

Harald Schöning

XML und Datenbanken

Konzepte und Systeme

HANSER



Inhalt

Vorwort	XI
1 Grundlagen von XML	2
1.1 Auszeichnungssprachen.....	2
1.2 Das XML-Dokument.....	4
1.2.1 Ein Beispiel.....	4
1.2.2 Die Struktur eines XML-Dokumentes	5
1.2.3 Die DTD	6
1.2.4 Der Prolog.....	11
1.2.5 Elementinhalt.....	12
1.3 Entities.....	12
1.4 Kommentare und Verarbeitungsanweisungen	14
1.5 Namensräume.....	14
1.6 Spezielle Attribute	17
1.7 Verweise in XML.....	19
1.7.1 Verweise in XML 1.0	19
1.7.2 XLink.....	20
1.7.3 XPointer.....	24
1.8 XML Kanonisierung.....	25
1.9 DOM und SAX.....	26
1.10 XML Information Set.....	29
1.11 XSL	29
1.12 Zusammenfassung	30
2 XML Schema	32
2.1 Überblick.....	32
2.2 Das Typkonzept.....	33
2.2.1 Vordefinierte Datentypen.....	33
2.2.2 Einfache und komplexe Typen	34
2.2.3 Komponenten einfacher Typen.....	35
2.2.4 Erzeugung von Datentypen.....	37
2.3 Definitionen.....	37

2.3.1	Definition einfacher Typen.....	37
2.3.2	Definition komplexer Typen.....	38
2.3.3	Gruppen.....	42
2.4	Deklarationen.....	43
2.5	Offene Schemata.....	44
2.6	Definition einschränkender Bedingungen.....	45
2.6.1	Eindeutigkeit von Werten.....	45
2.6.2	Referenzbeziehungen.....	46
2.6.3	Defizite des Konzepts.....	47
2.7	Namen und Namensräume.....	48
2.7.1	Gültigkeitsbereich.....	48
2.7.2	Namensräume.....	48
2.7.3	Syntaktische Repräsentation.....	49
2.8	Ableitung, Ersetzung, Import.....	52
2.8.1	Ersetzungsgruppen.....	53
2.8.2	Schema-Import.....	54
2.9	Zusammenhang zu DTDs.....	54
2.10	Instanzbezogene Konzepte.....	54
2.10.1	Das Attribut xsi:type.....	55
2.10.2	Das Attribut xsi:nil.....	55
2.10.3	Die Attribute xsi:schemaLocation und xsi:noNamespaceSchemaLocation.....	55
2.11	Physisches Schema.....	55
2.11.1	Entkoppelte Information.....	56
2.11.2	Proprietäre Schemasprache.....	56
2.11.3	Schema Adjunct.....	56
2.11.4	Schemasprachenerweiterung.....	56
2.12	Zusammenfassung.....	57
2.13	Andere XML-basierte Schemasprachen.....	59
2.13.1	Relax NG.....	59
3	Einsatz von XML.....	62
3.1	Nutzen von XML.....	62
3.2	Daten- und Dokumentsicht.....	64
3.2.1	Datenorientierte XML-Dokumente.....	64
3.2.2	Textorientierte XML-Dokumente.....	65
3.2.3	Wozu die Unterscheidung?.....	66
3.2.4	Mischformen.....	67
3.3	Beispiele für den Einsatz von XML.....	67
3.3.1	Publishing und Content Management.....	67
3.3.2	Elektronischer Handel.....	69
3.3.3	Workflow und integrierte Geschäftsprozesse.....	70
3.3.4	XML und Metadaten.....	71
3.3.5	XML und Multimedia.....	72

