ausgabe>.br
```

Eine weiter seitenbezogene Hilfsdatei enthält die Rubrik, der eine Seite zugeordnet ist. Der Name dieser Datei hat die Form

```
<Seite>-<Ausgabe>.ct
```

5.2.2 Strukturgenerator

Der Strukturgenerator erzeugt aus den vom HTML-Konverter ausgegebenen Zwischenformaten die WWW-Dokumentenstruktur des Online-Angebots. Zur Generierung der Dokumentenstruktur sind in erster Linie textuelle Modifikationen der vom Konverter erzeugten Zwischendateien und Operationen wie Dateizusammenführung und Manipulation von Dateinamen nötig. Da für derartige Operationen unter dem verwendeten Betriebssystem UNIX Standardhilfsmittel zur Verfügung stehen, wurde der Strukturgenerator als Bourne-Shell-Skript implementiert. Dieses Skript namens *tazbuild* hat eine Länge von ca. 400 Zeilen. Es wird aus dem übergeordneten Skript *tazmake* heraus aufgerufen.

Die Generierung der Seitenstruktur umfaßt folgende Funktionen (vgl. Kap. 4.1.1):

1. Bestimmung von Parametern der aktuellen Online-Ausgabe.
2. Erzeugung von Dokumententeilen, die Bestandteil jedes Gesamtverzeichnisses (*GV*) und des rubrikbezogenen Verzeichnisses (*SV*) der ersten Dokumentenseite sind.
3. Erzeugung der seitenorientierten Struktur, d.h., der Seitendokumente (*SDs*) und des seitenbezogenen Gesamtverzeichnisses (*SGV*)
4. Erzeugung der artikelorientierten Struktur, d.h., der Artikeldokumente (*ADs*) und des artikelbezogenen Gesamtverzeichnisses (*AGV*).
5. Erzeugung der rubrikorientierten Struktur, d.h., der Rubrik/Seitenverzeichnisse (*SVs*)

Vor der Erläuterung dieser Funktionen werden für die einzelnen Seitentypen die Formen der jeweiligen Dateinamen angegeben¹:

Artikeldateien/-dokumente (*ADs*):

```
<Seite>-<Ausgabe>@@T<Jahr><Monat><Tag>.<ARTIKEL-ID>~~.html
```

Seitendokumente (*SDs*):

```
<Seite>-<Ausgabe>.html
```

Artikelorientiertes Gesamtverzeichnis (*AGV*):

```
index.articles.html
```

Seitenbezogenes Gesamtverzeichnis (*SGV*):

```
index.page.html
```

Rubrik-/Seitenbezogenes Verzeichnisse (*SVs*):

```
index.<Seite>-<Ausgabe>.html
```

Bestimmung von Parametern der aktuellen Online-Ausgabe

Da die Seitenanzahl einer Ausgabe von Tag zu Tag differieren kann, muß diese zunächst vom Shellskript ermittelt werden. Hierzu wird mittels der Funktion `buildLinkFile()` die Datei `.taz-pages` erzeugt, die eine Liste der Namen aller Seitenverzeichnisdateien (*SVDs*, Dateierweiterung `*.dx`) enthält. Jeder Eintrag enthält - falls vorhanden - auch die Nummer der jeweils vorhergehenden und nächsten Seitenverzeichnisdatei. Aus dieser Liste läßt sich die Seitenanzahl direkt ablesen.

Erzeugung von Dokumententeilen für Verzeichnisse

Dokumententeile für Verzeichnisse enthalten das Ablaufdatum (HTML - EXPIRE) des Dokuments, das *taz*-Logo, Verweise auf aktuelle Informationen, die nicht Bestandteil der gedruckten Ausgabe sind, und *links* auf Abbildungen (Karikaturen/Photos).

Diese Elemente bilden den Vorspann des seitenorientierten Gesamtverzeichnisses (*SGV*), des artikelorientierten Gesamtverzeichnisses (*AGV*) und des Seitenverzeichnisses (*SV*) der ersten Seite.

1. Zum Format der in spitze Klammern eingeschlossenen Elemente vgl. vorhergehende Seite

Erzeugung der seitenorientierten Struktur

Zunächst werden die einzelnen Seitendokumente (*SDs*) generiert. Jedes Seitendokument besitzt einen Vorspann, der das *taz*-Logo, einen Copyrightvermerk, Navigationselemente und Rubrikbezeichnung und das lokale Inhaltsverzeichnis der auf der Seite vorhandenen Artikel enthält. Diesem Vorspann folgen die eigentlichen Artikel. Hierbei werden auch die vom Parser erzeugten seitenbezogenen Hilfsdateien (*SHDs*, Dateiendungen **.br* und **.ct*) der jeweilige Seite integriert.

Das lokale Inhaltsverzeichnis wird aus der zugehörigen Seitenverzeichnisdatei (*SVD*) erzeugt. Die Navigationselemente [**<**] und [**>**] innerhalb eines lokalen Inhaltsverzeichnisses verweisen auf das vorangehende bzw. folgende Seitendokument (*SD*), das Element [**^**] auf das seitenbezogenen Gesamtverzeichnis (*SGV*)

Jede zu einer bestimmten Seite gehörende Artikeldatei (*ADT*) wird durch die Funktion `link_article()` um die Navigationselemente [**<**] und [**>**] ergänzt. Die diesen Elementen zugeordneten *links* verweisen auf den jeweils auf der Seite vorausgehenden und nächsten Artikel. Das Element [**^**] verweist in Artikeln von Seitendokumenten immer auf den zugeordneten Eintrag im lokalen Inhaltsverzeichnis dieses Dokuments (*SD*). Durch diese Erweiterungen wird die Artikeldatei (*ADT*) zu einem vollständigen Artikeldokument (*AD*). Das Seitendokument entsteht durch abschließendes Zusammenfügen des Vorspanns, des Inhaltsverzeichnisses und der Artikeldokumente.

Generisch verweisen die im lokalen Gesamtverzeichnis vorhandenen und die mit den Navigationselementen der einzelnen Artikel verbundenen *links* auf *Artikeldokumente* (*ADs*, vgl. Kap. 5.2.1). Diese müssen in lokale Verweise, d.h., in Referenzen auf die im jeweiligen Seitendokument vorhandenen Artikel, geändert werden. Konkret bedeutet dies, daß alle Referenzen der Form

```
<XX>-<Ausgabe>@@<Jahr><Monat><Tag>.<ARTIKEL-ID>~~.html
```

zu Referenzen der Form

```
#<ARTIKEL_ID>
```

verkürzt werden, wobei *<XX>* die Seitennummer des jeweils bearbeiteten Dokuments angibt.

Das seitenbezogene Gesamtverzeichnis (*SGV*) besteht aus dem oben beschriebenen Vorspann und den aneinandergefügt Seitenverzeichnisdateien (*SVDs*). Die Inhaltsangaben verschiedener Seiten sind graphisch durch die in den zugehörigen **.br* und **.ct* Dateien enthaltenen Layoutelemente voneinander abgesetzt. Wie bei Generierung der Seitendokumente müssen auch im seitenbezogenen Gesamtverzeichnis die *links* auf Artikeldokumente in solche auf Artikel in Seitendokumenten umgewandelt werden. Konkret bedeutet dies, daß alle Verweise der Form

```
<Seite>-<Ausgabe>@@<Jahr><Monat><Tag>.<ARTIKEL-ID>~~.html
```

in solche der Form

```
<Seite>-<Ausgabe>#<ARTIKEL-ID>.html
```

geändert werden. Die Zielanker der Navigationselemente im seitenorientierten Gesamtverzeichnis und in Seitendokumenten lassen sich aus Spalte 1 bzw. 4 von Tab. 2 ablesen.

| | Seitenorientiertes Gesamtverzeichnis (SGV) | Artikelorientiertes Gesamtverzeichnis (AGV) | Rubrik-/Seitenverzeichnis (SV) | Seitendokument (SD) | Artikeldokument (AD) |
|----------|---|---|--------------------------------|--|---|
| [<] | Verzeichnis der vorhergehenden Seite innerhalb des SGVs | Verzeichnis der vorhergehenden Seite innerhalb des AGVs | SV für vorhergehende Seite | vorhergehender Artikel innerhalb des SDs | AD mit auf der Seite vorhergehendem Artikel |
| [>] | Verzeichnis der nächsten Seite innerhalb des SGVs | Verzeichnis der nächsten Seite innerhalb des AGVs | SV für nächste Seite | folgender Artikel innerhalb des SDs | AD mit auf der Seite folgendem Artikel |
| [^] | Tages-Leitseite | Tages-Leitseite | AGV | SD-lokales Inhaltsverzeichnis | SV |
| [Browse] | SV der Seite | n.v. | n.v. | SD | SV |
| [Paged] | n.v. | SGV | SGV | n.v. | n.v. |

TABELLE 2. Zielanker von Navigationselementen in allen Dokumententypen

Erzeugung der artikelorientierten Struktur

Die durch den HTML-Generator erzeugten Artikeldateien (*ADTs*) wurden bereits bei der Generierung der seitenorientierten Struktur zu Artikeldokumenten (*ADs*) ergänzt. Um das artikelbezogene Gesamtverzeichnis (*AGV*, vgl Kap. 4.2.2) zu erzeugen, wird analog zum seitenbezogenen Gesamtverzeichnis (*SGV*) verfahren. Die bei der Generierung des seitenbezogenen Inhaltsverzeichnisses (*SGV*) notwendigen Änderungen von *link*-Referenzen entfallen. Die Zielanker der im artikelorientierten Gesamtverzeichnis und in Artikeldokumenten vorhandenen Navigationselemente lassen sich aus Tab. 2 Spalte 2 bzw. 5 ablesen.

Erzeugung der rubrikorientierten Struktur

Zur Erzeugung dieser Struktur werden die Seitenverzeichnisdateien (*SVDs*) um einen Vorspann und Navigationselemente zu Seitenverzeichnissen (*SDs*) erweitert. Dies geschieht analog zur Erzeugung von Gesamtinhaltsverzeichnissen, wobei die einzelnen Rubrik/Seitenverzeichnisse natürlich nicht in einer gemeinsamen Datei zusammengefaßt werden. Die Zielanker der in den Rubrik-/Seitenverzeichnissen vorhandenen Navigationselemente lassen sich aus Tab. 2 Spalte 3 ablesen.

5.3 Authentifizierung

Die (optionalen) Authentifizierungsmerkmale der Dokumente werden in HTML-Kommentare eingeschlossen. Sie sind also beim Aufruf des Dokuments für den Benutzer nicht sichtbar.

Die Erweiterung der Dokumente um Authentifizierungsmerkmale erfolgt durch das Bourne-Shell Skript `tazauth.sh`. Dieses Skript liest zunächst die zur Entsperrung des *private keys* der *tageszeitung* nötige Paßwort aus der Datei `tazpgppass` ein. Der Signierungsvorgang eines Dokuments geschieht in folgenden Schritten:

1. Es wird eine Arbeitskopie des zu signierenden Verzeichnis-, Seiten- oder Artikeldokuments angefertigt.

2. Falls die erste Zeile des Dokuments die Direktive `<HTML>` enthält, wird diese entfernt. Gleiches gilt für die Direktive `</HTML>`, sofern sie die letzte Zeile des Dokuments bildet.
3. Dem Dokument wird eine *schließende* HTML-Kommentardirektive (`<!-->`) vorangestellt. Eine *öffnende* HTML-Kommentardirektive (`<!--`) wird als Schlußzeile des Dokuments angefügt.
4. Das Dokument wird mit dem Programm *PGP* signiert. Dieses Programm stellt dem Dokument die Zeile


```
-----BEGIN PGP SIGNED MESSAGE-----
```

 voran. Die erstellte Signatur wird von *PGP* an das Ende des Dokuments angefügt. Sie befindet sich zwischen den Zeilen


```
-----BEGIN PGP SIGNATURE-----
```

 und


```
-----END PGP SIGNATURE-----
```
5. An den Anfang des nunmehr signierten Dokuments wird eine *öffnende* HTML-Kommentardirektive und an das Ende des Dokuments eine *schließende* HTML-Kommentardirektive gesetzt. Alle von PGP erzeugten Zeilen befinden sich nunmehr innerhalb von HTML-Kommentarzeilen.
6. Das Originaldokument durch sein signiertes Äquivalent ersetzt.

