

Oliver Zimmermann

XML als neuer Standard der digitalen Markierungssprachen

Grundlagen, Syntax und Praxis

Studienarbeit

Bibliografische Information der Deutschen Nationalbibliothek:

Bibliografische Information der Deutschen Nationalbibliothek: Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de/> abrufbar.

Dieses Werk sowie alle darin enthaltenen einzelnen Beiträge und Abbildungen sind urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsschutz zugelassen ist, bedarf der vorherigen Zustimmung des Verlanges. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen, Auswertungen durch Datenbanken und für die Einspeicherung und Verarbeitung in elektronische Systeme. Alle Rechte, auch die des auszugsweisen Nachdrucks, der fotomechanischen Wiedergabe (einschließlich Mikrokopie) sowie der Auswertung durch Datenbanken oder ähnliche Einrichtungen, vorbehalten.

Copyright © 2000 Diplomica Verlag GmbH
ISBN: 9783832425135

Oliver Zimmermann

XML als neuer Standard der digitalen Markierungssprachen

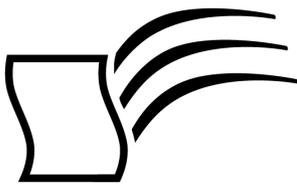
Grundlagen, Syntax und Praxis

Oliver Zimmermann

XML als neuer Standard der digitalen Markierungssprachen

Grundlagen, Syntax und Praxis

Hausarbeit
an der Fachhochschule Hamburg
Fachbereich Bibliothek und Information
3 Monate Bearbeitungsdauer
Mai 2000 Abgabe



Diplomarbeiten Agentur

Dipl. Kfm. Dipl. Hdl. Björn Bedey
Dipl. Wi.-Ing. Martin Haschke
und Guido Meyer GbR

Hermannstal 119 k
22119 Hamburg

agentur@diplom.de
www.diplom.de

ID 2513

Zimmermann, Oliver: XML als neuer Standard der digitalen Markierungssprachen: Grundlagen, Syntax und Praxis / Oliver Zimmermann ·

Hamburg: Diplomarbeiten Agentur, 2000

Zugl.: Hamburg, Fachhochschule, Hausarbeit, 2000

Dieses Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere die der Übersetzung, des Nachdrucks, des Vortrags, der Entnahme von Abbildungen und Tabellen, der Funksendung, der Mikroverfilmung oder der Vervielfältigung auf anderen Wegen und der Speicherung in Datenverarbeitungsanlagen, bleiben, auch bei nur auszugsweiser Verwertung, vorbehalten. Eine Vervielfältigung dieses Werkes oder von Teilen dieses Werkes ist auch im Einzelfall nur in den Grenzen der gesetzlichen Bestimmungen des Urheberrechtsgesetzes der Bundesrepublik Deutschland in der jeweils geltenden Fassung zulässig. Sie ist grundsätzlich vergütungspflichtig. Zuwiderhandlungen unterliegen den Strafbestimmungen des Urheberrechtes.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, daß solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Die Informationen in diesem Werk wurden mit Sorgfalt erarbeitet. Dennoch können Fehler nicht vollständig ausgeschlossen werden, und die Diplomarbeiten Agentur, die Autoren oder Übersetzer übernehmen keine juristische Verantwortung oder irgendeine Haftung für evtl. verbliebene fehlerhafte Angaben und deren Folgen.

Dipl. Kfm. Dipl. Hdl. Björn Bedey, Dipl. Wi.-Ing. Martin Haschke & Guido Meyer GbR

Diplomarbeiten Agentur, <http://www.diplom.de>, Hamburg 2000

Printed in Germany



Diplomarbeiten Agentur

Wissensquellen gewinnbringend nutzen

Qualität, Praxisrelevanz und Aktualität zeichnen unsere Studien aus. Wir bieten Ihnen im Auftrag unserer Autorinnen und Autoren Wirtschaftsstudien und wissenschaftliche Abschlussarbeiten – Dissertationen, Diplomarbeiten, Magisterarbeiten, Staatsexamensarbeiten und Studienarbeiten zum Kauf. Sie wurden an deutschen Universitäten, Fachhochschulen, Akademien oder vergleichbaren Institutionen der Europäischen Union geschrieben. Der Notendurchschnitt liegt bei 1,5.

Wettbewerbsvorteile verschaffen – Vergleichen Sie den Preis unserer Studien mit den Honoraren externer Berater. Um dieses Wissen selbst zusammenzutragen, müssten Sie viel Zeit und Geld aufbringen.

<http://www.diplom.de> bietet Ihnen unser vollständiges Lieferprogramm mit mehreren tausend Studien im Internet. Neben dem Online-Katalog und der Online-Suchmaschine für Ihre Recherche steht Ihnen auch eine Online-Bestellfunktion zur Verfügung. Inhaltliche Zusammenfassungen und Inhaltsverzeichnisse zu jeder Studie sind im Internet einsehbar.