3.3.6	Semantische Aufbereitung von Dokumenten.....	73
3.3.7	Gesundheitswesen.....	73
3.3.8	Komplex strukturierte Daten.....	73
3.4	Web Services.....	73
3.4.1	SOAP.....	74
3.4.2	WSFL.....	76
3.5	ebXML.....	76
3.6	RosettaNet.....	76
3.7	BizTalk.....	76
3.8	Zusammenfassung.....	77
4	Entwurf von XML-Vokabularen.....	80
4.1	XML-Vokabular aus einem relationalen Schema.....	80
4.1.1	Das Verfahren zur DTD-Erzeugung nach [Wi00].....	80
4.1.2	Erzeugen von XML-Schemabeschreibungen nach [KL02a].....	83
4.1.3	Bewertung.....	85
4.2	Neuentwurf von XML Schemata.....	86
4.2.1	Abbildungsverfahren nach [KL02].....	87
4.3	Ableitung aus Instanzen.....	93
4.4	Allgemeine Entwurfsrichtlinien.....	94
4.5	Entwurf mit XML Schema.....	97
4.5.1	Bekannte Konzepte?.....	97
4.5.2	Einzelfragen zur Modellierung mit XML Schema.....	98
4.6	Modellierungsstile mit XML Schema.....	100
4.6.1	Globale oder lokale Elementdeklarationen.....	100
4.6.2	Verwendung von Typen.....	102
4.6.3	Verwendung von Gruppen.....	102
4.6.4	Ähnliche Objekte: Typhierarchien oder Komposition?.....	103
4.7	Zusammenfassung.....	104
5	XML in Datenbanksystemen.....	106
5.1	Motivation.....	106
5.2	Anforderungen an ein XML-Datenbanksystem.....	108
5.2.1	Grundmodelle von XML-Datenbanken.....	108
5.2.2	Datenbankspezifische Anforderungen.....	110
5.2.3	XML-spezifische Anforderungen.....	114
5.3	Anfragetypen.....	120
5.4	Anfrage-Ergebnisse.....	125
5.5	Andere Operationen.....	127
5.6	Schemata in XML-Datenbanksystemen.....	128
5.6.1	Schemata in Datenbanken.....	128
5.6.2	Schemata für XML.....	128
5.6.3	Schemata in XML-Datenbanksystemen.....	129

5.7	Zusammenfassung	130
6	Anfragesprachen für XML	132
6.1	Anforderungen an eine XML-Anfragesprache	133
6.2	XML-QL	136
6.2.1	Sprachelemente	136
6.2.2	Mustersuche	140
6.2.3	Weitere Fähigkeiten von XML-QL	141
6.2.4	Bewertung	142
6.3	XPath	143
6.3.1	Das XPath-Datenmodell	143
6.3.2	XPath-Datentypen	144
6.3.3	Adressierung von Dokumentbestandteilen	145
6.3.4	Funktionsaufrufe	152
6.3.5	Zusammenfassung und Bewertung	153
6.4	XQuery	154
6.4.1	XQuery-Anfrage	155
6.4.2	Anwendung der Konzepte	160
6.4.3	Vordefinierte Funktionen	161
6.4.4	Funktionen	162
6.4.5	Der Anfrageprolog	162
6.4.6	Bewertung	163
6.5	Zusammenfassung	164
7	XML in relationalen Datenbanksystemen	166
7.1	Inhaltsorientierte Zerlegung	167
7.1.1	Generische Vorgehensweise	168
7.1.2	Definitorische Vorgehensweise	178
7.1.3	Bewertung der inhaltsorientierten Zerlegung	181
7.2	Strukturorientierte Zerlegung	183
7.2.1	Bewertung der strukturorientierten Zerlegung	184
7.2.2	Problematik der Zerlegung	186
7.3	Opake Speicherung	187
7.3.1	Bewertung der opaken Speicherung	188
7.3.2	Vergleich mit der strukturorientierten Zerlegung	191
7.4	SQL/XML	192
7.4.1	Abbildung von SQL-Konzepten auf XML-Konzepte	192
7.4.2	Der neue Datentyp XML	199
7.4.3	Operatoren auf dem Datentyp XML	199
7.4.4	Bewertung	203
7.5	XML-Unterstützung in Oracle	204
7.5.1	Opake Speicherung und inhaltsorientierte Zerlegung	204
7.5.2	XML Schema und XMLType	210