```

<HTML>
<!--
-----BEGIN PGP SIGNED MESSAGE-----
-->
(...Dokument...)
<!--
-----BEGIN PGP SIGNATURE-----
(...PGP - Signatur...)
-----END PGP SIGNATURE-----
-->

```

Abb. 20: Struktur eines signierten Dokuments

Die Form eines signierten Dokuments wird von Abb. 20 wiedergegeben. Die unsichtbar übertragene Signatur kann bei Bedarf durch den Benutzer kontrolliert werden. Dieser muß hierzu das Dokument lokal abspeichern und danach mit *PGP* überprüfen. Der dazu notwendige *public key* der *tageszeitung* kann vom Benutzer über WWW abgefragt oder auf Anfrage per *email* übermittelt werden.

Das Signieren aller Dokumente erfolgt, *bevor* das aktualisierte Angebot öffentlich zur Verfügung gestellt wird. Dies bringt im Vergleich zur unsignierten Ausgabe Verzögerungszeiten von etwa einer Stunde mit sich (vgl. Kap. 4.2.3).

5.4 Leserbriefe und Diskussionsforum

Leserbriefe und Diskussionsforum wurden auf der Grundlage der Programmpakets *Hypermail* von Kevin Hughes und Tom Gruber implementiert. *Hypermail* besteht aus den Komponenten *mail*, das ein einfaches *email*-Formular generiert, und *hypermail*, das einen *mailfolder*, d.h., eine *email*-Archivdatei, in einzelne, HTML-formatierte Dokumente zerlegt. Parallel dazu werden themen-, autoren- und ein datumsorientierte Indices aller *emails* erstellt (vgl Abb. 21). Das Programmpaket *Hypermail* ist in der Programmiersprache C implementiert.

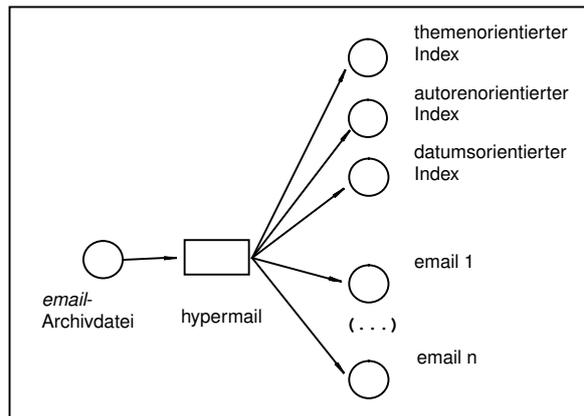


Abb. 21: Zerlegung und Indizierung einer email-Archivdatei durch *hypermail*

Für die Integration von *Hypermail* in das Online-Angebot der *tageszeitung* wurden verschiedene Erweiterungen vorgenommen. Zunächst wurde das auf S. 39 vorgestellte Formular entwickelt. Aus diesem Formular wurde ein `char*`-Array generiert, das im *Hypermail*-Modul `mail.c` von der Funktion `printforn()` anstelle des in `mail.c` vorgegebenen Standardformulars ausgegeben wird. Im Unterschied zum Standardformular hat die den Empfänger bezeichnende Variable `to` im neu entwickelten Formular nicht den Typ `text`, sondern den Typ `hidden`¹. Dies verhindert, daß der Benutzer die Empfängeradresse selbstständig manipuliert. Über den `SEND-button` des Formulars kann die *Preview*-Funktion aufgerufen werden. Zwei `buttons` namens `tazmail` und `noforum` dienen der Zuordnung der Empfängeradresse (private Nachricht/taz-Leserbriefredaktion, vgl. Kap. 4.2.4).

Um die verschiedenen Sendeoptionen des Leserbriefformulars korrekt behandeln zu können, war eine weitere Modifikation der *mail*-Komponente nötig. Eine Nachricht wird nur dann versandt, wenn die mit der *Preview*-Funktion verbundene Variable `SEND` den Wert `ON` hat, d.h., wenn der entsprechende `button` aktiviert ist (s.o., vgl. Kap. 4.2.4). Nachrichten werden immer dann an die Adresse des *taz*-Forums geschickt, wenn der Benutzer dies nicht explizit ausschließt. Dies wird vom *mail*-Programm durch die Überprüfung der Variablen `noforum` festgestellt. Hat die Variable `tazmail` den Wert `ON`, wird die Nachricht zusätzlich an die *email*-Adresse der Leserbriefredaktion der *tageszeitung* weitergeleitet. Die notwendigen Fallunterscheidungen wurden in die Funktion `cgi-main()` des Moduls `mail.c` eingefügt. Um eventuelle Verwechslungen mit dem

1. [Gra95], 102 ff.

UNIX-Standardprogramm *mail* zu vermeiden, wurde der Name der Komponente in *hypermailer* geändert.

Die Modifikationen der für das Konvertieren der *email*-Archivdatei nötigen *hypermail*-Komponente betrafen ausschließlich das Layout der generierten Dokumente und Indices. In den Funktionen `writearticles()`, `writedates()`, `writethreads()`, `writedesubjects()` und `writeauthor()` des Moduls `print.c` wird nun jeweils eine Zeile mehr generiert, die eine Referenz auf ein miniaturisiertes *taz-logos* enthält. In der Funktion `writearticles()` wurde die Darstellung der *Subject*-Zeile geändert und der Absenderangabe ein graphisches Symbol vorangestellt, über das ein Formular für Folgeartikel angefordert werden kann.

Das Modul *hypermailer* integriert alle für die Generierung von Leserbriefen, die Auswertung von Formularen und die Zuordnung von Empfängeradressen notwendigen Funktionen (vgl. Abb. 17 auf S. 45). Die Aktivierung dieser Komponente erfolgt sowohl beim Anfordern eines neuen wie auch beim Abschicken eines ausgefüllten Leserbriefformulars. Die Komponente *hypermail* (vgl. Abb. 21) wird indirekt aktiviert, und zwar durch das mit dem **[Forum]** - Element der Navigationsleiste (vgl. Abb. 11 auf S. 38) verbundene Skript `readmail`. Dieses Skript vergleicht das Alter der *email*-Archivdatei mit demjenigen der generierten, HTML-formatierten Indexdateien. Ist die Archivdatei jüngerer Datums, so wird das Programm *hypermail* gestartet. Dies bewirkt eine Konvertierung der neu eingetroffenen Nachrichten und Aktualisierung der Indexdateien. An den Benutzer wird standardmäßig der (ggf. aktualisierte) themenorientierte Index übertragen.

5.5 Stichwortsuche

Die Suchfunktion wird über das Bourne-Shell-Skript `taz-suche` realisiert. Bei einem Aufruf ohne Parameter wird von diesem Skript ein leeres Suchformular generiert. Sind Parameter vorhanden, wird eine Freitextsuche durchgeführt und das Ergebnis an den Benutzer übermittelt.

In Abstimmung mit Mitarbeitern der *tageszeitung* wurde nur eine vergleichsweise primitive Freitextsuche implementiert. Als Eingabe sind weder formale Deskriptoren noch reguläre Ausdrücke zugelassen. Eine Syntaxbeschreibung des Anfrageausdrucks erübrigt sich aus diesem Grunde. Erlaubt sind nur einfache Zeichenfolgen, nach denen alle Dokumente der aktuellen Ausgabe durchsucht werden.

Optional kann der Benutzer die Suche auf ganze Wörter einschränken. Standardmäßig wird nach der eingegebenen Zeichenfolge auch als Teilfolge von Worten gesucht. Die Semantik einer Anfrage lautet umgangssprachlich: *„Konstruiere eine HTML-formatierte Liste mit den Artikelüberschriften aller Artikeldokumente, die die eingegebene Zeichenfolge als Teilbegriff / als ganzes Wort enthält, und verbinde jedes Listenelement mit einem link auf das entsprechende Artikeldokument.“*.

Für Umlaute und Ligaturen in Suchbegriffen können Ersatzdarstellungen verwendet werden. Hierbei wird den entsprechenden Buchstaben das Fluchtsymbol ``` vorangestellt. Eine direkte Eingabe von Umlaut- und Ligaturzeichen ist ebenfalls zulässig. Da deren Transfersyntax von clientseitig verwendeten Browsertypen und Betriebssystemeinstellungen abhängt, muß die Suchanfrage

zunächst in eine Normalform überführt werden. Die Konvertierung in die vom HTML-Konverter erzeugte Standardform erfolgt durch einen *sed*-Filter.

Die Volltextsuche geschieht durch Filterung aller Artikeldokumente mit dem UNIX-Kommando *grep*. Da das Satzformat der *tageszeitung* keine Silbentrennungen enthält, ist eine Freitextsuche mit diesem Werkzeug recht zuverlässig. Eine wesentliche Einschränkung ist durch die zeilenorientierte Suche von *grep* gegeben: die gesuchte Zeichenfolge wird nur dann gefunden, wenn sie sich in ein und derselben Textzeile des Artikeldokuments befindet.

Vor dem Durchsuchen mittels *grep* werden die Artikeldokumente zunächst durch einen *sed* - Filter geschickt, der sämtliche in ihnen enthaltenen HTML-Direktiven entfernt. Die *grep* Option *-w* wird entsprechend der im Formular gewählten Suchoption "Nur ganzes Wort suchen" gesetzt. Enthält ein Artikeldokument die gesuchte Zeichenfolge, wird die zwischen den HTML-Direktiven *<H2>* und *</H2>* stehende Artikelüberschrift mittels eines *sed*-Filters extrahiert, mit einem *link* auf den entsprechenden Artikel versehen und in die Ergebnisliste übertragen. Das Ergebnisdokument besteht aus dem Suchformular und der Liste aller Artikel, in der die gesuchte Zeichenfolge gefunden wurde,.

6.0 Prototyping und Test

Grundansatz der Systementwicklung war, das Online-Angebot in einem mehrstufigen Prototyping-Prozeß zu optimieren, in dem die angebotenen Dokumentenstrukturen, Navigationsmechanismen und Dienste Benutzerbedürfnissen angepaßt werden sollten. Dabei wurde versucht, durch konsequente Berücksichtigung von Leserkomentaren einen akzeptablen Mittelweg zwischen den Extremen “mangelhaftes Design” und “Overengineering” zu finden.

Rückblickend betrachtet war diese Strategie erfolgreich. Die Leser standen den Online-Aktivitäten der *tageszeitung* in der Regel aufgeschlossen gegenüber (die *tageszeitung* war die erste deutschsprachige Tageszeitung im WWW und ist nach wie vor die einzige, die ihre Ausgabe im Internet vollständig und kostenlos zur Verfügung stellt). Viele Benutzer übermittelten Einschätzungen und Verbesserungsvorschläge. Es ist allerdings zweifelhaft, ob bei einem Nachfolgeprojekt mit einer derart intensiven Resonanz gerechnet werden darf – der Neuigkeitswert eines Projekts ‘Zeitung im Internet’ ist inzwischen nicht mehr in der damaligen Form gegeben.

6.1 Prototypingzyklen

Der Entwurfs- und Implementationsprozeß des in dieser Arbeit beschriebenen Systems gliederte sich in zwei Hauptzyklen. Während des ersten Zyklusses wurden, ausgehend von einem Minimal-system, sukzessive von Lesern vorgeschlagene Erweiterungen vorgenommen. Das so erweiterte System wurde während des zweiten Zyklusses einem Redesign unterzogen und um die Komponenten Leserbrief, Leserforum, Suchfunktion und Authentifizierung ergänzt. Der Vollständigkeit halber wird in diesem Kapitel auch auf das im Anschluß von der *tageszeitung* entwickelte Design eingegangen.

Erster Zyklus

Zunächst wurde die in Kapitel Kap. 4.1.1 beschriebene *seitenorientierte* Struktur implementiert, anfangs ohne die in Abb. 6 auf S. 34 dargestellten *links* zwischen aufeinanderfolgenden Dokumentenseiten und ohne vorangestellte lokale Seiteninhaltsverzeichnisse. Abgesehen von dem alle Artikelüberschriften umfassenden Hauptinhaltsverzeichnis entsprach diese Struktur einer 1:1 Übertragung von Zeitungs- auf WWW-Seiten (diese direkte Anlehnung an die Formatierung der Druckausgabe war für einen ersten Akzeptantest naheliegend). Das zugrundeliegende Benutzermodell setzte einen Leser voraus, der, nachdem er einen Artikel gelesen hat, die Überschriften der weiteren auf der Seite vorhandenen Artikel durch Vor- und Zurückrollen des Textes durchgeht und zum seitenbezogenen Gesamtverzeichnis durch Betätigen der in jedem Browser vorhandenen *Back*-Funktion zum Hauptinhaltsverzeichnis zurückspringt (bei dieser Vorgehensweise wird vom History-Mechanismus des Browsers Gebrauch gemacht).

Eine Reihe von Leserkomentaren deutete darauf hin, daß die Benutzerführung in mindestens zweierlei Hinsicht verbessert werden konnte. Erstens empfanden viele Online-Leser das Vor- und Zurückrollen innerhalb der Seitendokumente als unbequem. Auf ihre Anregungen hin wurde jedes Seitendokument um ein lokales Inhaltsverzeichnis mit *links* auf die in ihm enthaltenen Artikel erweitert. Dies erlaubte einen bequemerem Wechsel zwischen Titelübersicht und einzelnen Arti-

keln. Einige Leser wünschten sich eine stärkere Anlehnung der Benutzerführung an das 'Blättern' in gedruckten Zeitungen. Entsprechende Hinweise führten zur Einführung von Verweisen zwischen aufeinanderfolgenden Seitendokumenten. Das 'Umblättern' geschieht hierbei allerdings, ohne daß der Leser sicher ist, daß die von ihm angeforderte Folgeseite einen ihn interessierenden Artikel enthält. Dies ist nur für Benutzer geeignet, bei denen die Übertragung eines Seitendokuments nicht mehr als einige Sekunden in Anspruch nimmt. Es verwundert deshalb nicht, daß die Funktion des 'blinden Blätterns' von Online-Lesern gewünscht wurde, die über schnelle Verbindungen auf den WWW-Server des PRZ zurückgreifen konnten.

Viele Benutzer forderten eine erhöhte Selektivität bei der Übertragung von Dokumenten. Dies galt insbesondere für Leser, denen nur begrenzte Übertragungsbreiten zur Verfügung standen, z.B. wegen des Einsatzes eines Modems oder stark belasteter Transkontinentalverbindungen. Dies gab den Anstoß zur Implementation der *artikelorientierten* Struktur, bei der nur der angewählte Beitrag übertragen wird.

Weiterhin wurde angeregt, eine *themenbezogene* Übersicht einzuführen, die auf Inhaltsverzeichnisse einzelner Rubriken verweist. Interessanterweise wird diese Lösung jedoch von vielen Nutzern schmalbandiger Anschlüsse als suboptimal beurteilt. Trotz vergleichsweise langer Ladezeiten ziehen sie es in aller Regel vor, das etwa 35 kB große artikelorientierte Gesamtinhaltsverzeichnis anzufordern und von dort aus zu den sie interessierenden Artikeln zu springen (vgl. hierzu auch Kap. 8.0).

Zweiter Zyklus

Im Verlauf des ersten Prototypingzyklusses hatte sich ein genauer umrissenes Bild der von den Lesern gewünschten Navigationshilfen, Dokumentenstrukturen und Dienste herauskristallisiert. Auf dieser Grundlage wurde ein Neuentwurf des Systems vorgenommen. Dabei wurden alle wesentlichen Layoutelemente beibehalten, jedoch um die in Kap. 4.1.2 beschriebene einheitliche Navigationsleiste erweitert. Die verschiedenen Dokumentenstrukturen wurden in der in Kap. 4.1.1 und Kap. 4.1.2 beschriebenen Weise so integriert, daß während des Navigierens im Informationsangebot zwischen verschiedenen Darstellungsweisen gewechselt werden kann.