Individueller Service – Gerne senden wir Ihnen auch unseren Papierkatalog zu. Bitte fordern Sie Ihr individuelles Exemplar bei uns an. Für Fragen, Anregungen und individuelle Anfragen stehen wir Ihnen gerne zur Verfügung. Wir freuen uns auf eine gute Zusammenarbeit

Ihr Team der *Diplomarbeiten Agentur*

Dipl. Kfm. Dipl. Hdl. Björn Bedey —
Dipl. Wi.-Ing. Martin Haschke —
und Guido Meyer GbR —

Hermannstal 119 k —
22119 Hamburg —

Fon: 040 / 655 99 20 —
Fax: 040 / 655 99 222 —

agentur@diplom.de —
www.diplom.de —

Inhaltsverzeichnis

Abbildungen, Quelltexte, Zusammenfassungen.....	5
Abstract.....	6
Ziel der Diplomarbeit.....	7
1. Abschnitt: Die Entwicklung von SGML bis XML im Überblick..	8
1.0. Einleitung.....	8
1.1. Von SGML zu XML – Eine Markierungssprache entwickelt sich...	9
1.1.1. Die Grundideen von William Tunnicliffe und Stanley Rice.....	9
1.1.2. SGML wird ISO-Norm.....	9
1.1.3. HTML und die Geburt des WWW.....	10
1.1.4. Das WWW braucht neue Entwicklungen – XML entsteht.....	12
1.2. Markierungssprachen allgemein.....	14
1.2.1. Digitale Textverarbeitung – aber bitte mit Struktur.....	14
1.2.2. Anforderungen an eine Markierungssprache.....	15
1.3. Philosophie und Ziele von XML.....	17
1.4. Merkmale, Vorteile und Nachteile von XML.....	20
2. Abschnitt: Die XML-Syntax.....	25
2.0. Erste Berührung mit einem XML-Quelltext.....	25
2.1. Der grobe Aufbau eines XML-Quelltextes.....	27
2.2. Die erfolgreichsten Filme 1999 – ein konkretes Beispieldokument	30
2.2.1. Allgemeine Vorüberlegungen.....	30
2.2.2. Strukturelle Vorüberlegungen für die DTD.....	32
2.2.3. Die DTD für die Hitliste wird erstellt.....	33
2.2.3.1. Intern oder extern, das ist hier die Frage.....	34
2.2.3.2. Eingabe der Daten.....	34
2.2.3.3. Die Entities.....	36
2.2.3.4. Die Elemente.....	38
2.2.3.5. Die Attribute.....	42
2.2.4. Der Datenbereich der Hitliste.....	45

2.2.5.	Darstellung der Hitliste.....	48
2.2.5.1.	Darstellung als reines XML-Dokument.....	48
2.2.5.2.	Darstellung mit CSS.....	50
2.2.5.3.	Darstellung mit XSL.....	53
2.3.	Vorteile der Struktur von XML gegenüber HTML.....	58
2.4.	Ausblick – sonstige Arbeit an XML.....	60
2.4.1.	Externe Verweise mit XLink.....	61
2.4.2.	Interne Verweise mit XPointer.....	62
2.4.3.	Konfliktbeseitigung durch Namensräume.....	63
2.4.4.	Dokumente als Baumstrukturen interpretieren mit DOM.....	64
3.	Abschnitt: XML in der Praxis.....	65
3.0.	Einführung.....	65
3.1.	XML als systemunabhängige Markierungssprache.....	67
3.1.1.	Mobil ins Internet durch WAP.....	67
3.1.2.	Die CinemaxX AG im Überblick	68
3.1.3.	Kinokarten per Mobiltelefon.....	70
3.1.4.	Darstellungsmöglichkeiten von WAP-Angeboten.....	74
3.2.	XML als Metasprache.....	76
3.2.1.	Das International Press Telecommunications Council.....	76
3.2.2.	Die Grundidee von NewsML.....	78
3.2.3.	Funktion und erster Entwurf von NewsML.....	80
3.2.4.	Einsatzmöglichkeiten.....	83
3.3.	XML als Datenbasis.....	85
3.3.1.	Die ProSieben-Gruppe im Überblick.....	87
3.3.2.	Geschäftslösungen mit XML als Datenbasis.....	89
4.	Abschnitt: Schlusswort.....	93
Anhang A:	Glossar/Abkürzungen.....	95
Anhang B:	Literaturverzeichnis.....	103
Eidesstattliche	Versicherung.....	110

Abbildungen, Quelltexte, Zusammenfassungen

Abbildungen:

Abbildung 1.1.4. Beziehung zwischen SGML, XML und XHTML.....	13
Abbildung 1.2. Die drei Bestandteile eines Dokumentes.....	14
Abbildung 2.2.2.: Baumstruktur der Hitliste nach den Vorüberlegungen.....	33
Abbildung 2.2.5.1.: Auszeichnung als reines XML-Dokument.....	48
Abbildung 2.2.5.3.: Darstellung der Hitliste mit XSL.....	56
Abbildung 3.1.3.: Drei Ebenen des CinemaxX WAP-Angebotes.....	70
Abbildung 3.2.: Drei Wege vom Internet-Server ins Mobiltelefon.....	75
Abbildung 3.5.: XML als Datenbasis.....	89