7.5.3	Inhaltsorientierte Zerlegung von XMLType	212
7.5.4	Die Rolle der Textindizierung und Textsuche	215
7.5.5	Andere Indizierungen auf XMLType.....	218
7.5.6	Weitere Möglichkeiten zum Erzeugen von XML	219
7.5.7	XML-Schema-Generierung aus SQL-Objekttypen	225
7.5.8	URL-basierter Zugriff auf die XML-Sicht der Datenbank.....	226
7.5.9	Generische inhaltsorientierte Zerlegung: das XML SQL Utility	227
7.5.10	Das XDK	229
7.5.11	XSQL-Seiten.....	230
7.5.12	Andere Komponenten	233
7.5.13	Bewertung.....	233
7.6	XML-Unterstützung in DB2 UDB	235
7.6.1	Inhaltsorientierte Zerlegung mit DB2 XML-Extender	237
7.6.2	Opake Speicherung mit DB2 XML-Extender	241
7.6.3	Textsuche.....	246
7.6.4	XML-Werkzeuge	248
7.6.5	Bewertung	248
7.7	XML-Unterstützung in Microsoft SQL Server 2000	250
7.7.1	XML-Generierung	250
7.7.2	Zugriff mittels Template	255
7.7.3	Virtuelle XML-Sichten	256
7.7.4	HTTP-Zugriff	260
7.7.5	Änderungsoperationen	262
7.7.6	Programmatischer Zugriff auf XML-Dokumente	263
7.7.7	Bewertung	264
7.8	XML-Unterstützung in (objekt-)relationalen Datenbanksystemen	266
8	Reine XML-Datenbanksysteme	270
8.1	Tamino	271
8.1.1	Überblick	271
8.1.2	Speicherung von XML-Dokumenten	272
8.1.3	Schemabeschreibung der XML-Dokumente	273
8.1.4	Speicherung in Tamino	276
8.1.5	Zugriffsmodelle	278
8.1.6	Behandlung von Sortierordnungen	278
8.1.7	Indexierung.....	279
8.1.8	Anfragen	282
8.1.9	Sicherheit.....	286
8.1.10	Zugriff auf andere Datenbanken: X-Node	287
8.1.11	X-Tension	290
8.1.12	Schnittstellen.....	293
8.1.13	Werkzeuge	294
8.1.14	Internationalisierung	297

8.1.15	Funktionalität eines Datenbanksystems	298
8.1.16	Bewertung	298
8.2	eXcelon eXtensible Information Server	299
8.2.1	Überblick	299
8.2.2	Speicherung von XML	299
8.2.3	Schemabeschreibung der XML-Dokumente	301
8.2.4	Architektur von XIS	301
8.2.5	Indexierung	302
8.2.6	Anfragen	303
8.2.7	Sicherheit	306
8.2.8	Erweiterbarkeit	306
8.2.9	Schnittstellen	307
8.2.10	Werkzeuge	307
8.2.11	Internationalisierung	308
8.2.12	Funktionalität eines Datenbanksystems	309
8.2.13	Bewertung	309
8.3	Zusammenfassung	310
9	Benchmarks für XML-Datenbanken	312
9.1	Der XMark Benchmark	312
9.1.1	Zielsetzung	312
9.1.2	Beispielanwendung	312
9.1.3	Bewertung	314
9.2	Der Michigan Benchmark	314
9.2.1	Zielsetzung	314
9.2.2	Struktur des Benchmarks	315
9.2.3	Bewertung	316
9.3	Der XOO7 Benchmark	316
9.3.1	Zielsetzung	316
9.3.2	Beispielanwendung	316
9.3.3	Bewertung	317
9.4	Der XMach-1 Benchmark	317
9.4.1	Zielsetzung	317
9.4.2	Beispielanwendung	317
9.4.3	Bewertung	320
9.5	Zusammenfassung	320
	Glossar	323
	Literatur	327
	Index	335