Während des ersten Zyklusses war das Problem aufgetaucht, daß sich viele Leserbriefe nicht oder nur schwer jenen Artikeln zuordnen ließen, auf die sie sich bezogen. Aus diesem Grund wurde in die Navigationsleiste ein parametrisierter *link* integriert, der den Aufruf des Leserbriefformulars mit standardisierter *Subject*-Zeile erlaubte (vgl. Kap. 5.4). Dies war auch im Hinblick auf nachvollziehbare Überschriften in Beiträgen zum Diskussionsforum unerlässlich. Das Forum ist das einzige Element des Angebots, das nicht auf expliziten Wunsch von Lesern eingeführt wurde. Mit Hilfe dieses Dienstes sollte überprüft werden, wie weit ein *n:m* Medium, das allen Beteiligten sowohl das Empfangen wie auch das Senden von Nachrichten erlaubt, im Rahmen eines Tageszeitungsangebots angenommen wird (vgl. dazu Kap. 7.3.2).

Die implementierte Suchfunktion stellt einen Mittelweg zwischen Leserwünschen und einschränkenden Vorgaben der *tageszeitung* dar. Sie ist für viele Benutzer ein offensichtlich unzureichender Kompromiß (der mit Abstand am häufigsten gewünschte Zusatzdienst für die Onlineausgabe ist

die Möglichkeit einer Volltextrecherche innerhalb von Monats-, Jahres- oder Gesamtarchiven der *tageszeitung*).

Redesign des Systems durch die *tageszeitung*

Ausgehend vom beschriebenen Prototyp haben Mitarbeiter der *tageszeitung* einen eigenen Entwurf realisiert. Dieser ist *kein* Bestandteil der vorliegenden Arbeit, sondern in gewisser Hinsicht deren Folgeprodukt. In diesem Entwurf finden sich jedoch viele Elemente des Prototyps wieder, Vor- und Nachteile des in dieser Arbeit entwickelten Prototyps und des taz-Systems werden einander im nächsten Kapitel gegenübergestellt.

Das durch Mitarbeiter der *tageszeitung* entwickelte System verwendet einen eigenen HTML- und Strukturgenerator. Dieser operiert nicht mehr auf einer provisorischen Archivdatei, sondern verwendet als Eingabe Artikeldateien aus dem Redaktionssystem. Die HTML-Formatierung erfolgt auf einem Rechner des Redaktionssystems.

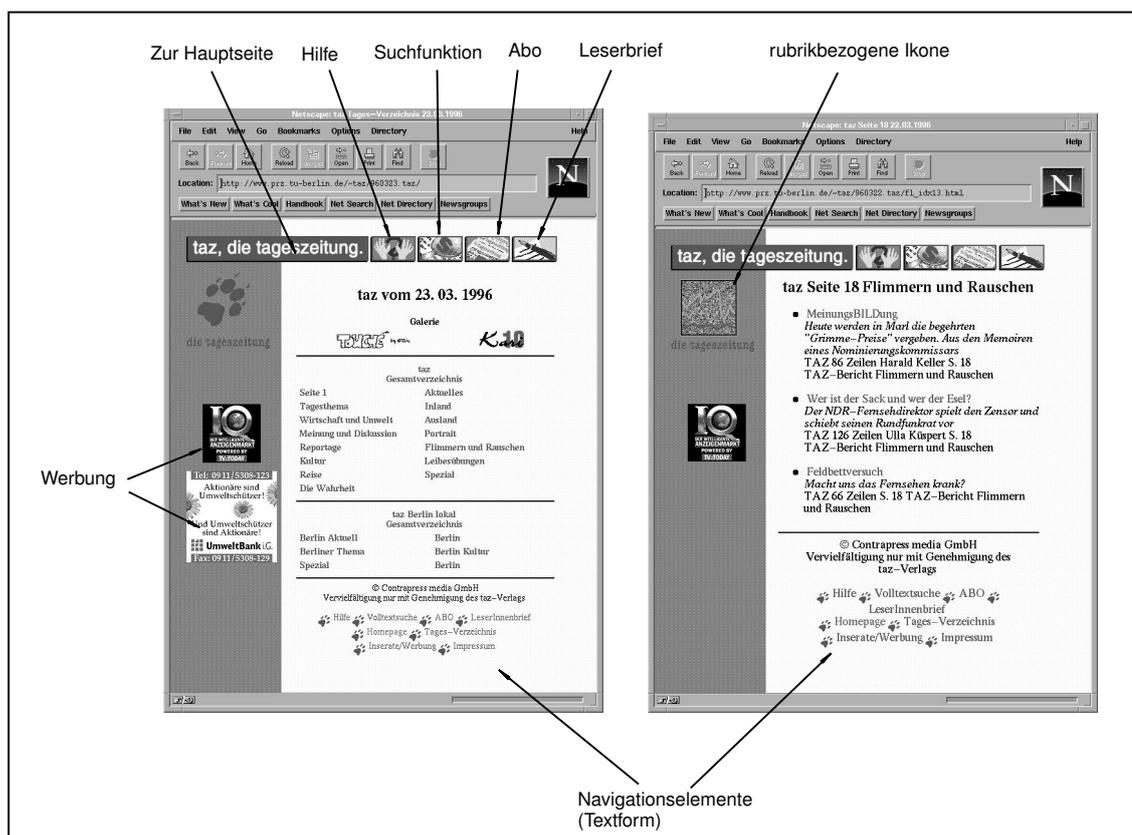


Abb. 22: Von der *tageszeitung* entwickeltes Layout

Für das neu entworfene Layout wurden z.T. nicht-standardisierte HTML-Direktiven wie Hintergrundgraphiken und Tabellen eingesetzt (vgl dazu Kap. 3.1.3). Dies wird dadurch gerechtfertigt, daß sich *Netscape* und *Microsoft Windows Explorer* mittlerweile als meistbenutzter Browser durchgesetzt hatten (eine Auswertung der während der Prototypingphasen angefallenen Server-Logdateien ergab, daß die genannten Browser am Ende des zweiten Zyklus von mehr als 90% aller Leser verwendet wurden).

Die von der *tageszeitung* entwickelte Dokumentenstruktur der Online-Ausgabe ist in drei Hierarchiestufen gegliedert. Ausgehend von einer Rubrikübersicht gelangt man zu den Rubrikinhaltsverzeichnissen und von dort zu den einzelnen Artikeln (das neue Layout stellte ursprünglich lediglich eine *artikelorientierte* Struktur zur Verfügung).

Ein ins Auge springendes Merkmal des neuen Layouts ist ein farblich abgesetzter linker Rand, der etwa 25% der Breite eines Dokuments umfaßt. Dieser Rand ist reserviert für die Platzierung von Werbe- und Informationsgrafiken sowie rubrikbezogenen Ikonen. Die Dienste Leserbrief, Suche und Hilfe werden über Ikonen am oberen Seitenrand aktiviert, ebenso das Abonnement-Bestellformular. Jedes dieser Navigationselemente ist auch am unteren Rand von Verzeichnissen und Artikeln vorhanden. Hier finden sich zusätzliche Verweise auf den nächsten Artikel, das Ressort- und Gesamtinhaltsverzeichnis und auf Informationen für potentielle Werbekunden. Abb. 22 zeigt das von der *tageszeitung* entwickelte Layout der Rubrikübersicht und eines Rubrikinhaltsverzeichnisses.

Alle Inhaltsverzeichnisse und Artikeltexte sind als *HTML-Tabellen* formatiert. Dies ermöglicht ein spaltenorientiertes Layout mit rechtsseitiger Textbegrenzung. Die Lesbarkeit der Texte erhöht sich hierdurch deutlich.

Es sei angemerkt, daß der Wunsch nach einer Integration kommerzieller Werbung ein entscheidenden Anstoß für den Neuentwurf des Layouts seitens der *tageszeitung* war. Diese Zielsetzung spielte jedoch bei der Realisierung des in dieser Arbeit entwickelten Systems keine Rolle.

6.2 Tests

Im Verlauf des ersten Prototypingzyklusses konnten eine Reihe Problemen isoliert und behoben werden, die im Zusammenhang mit unvollständigen oder fehlenden Informationen in der von der *tageszeitung* gelieferten Artikeldatei auftraten. Unbefriedigend blieb bis zum Schluß, daß aus technischen Gründen Artikel von Sonderseiten oder -beilagen zusammen mit den 'normalen' Artikeln der Tagesausgabe in einer gemeinsamen provisorischen Archivdatei übertragen wurden. Zumeist fehlten dabei Kontextinformationen, um automatisch zwischen 'normalen' Artikeln und solchen auf Sonderseiten unterscheiden zu können. Dieses Problem ließ sich nur durch mühsame manuelle Nachbearbeitung lösen. Hierbei wurde die provisorische Archivdatei zunächst in zwei Teildateien aufgespalten, von denen erste die 'normalen' Artikel, die zweite die Artikel der Sonderseiten enthielt. Die Teildateien wurden danach einzeln konvertiert und die konvertierten Dokumente in unterschiedlichen Verzeichnissen abgelegt.

Dateitransfer und die Dateikonvertierung erfolgten nach einer kurzen Experimentalphase zuverlässig. Übertragungsprobleme waren in aller Regel nachvollziehbar auf den Ausfall von Routern bzw. des WWW-Servers am PRZ zurückzuführen. Die am PRZ lokalisierten Schwachstellen wurden – teilweise durch Auswechseln von Hardware – weitestgehend eliminiert, so daß der Dienst nach einigen Wochen als stabil bezeichnet werden konnte.

Experimente mit digitaler Signierung von Artikeln und Spiegelung der Ausgabe wurde auf wenige Tage beschränkt. Dabei wurden beide in Kap. 5.3 erwähnten Verfahren getestet, d.h., das vorher-

rige Signieren aller Artikel vor Freigabe des Angebots und das Signieren nach Freigabe des Angebots. Je nach Umfang der Ausgabe und Auslastung des Servers betrug die Gesamtzeit für den Signierprozeß zwischen 30 und 60 Minuten. Da alle Artikel und Artikelverzeichnisse mit Authentifikationsmerkmalen versehen sein müssen, bevor sie gespiegelt werden können, stand das aktuelle Angebot auf den Spiegelservers beträchtlich später als auf dem Ausgangsserver am PRZ zur Verfügung.

Die Leserbrief- und Suchfunktionen waren praktisch auf Anhieb funktionstüchtig. Probleme tauchten lediglich bei der Indizierung und Konvertierung elektronischer Nachrichten an das Leserbriefforum auf (vgl. Kap. 5.4). Bei der ursprünglich gewählten `-u` Option von *hypermail*, die die Datenbasis um die neu hinzugekommenen Nachrichten erweitern soll, kam es mehrfach zu falschen Indizierungen. Konvertierungen einzelner *emails* waren fehlerhaft oder fehlten vollständig. Dies wurde durch Verwendung der *hypermail*-Option `-x` behoben, durch die eine Reindizierung und -konvertierung des gesamten *email*-Archivdatei durchgeführt wird. Dies ist ein vergleichsweise zeitaufwendiges Verfahren, das jedoch zuverlässig korrekte Ergebnisse lieferte.

7.0 Ergebnisse und Auswertung

Zunächst wird in Kap. 7.1 die Auswertung der Leserumfrage vorgestellt. In Kap. 7.2 wird der Bekanntheitsgrad des Online-Angebots der *tageszeitung* vergleichbaren deutschsprachigen Angeboten gegenübergestellt. Kap. 7.3 diskutiert die im Verlauf der Leserbefragung und des Prototypings gesammelten Ergebnisse und Erfahrungen. Dieses Kapitel stellt die unseres Wissens erste veröffentlichte Untersuchung von Benutzerprofilen und -verhalten bei WWW-gestützten Tageszeitungen in der Bundesrepublik dar.

7.1 Ergebnisse der Leserumfrage

Um die Akzeptanz des Online-Angebots der *tageszeitung* zu ermitteln, wurde ein Fragebogen in Form eines WWW-Formulars entwickelt. Hierdurch sollten Anhaltspunkte für folgende Fragestellungen gefunden werden:

- Über welche Anbieter und von wo aus greifen die Leser auf das Online-Angebot zu?
- Sind die Leser prinzipiell bereit, das Angebot zu abonnieren?
- Sind die Leser prinzipiell bereit, für das Angebot zu bezahlen, und falls ja, wieviel?
- Welches sind die wichtigsten Vor- und Nachteile des Online-Angebots im Vergleich zur gedruckten Ausgabe?

Der Wortlaut dieser Umfrage findet sich in Anhang C. Im Zeitraum von ca. 10 Monaten (5.5.1995 bis 23.3.1996) gingen etwa 3000 Fragebögen ein, von denen 2292 Formulare auswertbare Informationen enthielten. Die genaue Anzahl der jeweils zur Bearbeitung einer Fragestellung einbezogenen Formulare wurde in jeder der in diesem Kapitel abgebildeten Tabellen optisch hervorgehoben.

7.1.1 Internetanbieter und Zugriffsort

In Tab. 3 wurden die Fragebögen nach angegebenem Internetanbieter und Zugriffsort aufgeschlüsselt. Aus der Tabelle geht hervor, daß der überwiegende Teil der Leserschaft über Universitäreinrichtungen auf das Online-Angebot der *tageszeitung* zugreift. Aller Voraussicht nach wird sich der Anteil der über private Anbieter angeschlossenen Benutzer in Zukunft deutlich erhöhen. Darauf deuten Auswertungen der Server-Log-Dateien hin, die eine kontinuierliche Zunahme von Zugriffen kommerzieller Provider verzeichnen. Das Leserprofil der Online-Ausgabe dürfte in vielen Punkten demjenigen der Druckausgabe entsprechen (hoher Akademiker- und Ausbildungsanteil.)

| | Universtät | Medien | kommerz. Provider | Firma | SUMME |
|----------------------------|------------|--------|-------------------|-------|-------|
| Zugriff von der Arbeit aus | 1029 | 20 | 64 | 283 | 1396 |
| Zugriff von zuhause aus | 376 | 37 | 306 | 34 | 753 |
| Summe | 1405 | 57 | 370 | 317 | 2149 |

TABELLE 3. : Leser nach Internetanbietern und Zugriffsort

7.1.2 Abonnementsbereitschaft

Die zweite Fragestellung war, ob Online-Leser prinzipiell bereit sind, das WWW- Angebot der *tageszeitung* als elektronisches Abonnement zu beziehen. Eine Zusammenfassung der Antworten wird von Tab. 4 wiedergegeben. Ihr zufolge ist die Abonnementsbereitschaft jener Leser, die von zuhause aus auf das Angebot zugreifen, in der Regel höher als die jener Leser, die über Internetzugänge am Arbeits- oder Ausbildungsort verfügen. Bei der erstgenannten Gruppe ist das Bewußtsein dafür, daß die Nutzung eines Online-Informationsangebots mit Gebühren verbunden sein kann, offensichtlich ausgeprägter als bei jenen Lesern, die Internet-Ressourcen im Rahmen von Arbeit und Ausbildung kostenlos nutzen können.

| Internetanbieter | Zugriff v. | Abo ja | Abo nein |
|------------------|------------|--------|----------|
| Universität | der Arbeit | 328 | 619 |
| | zuhause | 132 | 226 |
| Medien | der Arbeit | 11 | 7 |
| | zuhause | 14 | 24 |
| komm. Provider | der Arbeit | 28 | 33 |
| | zuhause | 127 | 171 |
| Firma | der Arbeit | 98 | 169 |
| | zuhause | 21 | 18 |
| Summe | 2025 | 757 | 1268 |

TABELLE 4. Abonnementsbereitschaft

Erwähnenswert ist der Medienbereich, in dem die Abonnementsbereitschaft prozentual am höchsten ist. Der Zugang zu laufend aktualisierten, schnell verfügbaren Informationen aus vertrauenswürdigen Quellen am Arbeitsplatz stellt für diese Berufsgruppe natürlich einen Vorteil dar. Die für den Medienbereich angegebenen Werte sollten allerdings wegen der geringen Zahl entsprechender Antworten (31) vorsichtig interpretiert werden.

7.1.3 Finanzieller Wert

Von besonderem Interesse war die Fragestellung, wie die Leser den finanziellen Wert der Online-Ausgabe im Vergleich zur Druckausgabe einschätzen (die entsprechende Frage lautete: *Was wäre Ihnen die digiTaz im Monat wert?*) Die Zusammenfassung der Antworten findet sich in Tab. 5. Hierbei entsprechen die in Spalte 1 angegebenen Werte der Anzahl von Fragebögen, in denen explizit ein Betrag genannt wurde. Die in Spalte 2 angegebenen Werte entsprechen der Anzahl von Fragebögen, in denen dies nicht der Fall war. Spalte 3 liefert den Durchschnittsbetrag aller explizit genannten Beträge. Hierbei wurden nur Antworten berücksichtigt, bei denen der angegebene Betrag zwischen DM 0.00 und DM 50.00 liegt. Größere Wertangaben wurden als offensichtlich unsinnig verworfen. Während der in Spalte 3 angegebene Durchschnittsbetrag nur über jene Fragebögen gebildet wurde, in denen Beträge genannt wurden (entspricht der in Spalte 1 angegebenen

Anzahl), wurde in Spalte 4 der Mittelwert über *alle* Fragebögen einer Gruppe errechnet (entspricht der Summe von Spalte 1 und 2).

Die Werteinschätzung von Lesen im Ausland (525 von 2070 ausgewerteten Fragebögen, d.h., ca.25% der Leser) weicht nicht nennenswert von derjenigen der Gesamtstatistik ab. 228 von ihnen stehen einem Abonnement aufgeschlossen gegenüber, 232 gaben ein Betrag an (Mittelwert entsprechend Tab. 5, Spalten 3 und 4: DM 16.80 bzw. DM 7.42).