Quelltexte:

Quelltext 1: Einfaches HTML – Dokument.....	25
Quelltext 2: Einfaches XML – Dokument.....	26
Quelltext 3: Die DTD der Hitliste.....	35
Quelltext 4: Ein exemplarischer Datensatz aus der Hitliste.....	46
Quelltext 5: Beispielhaftes XSL-Dokument.....	54
Quelltext 6: Startseite des CinemaxX WAP-Angebotes.....	71
Quelltext 7: Auszug aus der Datei „center.asp“.....	74
Quelltext 8: Eine kategorisierte Nachricht mit NewsML.....	81
Quelltext 9: Löschung eines einzelnen Datensatzes.....	82
Quelltext 10: Beispiel für eine XML-Datei.....	91

Zusammenfassungen:

Zusammenfassung 1.2.2.: Anforderungen an eine Markierungssprache.....	16
Zusammenfassung 1.3.: Ziele von XML.....	19
Zusammenfassung 1.4.: Vorteile und Nachteile von XML.....	24
Zusammenfassung 2.1.: Bestandteile eines XML – Dokumentes.....	29
Zusammenfassung 2.2.3.4.1.: Übersicht der Definitionen in einer DTD.....	40
Zusammenfassung 2.2.3.4.2.: Mögliche Inhalte eines Elementes.....	41
Zusammenfassung 2.2.3.5.: Schlüsselwörter für Attribute.....	44
Zusammenfassung 3.3.: Vorteile und Nachteile vom neuen EDI.....	86

Abstract

Das Internet vollzieht eine rasante Entwicklung. Die Markierungssprache HTML, durch die das WWW seinen Siegeszug erst starten konnte, wird diesen Entwicklungen, vor allem in kommerziell genutzten Bereichen, längst nicht mehr gerecht. Eine neue Technik muss her. Seit 1996 arbeitet das W3-Consortium deshalb an der *Extensible Markup Language* (XML). Als Metasprache entwickelt, ist es mit XML möglich, neue, an individuelle Bedürfnisse angepasste Web-Sprachen zu generieren. Die vorliegende Arbeit soll aufzeigen, welches Potenzial in dieser neuen Sprache steckt. Dafür ist sie schwerpunktmäßig in drei Abschnitte unterteilt, die sich nacheinander mit den Grundlagen, der Syntax, sowie Beispielen aus der Praxis beschäftigen.

Schlagworte

Auszeichnungssprachen, Datenbasis, Datenstruktur, DTD, Markierungssprachen, Metasprachen, XML, XSL

Ziel der Diplomarbeit

Das Internet - und allem voran das WWW - wandelt sich mit rasanter Geschwindigkeit von einer Spielwiese für Computerfreaks, zu einer professionellen Plattform für Businesslösungen aus den Bereichen der Medien, der Wirtschaft oder des Handels (um nur drei Beispiele zu nennen). Die bislang dominierende Markierungssprache HTML hält den daraus resultierenden Ansprüchen schon lange nicht mehr Stand.

In der Mediendokumentation war die Arbeit mit der herkömmlichen WWW-Sprache beispielsweise stets schwierig, da der beschränkte Satz an Markierungsbefehlen keine komfortablen und professionellen Lösungen für die Indexierung und Recherche zulässt. Die Metasprache SGML bietet da schon bessere Lösungsmöglichkeiten, doch für einen Einsatz im Internet ist diese Sprache zu komplex und weitschweifig.

Abhilfe aus diesem Dilemma verspricht eine neue Entwicklung, die 1998 vom World Wide Web Consortium unter der Leitung von Jon Bosak standardisiert wurde – die *Extensible Markup Language* (XML). Diese neue Markierungssprache vereint nicht nur die Vorteile einer Metasprache wie SGML, sie ist auch so schlank und webfähig wie HTML.

Mit XML ist es möglich, eigene Markierungen und Attribute zu definieren. Außerdem trennt sie, anders als HTML, streng die Strukturierung von der Auszeichnung von Daten. Das macht XML vor allem für drei Anwendungsgebiete interessant: als Metasprache, als systemunabhängige Markierungssprache und als Datenbasis.

Die vorliegende Arbeit will einen möglichst allgemein verständlichen Überblick über die Grundlagen, die Syntax und die Praxis von XML bieten. Jedem dieser drei Bereiche ist dafür ein eigener Abschnitt gewidmet.

Ein Schwerpunkt wurde dabei auf die Arbeit mit XML in der Medienbranche gesetzt. Dies kommt hauptsächlich in der Beschreibung der Syntax und im Praxis-Abschnitt zum Tragen.