

Thomas Rusch

**Gegenüberstellung von HTML-Dialekten
zur Analyse und Konvertierung von
Informationen für DAB**

Diplomarbeit

BEI GRIN MACHT SICH IHR WISSEN BEZAHLT



- Wir veröffentlichen Ihre Hausarbeit, Bachelor- und Masterarbeit
- Ihr eigenes eBook und Buch - weltweit in allen wichtigen Shops
- Verdienen Sie an jedem Verkauf

Jetzt bei www.GRIN.com hochladen
und kostenlos publizieren



**Hochschule für
Technik und Wirtschaft
Mittweida (FH)**

Diplomarbeit

Rusch, Thomas

„Gegenüberstellung von HTML-Dialekten zur Analyse und Konvertierung von
Informationen für DAB“

1998

**Gegenüberstellung von HTML-Dialekten zur Analyse und Konvertierung
von Informationen für DAB**

Diplomarbeit

Hochschule für

Technik und Wirtschaft Mittweida (FH)

Fachbereich Mathematik / Physik / Informatik

Fachgruppe Informatik

eingereicht von: Thomas Rusch

Mittweida, den 26. Juni 1998

Gegenüberstellung von HTML-Dialekten zur Analyse und Konvertierung von Informationen für DAB

Hochschule für Technik und Wirtschaft Mittweida (FH),
Fachbereich Mathematik / Physik / Informatik,
Fachgruppe Informatik,
Diplomarbeit, 1998

Referat:

Die vorliegende Diplomarbeit beschäftigt sich mit der Analyse des HTML-Dialektes für das „Digital Audio Broadcasting“ (DAB) und mit der Ausarbeitung von alternativen Darstellungen, um HTML-Dokumente (Standard 3.2) mit möglichst wenig Informationsverlust im DAB darstellen zu können. Dazu ist ein Konvertierprogramm entwickelt worden, das diese Aufgabe übernimmt.

Inhaltsverzeichnis

1	Einführung	7
1.1	Veranlassung	7
1.2	Aufgabenstellung.....	7
1.3	Vorgehensweise.....	7
2	Überblick DAB.....	9
2.1.	Entwicklung und Geschichte.....	9
2.2.	Technische Grundlagen.....	9
2.3.	Endgeräte.....	11
3	Analyse des DAB – HTML – Dialektes.....	12
3.1	Vorbetrachtung	12
3.2	Beschreibung DAB - HTML – Tags	13
3.2.1	Kopfteil	13
3.2.2	Inhaltsteil	13
3.2.3	Überschriften	15
3.2.4	Zeichenketten.....	17
3.2.5	Textauszeichnung	18
3.2.6	Ausrichtung.....	21
3.2.7	Listenstrukturen	21
3.2.8	Links und Anker	23
3.2.9	Grafische Elemente.....	25
3.3	Überprüfung der Syntax einer HTML-Datei.....	25
3.3.1	Grundlagen.....	25
3.3.2	Syntaxgraph	26
3.3.3	Verwendete Klassen und deren Hierarchie.....	26
3.3.4	Struktogramm für Programmablauf.....	27
3.3.5	Programmumsetzung in C++.....	27
4	Alternative Darstellungsformen	29
4.1	Überblick nicht zugelassener Tags im DAB	29
4.2	Tabellen	30
4.2.1	Darstellung durch alternative Tags	30
4.2.2	Darstellung in einer Grafik	32

4.2.3	Darstellung mit Screenshot.....	33
4.3	Anker	33
4.4	Grafiken.....	36
4.4.1	Einfache Bilder	36
4.4.2	Image-Maps	37
4.5	Formulare	39
4.6	Indexe	39
4.7	Einfache Ersetzung mit anderen Tags	40
4.7.1	Große Schrift mit <BIG> und </BIG>	40
4.7.2	Kleine Schrift mit <SMALL> und </SMALL>	40
4.7.3	Nichtproportionale Schriftart.....	40
4.7.4	Abbildungsüberschrift mit <CAPTION> und </CAPTION>	41
4.7.5	Absätze mit <DIV> und </DIV>.....	41
4.8	Einfügen von Zeichen anstelle von Tags.....	41
4.8.1	Hoch- bzw. tiefgestellter Text	41
4.8.2	Durchgestrichener Text mit <STRIKE> und </STRIKE>	41
4.9	Nicht konvertierte Tags	41
4.10	Frames	42
5	Zusammenfassung	44
A	Vergleich HTML 3.2 – HTML-Subset für DAB.....	45
B	Beschreibung des Konvertierprogramms	46
B.1	Oberfläche	46
B.1.1	Optionen.....	46
B.1.2	Programmumgebung.....	47
B.2	Menüs	48
B.2.1	File	48
B.2.2	Edit.....	48
B.2.3	Tools	48
B.2.4	Help.....	49
B.3	Funktionsweise.....	49
C	Beschreibung Klassen, Funktionen und Variablen (Referenz).....	50

Abkürzungsverzeichnis.....	83
Quellenverzeichnis.....	84
Thesen.....	85

1 Einführung

1.1 Veranlassung

Der Anlaß für diese Diplomarbeit ist ein „Forschungsprojekt DAB“. Es beschäftigt sich mit der Weiterentwicklung der Möglichkeiten im DAB sowie mit deren Realisierung. Dabei sollen Probleme erkannt, nach Lösungen gesucht und diese umgesetzt werden.

1.2 Aufgabenstellung

Seiten, die in HTML-Standards bis einschließlich 3.2 erstellt wurden, sollen für die Darstellung im DAB nutzbar gemacht werden, damit diese auf Geräten früherer Generationen angezeigt werden können.

Man unterscheidet dabei verschiedene Empfängertypen (siehe auch Punkt 2.3):

- Typ B (Geräte ohne PC-Unterstützung),
- Typ C (Geräte mit PC-Unterstützung und Massenspeicher) und
- Typ D (PC-Einsteckkarten).

Die Typen B und C werden dabei hauptsächlich in Kraftfahrzeugen als mobile Stationen eingesetzt. Ich beschränke mich dabei auf die Umsetzung für die Typ-C-Geräte, da diese ausschließlich im Pilotprojekt des Freistaates Sachsen vorkommen und es sich dabei um die bisher am weitest entwickelten Geräte für den mobilen Einsatz handelt.

Für die Typ-D-Geräte gibt es keinerlei Beschränkungen, da die Darstellbarkeit durch den dazugehörigen HTML-Browser auf dem PC gewährleistet ist.

1.3 Vorgehensweise

Zunächst beschäftige ich mich mit der Analyse des DAB-HTML-Dialektes. Dieser Dialekt entspricht der HTML-Spezifikation 2.0 mit einigen Ausnahmen. Es wird ein Graph erstellt, der die Syntax dieses Dialektes veranschaulicht. Anhand dieses Graphen wird die Syntaxprüfung von HTML-Dateien vorgenommen. Dabei werden alle Tags der Quelldatei identifiziert. Syntaxfehler werden offenbar.

Zur Konvertierung der HTML-Dateien ist es notwendig, die Tags zu interpretieren. Nicht erlaubte Tags werden protokolliert. Erlaubte Tags werden beibehalten. Bestimmte Anweisungen können in Alternativdarstellungen konvertiert werden.