| Internetanbieter | Zugriff von | Betrag genannt | kein Betrag genannt | Ø Betrag der Gruppe (1) | Ø Betrag der Gruppen (1 und 2) |
|------------------|-------------|----------------|---------------------|-------------------------|--------------------------------|
| | | (1) | (2) | (3) | (4) |
| Uni | Arbeit | 378 | 636 | DM 15.78 | DM 5.46 |
| | zu Hause | 160 | 205 | DM 16.83 | DM 6.20 |
| Medien | Arbeit | 8 | 9 | DM 14.55 | DM 8.09 |
| | zu Hause | 13 | 19 | DM 11.07 | DM 4.08 |
| komm. Provider | Arbeit | 28 | 33 | DM 24.18 | DM 10.92 |
| | zu Hause | 144 | 162 | DM 18.42 | DM 7.85 |
| Firma | Arbeit | 113 | 160 | DM 17.97 | DM 6.51 |
| | zu Hause | 23 | 10 | DM 19.76 | DM 10.64 |
| Summe | 2070 | 836 | 1234 | | |

TABELLE 5. Subjektiver Wert der Online-Ausgabe pro Monat nach Internetanbieter und Zugriffsort

Für genaueren Überblick über die statistische Verteilung der genannten Beträge wurden die insgesamt 836 angegebenen Werte (vgl Tab. 5, Spalte 1) nach sieben Gruppen geordnet. Das Resultat wird von Tab. 6 wiedergegeben. Die vier linken Diagramme von Abb. 23 auf S. 71 enthalten graphische Darstellungen der in Tab. 6 aufgeschlüsselten Verteilung.

| Internetanbieter | Zugriff von | bis DM 5 | bis DM 10 | bis DM 15 | bis DM 20 | bis DM 25 | bis DM 30 | > DM 30 |
|------------------|-------------|----------|-----------|-----------|-----------|-----------|-----------|---------|
| Universität | Arbeit | 94 | 131 | 35 | 62 | 12 | 24 | 18 |
| | zu Hause | 40 | 48 | 19 | 27 | 4 | 18 | 4 |
| Medien | Arbeit | 2 | 1 | 0 | 2 | 0 | 2 | 1 |
| | zu Hause | 3 | 6 | 1 | 2 | 0 | 1 | 0 |
| komm. Provider | Arbeit | 1 | 4 | 2 | 6 | 3 | 6 | 6 |
| | zu Hause | 34 | 35 | 12 | 32 | 7 | 14 | 10 |
| Firma | Arbeit | 31 | 23 | 14 | 27 | 1 | 12 | 5 |
| | zu Hause | 4 | 8 | 2 | 2 | 3 | 0 | 4 |
| Summe | 861 | 211 | 256 | 85 | 160 | 30 | 77 | 42 |

TABELLE 6. Subjektiver Wert der Online-Ausgabe pro Monat nach Internetanbieter, Zugriffsort und Preis, Gesamtverteilung für alle ausgefüllten Fragebogen mit entsprechenden Angaben

Um überprüfen zu können, ob die Verteilung der Werteinschätzung bei jenen Lesern, die einem elektronischen Abonnement gegenüberstehen, von derjenigen der übrigen Leser abweicht, wurde für die entsprechenden 558 Angaben eine separate Verteilung ermittelt (vgl Tab. 7 bzw. die entsprechenden graphischen Darstellungen in der rechten Spalte von Abb. 23 auf S. 71).

| Internetanbieter | Zugriff von | bis DM 5 | bis DM 10 | bis DM 15 | bis DM 20 | bis DM 25 | bis DM 30 | > DM 30 |
|------------------|-------------|----------|-----------|-----------|-----------|-----------|-----------|---------|
| Universität | Arbeit | 40 | 87 | 26 | 39 | 11 | 23 | 11 |
| | zu Hause | 18 | 32 | 18 | 18 | 2 | 12 | 4 |
| Medien | Arbeit | 2 | 0 | 0 | 2 | 0 | 2 | 1 |
| | zu Hause | 0 | 5 | 1 | 2 | 0 | 1 | 0 |
| komm. Provider | Arbeit | 0 | 1 | 2 | 5 | 2 | 3 | 6 |
| | zu Hause | 14 | 15 | 10 | 25 | 6 | 10 | 10 |
| Firma | Arbeit | 11 | 13 | 13 | 22 | 1 | 10 | 3 |
| | zu Hause | 2 | 7 | 2 | 1 | 3 | 0 | 4 |
| Summe | 558 | 87 | 160 | 72 | 114 | 25 | 61 | 39 |

TABELLE 7. Subjektiver Wert der Online-Ausgabe pro Monat nach Internetanbieter, Zugriffsort und Preis, Verteilung für Leser, die einem Abonnement aufgeschlossen gegenüberstehen

7.1.4 Vor- und Nachteile des Online-Angebots

Der vierte Fragenkomplex bezog sich auf die wichtigsten Vor- und Nachteile der Online-Ausgabe. Dazu wurden eine Anzahl von Positiv- und Negativkriterien vorgegeben. Weitergehende Anmerkungen konnten in Freitextfelder eingetragen werden.

Wir beschränken uns auf die Auswertung der fest vorgegebenen Kriterien. Die Ergebnisse werden von den Tabellen Tab. 8.a und Tab. 8.b wiedergegeben.

| | | | | | | | | | | | | | | | | | | | | |
|--|--|-----|--------------------------|-----|---------------|----|------------------|-----|--------------|------|--|--|-----|--|-----|------------------|-----|--------------|------|--|
| <table border="1"> <tr> <td>Möglichkeit der gezielten Suche</td> <td>815</td> </tr> <tr> <td>frühes Erscheinen</td> <td>615</td> </tr> <tr> <td>Bilder</td> <td>56</td> </tr> <tr> <td>Sonstiges</td> <td>424</td> </tr> <tr> <td>Summe</td> <td>1910</td> </tr> </table> | Möglichkeit der gezielten Suche | 815 | frühes Erscheinen | 615 | Bilder | 56 | Sonstiges | 424 | Summe | 1910 | <table border="1"> <tr> <td>gewöhnungsbedürftige Handhabung</td> <td>174</td> </tr> <tr> <td>geringe Übertragungsgeschwindigkeit</td> <td>653</td> </tr> <tr> <td>Sonstiges</td> <td>533</td> </tr> <tr> <td>Summe</td> <td>1360</td> </tr> </table> | gewöhnungsbedürftige Handhabung | 174 | geringe Übertragungsgeschwindigkeit | 653 | Sonstiges | 533 | Summe | 1360 | <p>Summen bezogen auf eine Gesamtanzahl von 2292 ausgewerteten Formularen. Die aufgeführten Optionen stellten Alternativen dar, von denen nur jeweils eine ausgewählt werden konnte.</p> |
| Möglichkeit der gezielten Suche | 815 | | | | | | | | | | | | | | | | | | | |
| frühes Erscheinen | 615 | | | | | | | | | | | | | | | | | | | |
| Bilder | 56 | | | | | | | | | | | | | | | | | | | |
| Sonstiges | 424 | | | | | | | | | | | | | | | | | | | |
| Summe | 1910 | | | | | | | | | | | | | | | | | | | |
| gewöhnungsbedürftige Handhabung | 174 | | | | | | | | | | | | | | | | | | | |
| geringe Übertragungsgeschwindigkeit | 653 | | | | | | | | | | | | | | | | | | | |
| Sonstiges | 533 | | | | | | | | | | | | | | | | | | | |
| Summe | 1360 | | | | | | | | | | | | | | | | | | | |
| <p>TABELLE 8. a Vorteile des Angebots</p> | <p>TABELLE 8. b Nachteile des Angebots</p> | | | | | | | | | | | | | | | | | | | |

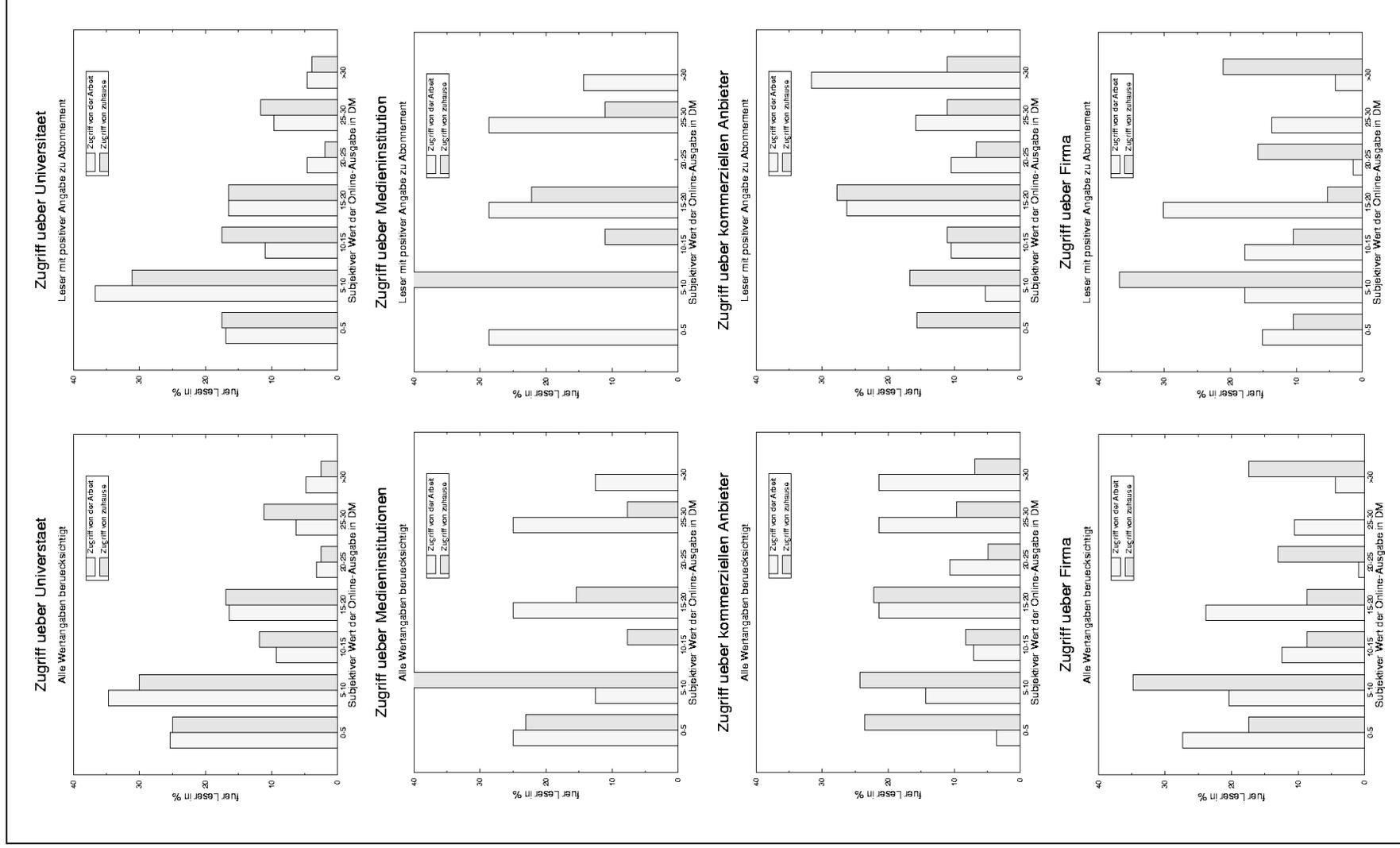


Abb. 23: Normalisierter subjektiver Wert der Online - Ausgabe pro Monat nach Internetanbieter und Zugriffsort

7.2 Ermittlung des Bekanntheitsgrads

In diesem Unterkapitel wird versucht, den Bekanntheitsgrad des Online-Angebot der *tageszeitung* zu ermitteln und vergleichbaren deutschsprachigen Online-Angeboten gegenüberzustellen. Das hierfür verwendete Kriterium ist die *Anzahl der Verweise auf Online Angebote* (vgl Tab. 9).

Die Anzahl der Verweise wurde über Anfragen an den *Alta Vista Index*¹ bestimmt. Dieser WWW-Index erlaubt die Eingabe von *URLs* als Suchbegriffen und liefert u.a. die Anzahl derjenigen

| Zeitung/Zeitschrift | Überprüfte URLs (14.04.95, 11.00 - 12.00h) | Anz. Verweise |
|-------------------------|--|---------------|
| Der Spiegel | http://www.spiegel.de http://muenchen.bda.de:80/bda/int/sponel.de | ca. 2000 |
| Focus | http://www.focus.de | ca. 500 |
| Holstein-Courier | http://www.courier.de | 121 |
| Mittelbayrische Zeitung | http://www.donau.de/mz | 93 |
| Morgenpost Hamburg | http://www.mopo.de | ca. 500 |
| Neue Passauer Presse | http://www.vgp.de | 29 |
| Rheinischer Merkur | http://www.merkur.de | 0 |
| Rheinzeitung | http://primus.cicero.de | ca. 400 |
| Saarbrücker Zeitung | http://www.sz-sb.de | ca 800 |
| Singener Wochenblatt | http://www.emb.net/wochenblatt | 0 |
| Schwäbisches Tagblatt | http://www.cityinfonetz.de/tagblatt | 126 |
| Schweriner Volkszeitung | http://www.hansenet.de/SVZ/ http://muenchen.bda.de/bda/nat/hn/SVZ | 49 |
| Stern | http://www.stern.de | ca. 800 |
| Süddeutsche Zeitung | http://www-dw.gmd.de/sz | ca. 400 |
| Super-Illu | http://www.super-illu.de/ | 3 |
| Tagesspiegel Berlin | http://www.tagesspiegel-berlin.de | ca. 500 |
| die tageszeitung | http://www.prz.tu-berlin.de/~taz http://www.taz.de | ca. 2000 |
| TV Movie | http://www.tvmovie.de | ca. 1000 |
| TV Today | http://www.wochenpost.de | ca. 1000 |
| Die Welt | http://www.welt.de | ca 3000 |
| Wochenpost | http://www.wochenpost.de | 900 |
| Die Zeit | http://www.zeit.d | 88 |

TABELLE 9. Bekanntheitsgrad der Online-Angebots deutschsprachiger Zeitungen

Dokumente zurück, die die eingegebene *URL* als *link* enthalten. Die von *Alta Vista* ermittelten Werte sind nicht exakt reproduzierbar, weil die Suchtiefe der Datenbankapplikation abhängig von-

1. URL: <http://altavista.digital.com/>

der Auslastung des Servers ist¹. Zu verschiedenen Tageszeiten durchgeführte Tests ergaben für den gleichen Suchbegriff Abweichungen von etwa +/-10%.

Die Anfragen erfolgten in allen Fällen als *Simple Query*² mit der von *Alta Vista* vorgegebenen Syntax zur Ermittlung von Verweisen auf WWW-Dokumente³. Die Ergebnisse werden von Tab. 9 wiedergegeben. Dabei muß beachtet werden, daß von *Alta Vista* alle Zahlenangaben gerundet werden, die größer als 200 sind. Um die Werte für *tageszeitung* und *Welt* zu präzisieren, wurden für beide Angebote eine Reihe von *Advanced Query*⁴ Anfragen für wöchentliche Zeitintervalle seit März 1995 gestellt. Die zurückgelieferten Resultate waren in diesem Fall immer kleiner als 200, wodurch die Rundungsfunktion von *Alta-Vista* ausgeschaltet wurde. Die Summierung der erhaltenen Werte ergab für die *tageszeitung* 2222 Referenzierungen, für die *Welt* 2952 (beide Zahlen vom 14.04.1996. 12:00h).

7.3 Auswertung

Aus Tab. 9 geht hervor, daß das Online-Ausgabe der *tageszeitung* (hinter demjenigen der *Welt*) das am zweithäufigsten referenzierte WWW-Angebot bundesdeutscher Tageszeitungen ist. Für eine ungefähre Bestimmung der Anzahl täglicher Online-Leser können die in Tab. 10 angegebenen Zahlen für Zugriffe auf die Gesamtverzeichnisse herangezogen werden. Da sich die monatlichen Zugriffstatistiken stark ähneln, wurde die Auswertung vom Oktober als repräsentativ herangezogen. Unter der Voraussetzung, daß jeder Leser genau einmal ein Gesamtinhaltsverzeichnis anfordert, ergibt sich z.B. für den Oktober 1995 eine Zahl von durchschnittlich 1559 Lesern pro Tag. Dieser Wert ist als *untere Schranke* anzusehen, da einmal geholte WWW-Seiten in vielen Fällen in Caches von Proxy-Servern zwischengespeichert werden (vgl Kap. 3.1.2).

| | |
|---|---------------|
| Zugriffe auf seitenorientiertes Gesamtverzeichnis | 33881 (68.8%) |
| Zugriffe auf artikelorientiertes Gesamtverzeichnis | 14439 (29.3%) |
| Zugriffe auf rubrik- und artikelbezogenes Verzeichnis der Seite 1 | 941 (1.9%) |
| Aufrufe der Suchfunktion | 4287 |
| Aufrufe des Leserforums | 1663 |

TABELLE 10. Zugriff auf ausgewählte Seiten des Angebots im Zeitraum vom 1.10. - 31.10.1995

Während des Projektzeitraums konnten weder positive noch negative Auswirkungen auf Abonnements- und Kioskverkaufszahlen festgestellt werden. Dies gilt auch für die Nutzung des über GENIOS angebotenen, kostenpflichtigen Archivdienst der *tageszeitung*. Die Möglichkeit einer jahrgangübergreifenden Volltextsuche ist der von der Online-Leserschaft am häufigsten gewünschten zusätzliche Dienst.

1. persönliche Mitteilung des Systemadministrators von *altavista.digital.com*

2. <http://altavista.digital.com/cgi-bin/query/pg=q&what=web>

3. "link:<Ausdruck> -url:<Ausdruck>", vgl URL <http://altavista.digital.com/cgi-bin/query?pg=tips>

4. URL <http://altavista.digital.com/cgi-bin/query?pg=aq&what=web>

7.3.1 Auswertung der Leserumfrage

Tab. 3 zeigt, daß mehr als ein Drittel aller Benutzer das Online-Angebot der *tageszeitung* von zuhause liest, also nur über vergleichsweise schmalbandige Verbindungen (beim derzeitigen Stand der Technik sind dies Modems mit 14400 baud) verfügen. Für diese Benutzergruppe dürfte ein auf möglichst geringen Datendurchsatz hin optimiertes Layout besonders hohe Bedeutung haben, da neben Wartezeiten auch Telefongebühren anfallen.

Laut Tab. 8.b werden Zugriffs- und Übertragungsgeschwindigkeiten von fast einem Drittel der Benutzer als zu gering eingeschätzt. Auch dies stützt legt nahe, die Menge der übertragenen Daten möglichst klein zu halten. Erhöhte Selektivität ("gezielt") ist laut Tab. 8.a das wichtigste Positivmerkmal. Andererseits führten viele Benutzer die *erzwungene* Selektivität im Freitextfeld als Negativmerkmal auf (diese Aussage wird nicht quantitativ präzisiert, da dies nur durch manuelle Auswertung des entsprechenden Freitextfeldes möglich wäre).

Die in Tab. 4 dargestellten Werte besagen, daß mehr als ein Drittel der Online-Leserschaft prinzipiell bereit ist, das Angebot als Abonnement zu beziehen. Dieser Wert läßt sich wohl nicht ohne weiteres auf die potentielle *Zahlungswilligkeit* der Leser übertragen. Solange keine bequemen Zahlungsverfahren und Mikroabrechnungsmethoden (Abbuchung von Kleinbeträgen pro angefordertem Artikel) zur Verfügung stehen, dürfte die Zahl jener Leser, die tatsächlich eine Abonnementsgebühr an die *tageszeitung* abführen würden, geringer sein als der oben genannten Wert. Darauf zumindest deuten viele Kommentare in den Freitextfeldern des Fragebogens hin. Außerdem kann eine mögliche systematischen Verzerrung des Ergebnisses nicht ausgeschlossen werden: denkbar ist, daß die antwortenden Personen durch bewußt falsche Angaben die Chancen für eine Weiterführung des Online-Dienstes zu erhöhen versuchen. Diese Einschränkung gilt auch für alle folgenden Ergebnisse.

Aus den Gegenüberstellung der Diagramme in Abb. 23 ist ablesbar, daß die von den Lesern angegebenen Summen stark streuen. Die in Tab. 5 angegebenen *Mittelwerte* sind aus diesem Grund nur eingeschränkt als Grundlage für eine Preisgestaltungsstrategie brauchbar. Es böte sich an, verschiedene Preiskategorien zu bilden. Dabei könnten höhere Gebühren mit zusätzlichen Dienstangeboten verknüpft werden, etwa der Option der Recherche im Gesamtarchiv, früherer Zeitpunkt des Zugriffs, Authentifizierungsmechanismen und individuell zugeschnittene Vorauswahl von Artikeln anhand von Benutzerprofilen. Eine derartige Staffelung von Abonnementspreisen gibt es bereits bei der Druckausgabe der *tageszeitung*, die zu drei unterschiedlichen Sätzen vertrieben wird (ermäßigter, normaler und politischer Preis). Hier erhält jedoch jeder Abonnent die Zeitung in gleichem Umfang und gleicher Qualität zugestellt - unabhängig vom gezahlten Tarif. Ob für das Online-Angebot eine Preisstaffelung mit unterschiedlichen Dienstqualitäten verknüpft werden sollte, ist eine nicht-triviale Marketingfrage.

Im Rahmen dieser Arbeit konnte aus Zeitgründen nur eine vergleichsweise rudimentäre Fragebogenauswertung vorgenommen werden. So mußte leider vollständig auf eine Kategorisierung und Darstellung der in den Freitextfeldern des Formulars enthaltenen Hinweise verzichtet werden. Die detaillierte Auswertung des Datenmaterials könnte Bestandteil einer weiterführenden Arbeit sein

7.3.2 Auswertung der Prototypingzyklen und des laufenden Betriebs

Ein hypertextgestütztes Online-Angebot für Tageszeitungen ist medienbedingt stärker strukturiert ist als ihr gedrucktes Korrelat. Benutzer möchten mit möglichst wenigen Aktionen zu den sie interessierenden Beiträgen gelangen. Hierzu ist eine starke Hierarchisierung der Dokumentenstruktur nötig, die tendenziell den selektive Umgang mit dem Informationsangebot verstärkt. Diese Beobachtung wird durch entsprechende Ergebnisse der Fragebogenauswertung gestützt (vgl. Kap. 7.1.4).

Der *Zwang* zu selektivem Vorgehen wird allerdings von vielen Lesern als notwendiges Übel angesehen. Diese Aussage wird durch die exemplarische Auswertung der Logdateieinträge untermauert (vgl. Tab. 10). Es erfolgten 68.8% aller Verzeichniszugriffe auf den *seitenorientierten* Gesamtindex (vgl. Kap. 4.1.1), jedoch nur 29.3% auf den *artikelorientierten* Gesamtindex. Der rubrik- und artikelorientierte Index der Titelseite wurde von lediglich 1.9% der Leser genutzt. Trotz der zur Verfügung stehenden Option, einzelne Artikel gezielt abzurufen, zog die Mehrzahl der Benutzer vor, mit dem ausgewählten Artikel gleichzeitig dessen Kontext, d.h., alle Artikel der entsprechenden Seite, übertragen zu bekommen. Die daraus ableitbare Schlußfolgerung lautet: längere Wartezeiten werden tendenziell in Kauf genommen, wenn die gleichen Art und Anzahl von Benutzeraktion zur Übertragung von mehr Information (Artikeln) führte.

Die von Online-Lesern vorgebrachte Kritik am (in der zweiten Prototypingphase) überarbeiteten Design bemängelte in der Regel die große Anzahl der in Gesamtinhaltsverzeichnissen und Seitendokumenten integrierten Navigationselemente. Positiv wurde vermerkt, daß mehrere Darstellungsalternativen verfügbar waren, von denen der Benutzer die ihm genehme wählen kann.

An dieser Stelle bietet sich ein Vergleich mit den Benutzerreaktionen nach Einführung des von der *tageszeitung* überarbeiteten Layouts an. Dieses stellte anfangs ausschließlich rubrikbezogene Inhaltsverzeichnisse und eine artikelorientierte Dokumentenstruktur zur Verfügung, wobei man davon ausging, daß diese Struktur für alle Leser angemessen ist (zum damaligen Zeitpunkt war die in dieser Arbeit vorgestellte statistische Auswertung noch nicht verfügbar). Dies stellte sich jedoch als Fehlannahme heraus: es kam zu einer Anzahl von Leserprotesten. Mittlerweile wird deshalb zusätzlich das vom Prototyp her bekannte, nach Rubriken gegliederte, artikelorientierte Gesamtverzeichnis angeboten. Es kann aus der Rubrikübersicht heraus aufgerufen werden.

Das von der *tageszeitung* entwickelte neue Layout führt besonders bei Inhaltsverzeichnissen zu einer Vervielfachung der übertragenen Datenmenge. Es wird eine Dokumentenstruktur erzeugt, in der es nicht mehr möglich ist, mehrere Artikel einer Rubrik mit nur einer Aktion anzufordern. Die Analyse der Logdateien (vgl. Tab. 10) legt jedoch nahe, daß genau diese Option angeboten werden sollte. Die in dieser Arbeit entwickelten multiplen Dokumentenstrukturen bieten hingegen verschiedene Navigationsalternativen für den gleichen Informationskorpus an. Diese Entwurfsentscheidung wird durch die empirisch nachweisbaren, unterschiedlichen Leservorlieben (ca. 70% seitenorientierte Navigation, ca. 30% artikelorientierte Navigation) gerechtfertigt. Das (auch beim Prototypen notwendige) ständige Umschalten zwischen Inhaltsverzeichnis und Artikels wurde

von verschiedenen Benutzern als lästig empfunden. Im Schlußkapitel dieser Arbeit wird kurz auf eine mögliche Lösung dieses Problems eingegangen.

Kritisiert wurden vor allem die starke graphische Orientierung des von der *tageszeitung* entwickelten Layouts, die den für die Übertragung einer Seite notwendigen Datendurchsatz etwa verdreifacht. Während der vermehrte Einsatz von Grafiken und Ikonen von Teilen der Leserschaft als ästhetische Zumutung empfunden wurde, stieß diese Entwurfsentscheidung bei anderen Nutzern auf ungeteilte Zustimmung.

Stichwortsuche und Nutzung des Leseforums bilden einen vergleichsweise kleinen Anteil der Leseraktivitäten (zusammen ca. 1% aller Aufrufe). Im Fall der Suchfunktion dürfte dies darauf zurückzuführen sein, daß lediglich die aktuelle Tagesausgabe nach Stichworten durchsucht werden kann. Es läßt sich indes nur schwer abschätzen, wie stark die Nutzung der Funktion zunehmen würde, wenn z.B. ein kompletter Jahrgang oder das Gesamtarchiv der *tageszeitung* durchsucht werden könnten.

Wie aus Tab. 10 hervorgeht, wurde das Leserforum im Oktober 1995 insgesamt 1663 mal aufgerufen. Im Verlauf von vier Monaten wurden 164 Nachrichten (1-2 Artikel pro Tag) an das Forum geschrieben, wobei sich 13 Nachrichten aufeinander bezogen. Das elektronische Leserforum wurde demnach fast ausschließlich in der Art und Weise traditioneller Leserbriefe (d.h., als Forum für *Statements*) genutzt, weniger im Sinne einer USENET-Newsgroup (d.h., als Forum für *Diskussionen*). Die relativ geringe Akzeptanz des Forums dürfte hauptsächlich auf das vom Programm *hypermailer* erzeugte, unzulängliche Layout zurückzuführen sein (zum Vergleich: im hervorragend gestalteten WWW-Leserforum der Tageszeitung *Die Welt* gingen im Zeitraum vom 28. März bis zum 11 April 1996 die Anzahl von 149 Nachrichten ein, was einem Aufkommen von ca. 10 Leserbeiträgen pro Tag entspricht).

Im Gegensatz zur vergleichsweise geringen Zahl an Beiträgen im Leserforum stand die Menge an Nachrichten, die direkt per *email* an die *tageszeitung* gerichtet wurden (10-20 pro Tag). Während der Anfangsphase des Projekts waren die *taz*-internen Zuständigkeiten zur Bearbeitung von *emails* noch nicht ausreichend geregelt. In dieser Zeit wurde ein großer Teil der elektronischen Leserpost am PRZ bearbeitet, was z.T. mehrere Stunden pro Tag in Anspruch nahm. Bei optimierter Arbeitsorganisation kann dieser Zeitraum sicherlich verkürzt werden. Unserer Erfahrung nach erfordert die (inhaltliche) Bearbeitung der Lesermittelungen und -anfragen jedoch ein Vielfaches jener Zeit, die für die laufende (technische) Wartung des Angebots nötig ist.

Im Rahmen des beschriebenen Projekt konnte demonstriert werden, daß die Bereitstellung eines aktualisierten, WWW-gestützten Online-Angebots für eine Tageszeitung weitgehend ohne zusätzlichen Personalaufwand möglich ist. Da das Quellformat ASCII-orientiert war, hielt sich der Implementationsaufwand mit etwa drei Mann-Monaten in Grenzen. Auf der Grundlage der in diesem Projekt gesammelten Erfahrungen war eine vollständige Überarbeitung des Online-Angebots durch die *tageszeitung* mit etwa als einem Mann-Monat möglich.

8.0 Zusammenfassung und Ausblick

Die technische Machbarkeit einer automatisch erstellten, täglich aktualisierten, WWW-gestützten Online-Zeitung in einem verteilten System konnte in allen vorgegebenen Punkten (vgl. Kap. 3.3) demonstriert werden. Als praktisches Ergebnis konnte ein stabil arbeitender Prototyp präsentiert werden, der im täglich von durchschnittlich 1500 Personen genutzt wurde. Die Größenordnung der monatlich übertragenen Dokumente und Graphiken lag bei etwa einer halben Million (ca. 6 GB).

Das in dieser Arbeit beschriebene System war der bundesweit erste Prototyp eines Tageszeitungsangebots im WWW. Seine technische Besonderheit liegt darin, daß auf das zur Verfügung gestellte Informationsangebot über mehrere parallel angebotene Dokumentenstrukturen zugegriffen werden kann. Die empirische Untersuchung des Leserverhaltens belegt, daß diese Gestaltung den Bedürfnissen großer Nutzergruppen entgegenkommt. Teile der in dieser Arbeit entwickelte Strukturierung und Gestaltung wurde von der *tageszeitung* in ein überarbeitetes System übernommen.

Aus der durchgeführten Leserbefragung geht hervor, daß bei der Gestaltung eines Online-Angebots nach wie vor auf die Minimierung des Datendurchsatzes geachtet werden sollte. Weiterhin sollte den Lesern die Option angeboten werden, die auf eine Benutzeraktion hin übertragene mittlere Datenmenge selbst zu wählen.

Laut Leserumfrage steht mehr als ein Drittel der Teilnehmer einem elektronischen Abonnement aufgeschlossen gegenüber. Für etwa Viertel aller Teilnehmer ist vorstellbar, ein elektronisches Abonnement zu bezahlen. Der durch diese Gruppe eingeschätzte Wert der Online-Ausgabe entspricht etwa der Hälfte des für ein Abonnement der gedruckten Ausgabe zahlbaren Betrags, d.h., in einer Größenordnung von DM 15.00. Die Aussichten zur Akquisition kommerzieller Anzeigen werden von der Marketingabteilung der *tageszeitung* als erfolgversprechend eingeschätzt. Die durch die Umfrage erhobenen Daten können als Ausgangsbasis für künftige Marketing- und Preisgestaltungsstrategien verwendet werden.

Eine Weiterentwicklung des Systems sollte sich zunächst auf die Minimierung von Benutzeraktionen konzentrieren. Bislang ist nach jedem Aufruf eines Artikels oder einer Seite ein Rücksprung zu einem Inhaltsverzeichnis notwendig, wenn weitere Beiträge oder Seiten gezielt aufgerufen werden sollen. Dieses Problem ist von viele Lesern bemängelt worden und bei allen vorgestellten Gestaltungsalternativen bisher unbefriedigend gelöst.

Eine Alternative könnte hier die Gestaltung der Gesamtinhaltsverzeichnisse als *HTML-Formular* sein. Zusätzlich zu einem direkten *link* auf den zugeordneten Artikel würde hierbei neben jeder Artikelüberschrift ein *Button* plaziert werden, der vom Benutzer beim Durchlesen des Inhaltsverzeichnisses gesetzt werden kann. Das Formular würde sukzessive ausgefüllt und schließlich an den Server übertragen. Dort würde ein Dokument zusammengestellt, das ein lokales Inhaltsverzeichnis aller gewählten Beiträge, die gewählten Beiträge selbst und Hilfen zum Navigieren zwischen Artikeln enthält. Dieses Dokument wäre gewissermaßen eine vom Benutzer zusammengestellte *Digital Me* (vgl. Kap. 2.2.2).

Vergleichsweise geringen zusätzlichen Aufwand würde ein dynamisch aktualisiertes Angebot erfordern, d.h. der Transfer von Artikeln aus dem Redaktionssystem der *tageszeitung* in das Online-Angebot in Echtzeit. Eine derartige, fortlaufende Aktualisierung kann unserer Erfahrung nach in Ausnahmesituationen, d.h., zu Zeiten hoher 'Nachrichtendichte', von Vorteil sein. (Beispiel während des Projektzeitraums: Atomversuche im Mororua-Atoll).

In Verbindung mit Mechanismen zur (anonymisierten) Benutzeridentifikation könnten bei dieser Form des Angebots sukzessive Informationsprofile aufgebaut werden. Dies würde den Aufbau vorinitialisierter Verzeichnis-Formulare erlauben, in denen bereits jene Artikel markiert wären, die den Benutzer voraussichtlich interessieren dürften. Die Zuordnung von Benutzerprofilen zu Artikeln sollte möglichst auf Grundlage der von den Archivmitarbeitern manuell erstellten Index- und Schlagwortinformationen geschehen. Die Qualität der getroffenen Auswahl dürfte in diesem Fall klar über derjenigen automatischer Textfilter liegen (vgl. Kap. 2.2).

Voraussetzung für die Verwirklichung eines derartigen Konzepts wäre, daß die Indizierung und Verschlagwortung von Artikeln bereits *während* der Produktion der Ausgabe geschieht und gleichzeitig mit (oder kurz nach) dem Bereitstellen des Online-Angebots verfügbar ist. Wie entsprechende Experimente anderer Tageszeitungen belegen, wäre hierfür vermutlich eine Reorganisation des bestehenden Archivwesens sowie eine Veränderung von Arbeitsprozessen und -zeiten nötig. Bei der in Dublin herausgegebenen *Irish Times* wird täglich eine speziell auf die Randbedingungen des Online-Angebots hin optimierte elektronische Ausgabe erzeugt. Dies erfordert mehrstündige Nacharbeit dreier qualifizierter Redakteure, deren Aufgabe u.a. darin besteht, zu jedem Artikel drei- bis fünfzeilige Infogramme zu erstellen. Diese Tätigkeit kommt derjenigen einer Verschlagwortung sehr nahe (die Infogramme werden als Unterüberschriften eingesetzt, um die Informationsauswahl der Leser zu unterstützen)¹.

Obwohl der laufende technische Betrieb eines Online-Angebots ohne nennenswerten Personalaufwand möglich ist, deuten sowohl eigene Erfahrungen mit verschiedenen hypertextgestützten Zeitungen als auch eine Vielzahl von Leserbriefen darauf hin, daß zusätzlicher *inhaltlicher* Aufwand nötig ist, um eine Online-Tageszeitung attraktiv und benutzerfreundlich zu machen. Dies soll kurz an einem Beispiel demonstriert werden.

Die bisherige Form der Inhaltsverzeichnisse in der Online-Ausgabe - ein Dachzeilenbegriff, gefolgt von einer zumeist kurzen Schlagzeile und ggf. einem einzeiligen Untertitel - orientiert sich notgedrungen an der Aufmachung der gedruckten Ausgabe. Bei der Papierversion kann ein Leser bereits nach einem kurzen Blick auf den Text entscheiden, ob der Beitrag hält, was sein Titel verspricht. Bei hierarchisch aufgebauten Online-Angeboten ist dies nicht der Fall: um eine entsprechende Entscheidung treffen zu können, muß der Leser das entsprechende Dokument zunächst anfordern.

Traditionelle Schlagzeilen und Untertitel geben in Inhaltsverzeichnissen von Online-Ausgaben zumeist nur unzureichenden Aufschluß über den Inhalt des Artikels. Abhilfe schaffen könnte hier

1. Persönliche Mitteilung von Joe Breen, Duty Editor Electronic Publishing der *Irish Times*, beim Roundtable "Managing a Digital Newspaper" des European Journalism Center vom 28-29 März in Dublin

eine zusätzliche Gliederungsstufe. d.h., ein kurzer Textblock, der in wenigen Zeilen den Inhalt des Artikels referiert. Der dazu notwendige, zusätzliche redaktionelle Aufwand (vgl. das oben erwähnte Beispiel der *Irish Times*) ist kaum zu umgehen, wenn man sowohl die Spezifika des elektronischen Mediums wie auch die Bedürfnisse der Online-Leserschaft ernstnimmt.

Dies gilt ebenso für die wünschenswerte Integration von Verweisen auf Informationen, die zu in Artikeln behandelten Themen im Internet verfügbar sind. Die Recherche derartiger Informationen sollte möglichst eng mit der Arbeit des Archivs verbunden werden. Auf diese Weise könnten einmal gefundene Quellen der Indexstruktur des Archivs entsprechend katalogisiert werden. Die Güte der Quelle (ihre 'Zitierfähigkeit') würde dabei von Redakteuren oder Archivmitarbeitern beurteilt. Verschiedene kommerzielle Internet-Indexanbieter zeigen, daß derartige Sammlungen klassifizierter und bewerteter Ressourcen bereits heute ein vermarktbare Produkt darstellen.

Die sukzessive Evaluierung des im Internet verfügbaren Quellenmaterials dürfte die Leistungsfähigkeit einzelner Zeitungsarchive überfordern. Zukünftige Perspektiven für Zeitungsverlage könnten deshalb vermehrt in Kooperationsmodellen ähnlich dem in Kap. 2.1 erwähnten Projekts *Cental Station* liegen. Denkbar wären beispielsweise *zeitungsübergreifende*, individualisierte Informationsangebote überregionaler Tages- und Wochenzeitungen in Verbindung mit den Lokalnachrichten regional verankerter Blätter. Ein gemeinsames Vorgehen wäre auch beim Aufbau von Internet-Quellenindices sinnvoll.

Viele Leserbriefe deuten darauf hin, daß elektronische Zeitungsausgaben besonders für Personen wertvoll sind, die sich außerhalb des geographischen Verbreitungsgebiets der gedruckten Zeitung befinden. Um Online-Ausgaben künftig auch innerhalb des Vertriebsgebiets für breitere Leserkreise interessant zu machen, scheinen zunächst einschneidende Verbesserungen bei den technischen Interfaces und eine radikale Senkung des Angebotspreises nötig zu sein. In den folgenden Abschnitten soll kurz auf diese beiden Determinanten eingegangen werden.

Die künftigen Entwicklungen von Online-Tageszeitungen gesetzten Grenzen sind vorrangig ökonomischer Natur. Zwar lassen sich aus Sicht der Verlage Druck- und Vertriebskosten der Ausgabe einsparen, doch die beträchtlichen fixen Kosten für Redaktion und Infrastruktur bestehen weiterhin. Sollen diese Kosten durch den Verkauf der elektronischen Ausgabe erwirtschaftet werden, müßte beispielsweise die *tageszeitung* von etwa 60.000 regelmäßiger Online-Leser täglich etwa DM 0.80-DM 1.00 einnehmen. Dies entspricht zwar nur dem halben Kiosk-Verkaufspreis, doch selbst ein Bruchteil dieses Betrages ist im Internet derzeit nicht durchsetzbar. Obwohl es in den USA Beispiele für WWW-Angebote gibt, die tägliche Konzessions- und Werbeeinnahmen in der oben angegebenen Größenordnung erzielen, sind derartige Beträge auf dem deutschsprachigen Markt mit einem Tageszeitungsangebot wohl kaum zu erreichen.

Aus Kundensicht besteht eines der Hauptprobleme zur Zeit darin, daß selbst Reduzierungen des Verkaufspreises um die oben genannten 50 Prozent bei ihm nicht ankommen. Der eingesparte Betrag wird durch hohe Investitionskosten für die nötige technische Ausstattung und laufende Gebühren für Telekommunikationsverbindungen und Dienstleister aufgezehrt. Auch wenn die komplette Ausgabe für den Preis eines Ortsgesprächs zum Kunden transferiert und für einen Pfen-

nig pro Seite ausgedruckt werden könnte, lägen die aufzubringenden Gesamtkosten (Gebühren, Material, Energie) bereits bei ca. DM 0.70. Dies entspricht der Größenordnung heutiger Druck- und Vertriebskosten der *tageszeitung*. Berücksichtigt man die real existierenden Wartezeiten und die Abschreibung der technischen Ausstattung, liegt dieser Wert sogar beträchtlich höher.

Die Qualität und die Kosten künftiger Netzinfrastruktur sind demnach wichtige Faktoren für die zukünftige Entwicklung von Online-Informationsangeboten. Dies gilt in mindestens ebensogroßem Maße für künftige Benutzerinterfaces. Eine kurzfristig verfügbare Alternative zu hochauflösenden elektronischen Displays könnte dabei das Home-Printing sein. Die graphische Auflösung in Zeitungsqualität stellt für Computerdrucker kein Problem mehr dar. Ideal für die Ausgabe von Zeitungen wäre ein doppelseitiger DIN A3 Drucker, deren Erstellungskosten im Bereich von DM 0.01 pro Druckseite liegen. Da es sich bei Zeitungen um typische Wegwerfprodukte handelt, sollte hierbei eine bei Computerdruckern längst überfällige, umweltfreundliche Technologie eingesetzt werden, nämlich wiederverwertbare Medien. Denkbar wäre der Einsatz von qualitativ hochwertigem Papier zusammen mit Druckertinte, die sich bei z.B. bei hohen Temperaturen, starker UV-Strahlung oder elektrostatische Einwirkung farblos zersetzt. Es sei angemerkt, daß für alle künftigen Szenarien eine Gesamt-Ökobilanz wünschenswert wäre. Zwar haben viele Online-Leser die von ihnen vermutete ressourcenschonende (weil papierlose) Form des Angebots begrüßt, doch der Beweis ihrer Vermutung steht noch aus.

Abschließend seien kurz die ethischen Grenzen von Online-Angeboten erwähnt. Dies betrifft weniger deren Inhalte als vielmehr den Umgang mit Benutzerdaten. Angesichts unklarer Finanzierungsmöglichkeiten sehen einige Zeitungsverleger eine künftige Verdienstquelle im Erstellen von Kunden-Informationsprofilen, die z.B. für Direktmarketing eingesetzt werden können. Die Verantwortlichen sollten sich vergegenwärtigen, daß die Aufgabe und Existenzgrundlage von Presseorganen nicht darin besteht, Information *über* ihre Leser zu sammeln, sondern darin, eben dies *für* ihre Leser zu tun.

9.0 Quellen

- Bae85 Baer, Walter: Defining a new Business. In: Electronic Publishing Plus, White Plains, N.Y., 1985
- Bal94 Balasubramanian, V.: State of the Art Review on Hypermedia Issues And Applications. Graduate School of Management, Rutgers University, Newark, New Jersey, März 1994. URL: http://www.isg.sfu.ca/~duchier/misc/hypertext_review/chapter1.html
- Beg96 Berghel, H.: "The Client's Side of the World Wide Web". In: Communications of the ACM Vol. 39, No. 1, Januar 1996, S.34
- Ber01 Berners-Lee, Tim: WorldWideWeb-Summary. Ohne Datum. URL: <http://www.w3.org/hypertext/WWW/Summary.html>
- Ber02 Berners-Lee, T.: Navigational Techniques and Tools. Ohne Datum. URL: <http://www.w3.org/hypertext/WWW/DesignIssues/Navigation.html>
- Ber91 Berners-Lee, Tim.: Design Constraints. 1991 URL: <http://www.w3.org/hypertext/WWW/MarkUp/HTMLConstraints.html>
- Ber94 Berners-Lee, T. et al: Uniform Resource Locators (URL). RFC 1738, Dez. 1994. URL: <http://ds.internic.net/rfc/rfc1738.txt>
- Ber95a Berners-Lee, Tim et al: Hypertext Transfer Protocol -- HTTP/1.0. Internet Draft, Okt. 1995. URL: <ftp://ietf.cnri.reston.va.us/internet-drafts/draft-ietf-http-v10-spec-04.txt>
- Ber95b Berners-Lee,T; MIT/W3C; Conolly, D.: Hypertext Markup Language - 2.0. RFC 1866, November 1995. URL: <http://www.cis.ohio-state.edu/htbin/rfc/rfc1866.html>
- Ble95 Bleuel, Jens: Online Publizieren im Internet. Pfungstadt und Bernsheim, 1995
- Bre91 Breilinger-Gusbeth: Bildschirmtext im Zeitungswesen. In: Breilinger-Gusbeth: Bildschirmtext und seine Anwendung. Loseblattsammlung, RS Schulz. 1991
- Brö95 Brössler, Daniel: Zeitung und Multimedia. München 1995.
- Cai95 Cailliau, Robert: A Little History. Oktober 1995. URL: <http://www.w3.org/hypertext/WWW/History.html>
- Eit95 Enterprise Integration Technologies (EIT): Hypermail. Okt. 1995 URL: <http://www.eit.com/software/hypermail/>
- Eis94 Eisenhart, Douglas M.: Publishing in the Information Age. Westport, CT, 1994.
- Fie95 Fielding, R.: Relative Uniform Resource Locators. RFC 1808, Juni 1995. URL: <http://ds.internic.net/rfc/rfc1808.txt>

- Fry95 Frystyk, Henrik: W3C Reference Library. Dezember 1995.
URL: <http://www.w3.org/pub/WWW/Library>
- Gra95 Graham, I.S.: HTML Sourcebook. New York, 1995.
- Gro60 Groth, Otto: Die unerkannte Kulturmacht. Berlin 1960.
- Har96 Hardy, D.R.: "Running a Cache hierarchy". In: Harvest User's Manual". Boulder, Colorado, Januar 1996.
URL: <http://harvest.cs.colorado.edu/harvest/user-manual/node121.html>
- Hed95 Hedlund, M.: The Common Gateway Interface (CGI) Frequently Asked Questions (FAQ) List - Implementation Issues. Oktober 1995.
URL: <http://www.best.com/~hedlund/cgi-faq/new/faq.7-implement.html>
- Ics94 ICS (Department of Information and Computer Science, University of California, Irvine): wwwstat -- Distribution Information. Jul. 1994.
URL: <http://www.ics.uci.edu/WebSoft/wwwstat/>
- Iso86 ISO 8879. Information Processing -- Text and Office Systems - Standard Generalized Markup Language (SGML), 1986. URL: <http://www.iso.ch/cate/d16387.html>
- Joh93 Johns, M. S.: Identification Protocol. RFC 1413, Februar 1993.
URL: <http://ds.internic.net/rfc/rfc1413.tx>
- Kat94 Katz, Jon: Online or not, Newspaper suck. In: Wired XXX
- Kuhn96 Kuhn, Thomas: Die Struktur wissenschaftlicher Revolutionen. Ffm 1996
- Kuh91 Kuhlen, Rainer: Hypertext. Berlin, 1991
- Mit01 MIT Media Lab. URL: <http://nif.www.media.mit.edu/abs.ht>
- Nau94 Naumann, Arthur C.: Es wird viele Bees geben. In: Zeitung der Zukunft. Hg. Bundeszentrale für politische Bildung, Bonn 1994, S.2
- Noe89 Noelle-Neumann u.a. (Hrsg.): Fischer Lexikon Publizistik Massenkommunikation. Frankfurt 1981
- Pri96 Price, Roger (Ed.): Proposed ISO/IEC International Standard for HTML. Januar 1996.
URL: <http://www.w3.org/hypertext/WWW/MarkUp/mheg-2.ps>
- Rie95 Riefler, Katja: Zeitung Online. Bonn 1995.
- Tur96 Turner, Stephen: The analog program. University of Cambridge Statistical Laboratory, 1996. URL: <http://www.statslab.cam.ac.uk/~sret1/analog/>

Spi01 Der Spiegel. Ohne Datum. URL:

[http://www.hamburg.pop.de:800/bda/int/spon/online/meta.html#Zeitungen und Magazine](http://www.hamburg.pop.de:800/bda/int/spon/online/meta.html#Zeitungen%20und%20Magazine)

Wwc01 World Wide Web Consortium: Hypertext Terms. April 1995. URL:

<http://www.w3.org/hypertext/WWW/Terms.html>

Wwc02 World Wide Web Consortium: HTML Publication History. Ohne Datum.

URL: <http://www.w3.org/hypertext/WWW/MarkUp/HTML.html>

Anhang A: BNF-Grammatik des *taz*-Archivformats

BNF-Grammatik des *taz*-Archivformats

Zur Darstellung von Sonderzeichen wurde die in der Programmiersprache “C” gebräuchliche Notation gewählt (z.B. ‘\n’ für Zeilenumbruch). Einzelne Zeichen werden von Anführungszeichen eingeschlossen (z.B. ‘ä’). Für Alternativen zwischen fortlaufenden *ISO-Latin-8859-1* Symbolen wird eine verkürzende Notation verwendet (die Semantik der Notation ‘[A..Z]’ lautet umgangssprachlich: ‘Ein und nur ein großer Buchstabe des lateinischen Alphabets’).

Diese empirisch ermittelte Grammatik erhebt nicht den Anspruch, das tatsächliche Archivformat der *tageszeitung* vollständig und korrekt zu beschreiben. Sie war jedoch ausreichend, um jede während des Projektzeitraums bearbeitete Archivdatei fehlerfrei zu parsen und ohne manuelle Korrekturen in HTML-formatierte Dokumente umzusetzen.

```
W_SP      ::= W_SP | WHITESPACE
WHITESPACE ::= CR | TAB | SPACE
CR        ::= '\n'
TAB       ::= '\t'
SPACE     ::= ' '
DIGITS    ::= DIGITS | DIGIT
DIGIT     ::= [0..9]
ALL_CHARACTERS ::= CHARACTER | DIGIT | SPECIALS | WHITESPACE
CHARACTER ::= [A..Z] | [a..z] | 'ä' | 'ö' | 'ü' | 'Ä' | 'Ö' | 'Ü' | 'ß' |
SPECIALS  ::= '\\ ' | '\ ' | '! ' | '@ ' | '#' ' | '$ ' | '%' ' | '^ ' | '& ' | '*' ' |
            '(' ' | ')' ' | '-' ' | '_' ' | '=' ' | '+' ' | ';' ' | ':' ' | '"' ' | '\'' ' |
            '<' ' | '>' ' | '/' ' | '?' ' | '\\ ' | '\ ' | '[' ' | ']'
LINESTRING ::= STRING | CHARACTER | TAB | SPACE

FILE      ::= ARTIKEL_LISTE
ARTIKEL_LISTE ::= W_SP ARTIKEL ARTIKEL_LISTE | W_SP ARTIKEL W_SP
ARTIKEL   ::= "{" LAEN_FELD CHSET_FELD IDENT_FELD QUELLE_FELD NR_FELD
            SEITE_FELD DATUM_FELD ZEILEN_FELD RESS_FELD ART_FELD BEM_FELD
            AUTOR_FELD KLASS_FELD TITEL_FELD BROT_FELD W_SP "}" CR
LAEN_FELD ::= P_LAENGE | VOID
P_LAENGE  ::= W_SP "{" "Laen" LAEN_ID W_SP "}"
LAEN_ID   ::= "lang" | "mittel" | "kurz"
CHSET_FELD ::= P_CHSET | VOID
P_CHSET   ::= W_SP "{" "ChSet" W_SP CHSET_TYPE W_SP "}"
CHSET_TYPE ::= "ISO-Latin-8859-1"
IDENT_FELD ::= P_IDENT | VOID
P_IDENT   ::= W_SP "{" "Ident" W_SP LINESTRING "}"
QUELLE_FELD ::= P_QUELLE | VOID
P_QUELLE  ::= W_SP "{" "Quelle" W_SP LINESTRING "}"
NR_FELD   ::= P_NR | VOID
P_NR      ::= W_SP "{" "Nr" W_SP DIGITS W_SP "}"
SEITE_FELD ::= P_SEITE | VOID
P_SEITE   ::= W_SP "{" "Seite" W_SP DIGITS W_SP "}"
DATUM_FELD ::= P_DATUM | VOID
P_DATUM   ::= W_SP "{" "Datum" W_SP LINESTRING "}"
ZEILEN_FELD ::= P_ZEILEN | VOID
P_ZEILEN  ::= W_SP "{" "Zeilen" W_SP DIGITS W_SP "}"
RESS_FELD ::= P_RESS | VOID
P_RESS    ::= W_SP "{" "Ress" W_SP LINESTRING "}"
ART_FELD  ::= P_ART | VOID
```

```

P_ART      ::= W_SP "{" "Art" W_SP LINESTRING }"
BEM_FELD  ::= P_BEM | VOID
P_BEM     ::= W_SP "{" "Bem: W_SP LINESTRING }"
AUTOR_FELD ::= P_AUTOR | VOID
P_AUTOR   ::= W_SP "{" "Autor" W_SP LINESTRING }"
KLASS_FELD ::= P_KLASS | VOID
P_KLASS   ::= W_SP "{" "Klass" W_SP LINESTRING }"
TITEL_FELD ::= P_TITEL | VOID
P_TITEL   ::= W_SP "{" "Titel" W_SP A_FORMAT_TEXT }"
A_FORMAT_TEXT ::= A_FORMAT_ITEM A_FORMAT_TEXT | A_FORMAT_ITEM
A_FORMAT_ITEM ::= A_FORMAT | TEXT
A_FORMAT     ::= W_SP "{" A_FORMAT_ID W_SP TEXT }"
TEXT        ::= TEXT_CHARS | VOID
TEXT_CHARS  ::= ALL_CHARACTERS TEXT_CHARS | ALL_CHARACTERS
A_FORMAT_ID ::= "A_Vorspann" | "A_Dachzeile" | "A_Ueberschrift" |
               "A_Unterzeile"
BROT_FELD  ::= P_BROT | VOID
P_BROT     ::= W_SP "{" W_SP PAR_TEXT }"
Z_FORMAT_TEXT ::= Z_FORMAT_TEXT Z_FORMAT_ITEM | Z_FORMAT_ITEM
Z_FORMAT_ITEM ::= Z_FORMAT | TEXT
Z_FORMAT     ::= W_SP "{" Z_FORMAT_ID W_SP TEXT }"
Z_FORMAT_ID  ::= "A_Zwischen" | "Z_Autor" | "Z_Korres" | "Z_Fett" |
               "Z_Kursiv" | "Z_Fett_Kursiv" | "Z_Fuss" | "Z_Initial"
VOID        ::= W_SP | ""

```

Beispiel: Artikel im Archivformat der *tageszeitung*

```

{Artikel
  {ChSet ISO-Latin-8859-1}
  {Laen mittel}
  {Ident T951127.4}
  {Quelle TAZ}
  {Nr 4784}
  {Seite 1}
  {Datum 27.11.1995}
  {Zeilen 72}
  {Ress Berlin}
  {Autor Karl Gersuny}
  {Titel
    {A_Ueberschrift
      In Zentralbosnien brennen die Häuser}
    {A_Unterzeile
      Kroatische Soldaten plündern Dörfer und Städte, die nach dem
      Friedensvertrag von Dayton an die Serben abgegeben werden müssen.
      Tausende Serben demonstrieren für ein "serbisches Sarajevo". Karadzic
      droht der Nato}
    }
  {Brot
    {Z_Fett Wien (taz) - }Nur wenige Tage nach der Paraphierung des
    Friedensabkommens für Bosnien wird dieses erneut in Frage gestellt. So
    forderte der Führer der bosnischen Serben Radovan Karadzic neue Verhandlungen
    über den Status von Sarajevo. In den serbischen Teilen der bosnische
    Hauptstadt demonstrierten Tausende für ein "serbisches Sarajevo".
    {A_Zwischen Ausschreitungen auf kroatischer Seite}
    Die schwersten Zwischenfälle ereigneten sich nach UNO-Angaben aus Sarajevo
    sowie Berichten der BBC südlich von {Z_Kursiv Banja Luka}, wo kroatische Verbände
    in den Städten Mrkonjic Grad und Sipovo ganze Straßenzüge in Brand steckten. }
  }
}

```

Anhang B: Liste der Systemmodule

Entpacken der Archivdatei, Aufruf des Formatkonvertierers, des Strukturgenerators und des Authentifikationsgenerators:

tazmake.sh Bourne-Shell-Skript

Formatkonversion:

Makefile
parser.c C-Sourcdatei
parser.h C-Sourcdatei
action.c C-Sourcdatei
action.h C-Sourcdatei
tazbuild kompilierte Binärdatei

Strukturgenerator:

tazbuild.sh Bourne-Shell-Skript

Authentifizierung:

tazauth.sh Bourne-Shell-Skript

CGI-Skripte

Stichwortsuche

taz-suche Bourne-Shell-Skript

Leserforum

readmail Bourne-Shell-Skript
Hypermail 1.02 Sourcedistribution, angepaßt wurde das Modul print.c
hypermail kompilierte Binärdatei der Distribution hypermail 1.02

Leserbrief

Hypermail 1.02 Sourcedistribution, angepaßt wurde das Modul mail.c
hypermailer kompilierte Binärdatei der Distribution hypermail 1.02

Fragebogen

Taz-Fragebogen.pl Perl-Skript

Programmierungsumgebung

gcc 2.6.3
Perl 5.0
bourne shell mit Standardhilfsprogrammen (sed, grep, awk, expr)

Anhang C: Wortlaut der Leserumfrage

digiTaz -- Ihre Meinung ist gefragt!

Hallo!

Sie sind im Begriff die digitale Ausgabe der *tageszeitung* zu lesen. Wir, die PRZ¹ der TU Berlin, bieten mit Unterstützung der DeTeBerkom² und in Zusammenarbeit mit der *tageszeitung* seit kurzem diesen Dienst inoffiziell an.

Am 12. Mai soll im Rahmen einer Pressekonferenz die digiTaz offiziell vorgestellt werden. Dazu hätten wir einige Fragen an Sie. Wir würden uns freuen, wenn Sie sich die Zeit nehmen und diese beantworten: Nur mit Ihrer Hilfe können wir besser werden. Alle Angaben werden anonym behandelt, siehe auch die Hinweise am Ende dieser Seite.

- Was bedeutet es für Sie die taz online im Internet zu lesen?
(Freitextfeld)
- Wo lesen Sie die digiTaz?
(Nur eine Antwort möglich)
 - Während der Arbeit
 - Zu Hause
- Von welchem Land aus lesen Sie die digiTaz?
(Freitextfeld, voreingestellt war "Deutschland")
- Wie oft lesen Sie die digiTaz?
(Nur eine Antwort möglich)
 - Ist das erste Mal
 - gelegentlich
 - täglich
- Haben Sie vor, die taz zukünftig noch öfter auf diesem Weg zu lesen?
(Nur eine Antwort möglich)
 - Ja
 - Nein
- weil ...
(Freitextfeld)
- Was für eine Organisation stellt Ihnen den Zugriff auf das Internet zur Verfügung?
(Nur eine Antwort möglich)
 - Ein privater kommerzieller Anbieter
 - Ihre Universität
 - Ihr Arbeitgeber
 - Eine Firma
- Wie bewerten Sie diese neue Leseerfahrung digiTaz?
(Freitextfeld)
- Was gefällt Ihnen an der digiTaz besonders gut?
(Nur eine Antwort möglich)
 - Die Möglichkeit der gezielten Suche
 - Das frühe Erscheinen
 - Die Bilder
 - Was anderes... (Freitextfeld)

1. Prozeßrechnerverbund-Zentrale
2. Deutsche Telekom: Berliner Kommunikationssysteme

- Was mißfällt Ihnen an der digiTaz besonders?
 (Nur eine Antwort möglich)
 - Die gewöhnungsbedürftige Handhabung
 - Die geringe Übertragungsgeschwindigkeit (lange Wartezeiten)
 - Was anderes... (Freitextfeld)

- Wie haben Sie von der digiTaz erfahren?
 (Freitextfeld)

- Wären Sie bereit fuer diese Form der taz Geld auszugeben?
 (Nur eine Antwort möglich)
 - Ja
 - Nein
 weil...
 (Freitextfeld)

- Könnten Sie sich vorstellen, die digiTaz zu abonnieren, wenn sie täglich erscheint?
 (Nur eine Antwort möglich)
 - Ja
 - Nein
 weil...
 (Freitextfeld)

- Was wäre Ihnen die digiTaz im Monat wert?
 Etwa DM
 (Freitextfeld)

Hinweise

Anonymität

Die Teilnahme an dieser Umfrage ist freiwillig und verpflichtet Sie zu nichts. Alle von Ihnen gemachten Angaben kommen bei uns anonymisiert an - wir wissen nicht wer Sie sind, wir speichern keinerlei persönliche Daten über Sie (weil wir sie nicht haben), wir nehmen uns nur Ihrer Kommentare an und versuchen, Ihre Hinweise und Verbesserungsvorschläge aufzugreifen. Wenn Sie den Dialog mit uns suchen, schreiben Sie bitte an eine der unten genannten Mail-Adressen.

Wie Sie uns erreichen können

Die digiTaz

Bei Fragen oder Problemen, die die Web-Ausgabe der tageszeitung "digiTaz" betreffen, z.B. Probleme mit der Darstellung unter Ihrem WWW-Browser, wenden Sie sich bitte per E-Mail an taz@prz.tu-berlin.de.