

Contents at a Glance

| | |
|--|------------|
| <i>Introduction</i> | 1 |
| <i>Part I: XML Basics</i> | 9 |
| Chapter 1: Getting to Know XML | 11 |
| Chapter 2: Using XML for Many Purposes | 23 |
| Chapter 3: Slicing and Dicing Data Categories: The Art of Taxonomy | 33 |
| <i>Part II: XML and the Web</i> | 45 |
| Chapter 4: Adding XHTML for the Web | 47 |
| Chapter 5: Putting Together an XML File | 65 |
| Chapter 6: Adding Character(s) to XML | 83 |
| Chapter 7: Handling Formatting with CSS | 95 |
| <i>Part III: Building In Validation with DTDs and Schemas</i> | 109 |
| Chapter 8: Understanding and Using DTDs | 111 |
| Chapter 9: Understanding and Using XML Schema | 135 |
| Chapter 10: Building a Custom XML Schema | 157 |
| Chapter 11: Modifying an Existing Schema | 173 |
| <i>Part IV: Transforming and Processing XML</i> | 195 |
| Chapter 12: Handling Transformations with XSL | 197 |
| Chapter 13: The XML Path Language | 215 |
| Chapter 14: Processing XML | 235 |
| <i>Part V: XML Application Development</i> | 245 |
| Chapter 15: Using XML with Web Services | 247 |
| Chapter 16: XML and Forms | 259 |
| Chapter 17: Serving Up the Data: XML and Databases | 271 |
| Chapter 18: XML and RSS | 285 |
| <i>Part VI: The Part of Tens</i> | 299 |
| Chapter 19: XML Tools and Technologies | 301 |
| Chapter 20: Ten Top XML Applications | 313 |
| Chapter 21: Ten Ultimate XML Resources | 321 |
| <i>Glossary</i> | 329 |
| <i>Index</i> | 347 |

Table of Contents

| | |
|---|-----------|
| <i>Introduction</i> | 1 |
| About This Book | 1 |
| Conventions Used in This Book | 2 |
| Foolish Assumptions | 3 |
| How This Book Is Organized | 4 |
| Part I: XML Basics | 4 |
| Part II: XML and the Web | 4 |
| Part III: Building in Validation with DTDs and Schemas | 5 |
| Part IV: Transforming and Processing XML | 6 |
| Part V: XML Application Development | 6 |
| Part VI: The Part of Tens | 7 |
| Glossary | 7 |
| Icons Used in This Book | 7 |
| Where to Go from Here | 8 |
| | |
| <i>Part I: XML Basics</i> | 9 |
| | |
| Chapter 1: Getting to Know XML | 11 |
| XML (eXtremeMely cool) | 12 |
| Mocking up your own markup | 12 |
| Separating data and context | 12 |
| Making information portable | 13 |
| XML means business | 13 |
| Figuring Out What XML Is Good For | 14 |
| Classifying information | 14 |
| Enforcing rules on your data | 15 |
| Outputting information in a variety of ways | 16 |
| Using the same data across platforms | 17 |
| Beyond the Hype: What XML Isn't | 18 |
| It's not just for Web pages anymore | 19 |
| It's not a database | 20 |
| It's not a programming language | 20 |
| Building XML Documents | 21 |
| | |
| Chapter 2: Using XML for Many Purposes | 23 |
| Moving Legacy Data to XML | 23 |
| The Many Faces of XML | 24 |
| Creating XML-enabled Web pages | 24 |
| Print publishing with XML | 25 |

| | |
|---|----|
| Using XML for business forms | 28 |
| Incorporating XML into business processes | 29 |
| Serving up XML from a database | 31 |
| Alphabet Soup: Even More XML | 31 |

Chapter 3: Slicing and Dicing Data Categories:

The Art of Taxonomy 33

| | |
|---|----|
| Taking Stock of Your Data | 33 |
| Looking at business practices and partners | 34 |
| Gathering some content | 34 |
| Checking whether a DTD or schema already exists | 35 |
| Searching for a schema repository | 36 |
| Breaking Down Data in Different Ways | 37 |
| Winnowing out the wheat from the chaff | 38 |
| Types of data that can be stored in XML | 39 |
| Developing Your Taxonomy | 39 |
| Testing Your Taxonomy | 41 |
| Using trial and error for the best fit | 41 |
| Testing your content analysis | 42 |
| Looking Ahead to Validation | 43 |

Part II: XML and the Web 45

Chapter 4: Adding XHTML for the Web 47

| | |
|--|----|
| HTML, XML, and XHTML | 47 |
| What HTML does best | 48 |
| The limits of HTML | 49 |
| Comparing XML and HTML | 50 |
| Using XML to describe data | 51 |
| The benefits of using HTML | 53 |
| The benefits of using XML | 53 |
| XHTML Makes the Move to XML Syntax | 54 |
| Making the switch | 55 |
| Every element must be closed | 56 |
| Empty elements must be formatted correctly | 56 |
| Tags must be properly nested | 57 |
| Case makes a difference | 57 |
| Attribute values are in quotation marks | 58 |
| Converting a document from HTML to XHTML | 59 |
| The Role of DOCTYPE Declarations | 62 |

Chapter 5: Putting Together an XML File 65

| | |
|---|----|
| Anatomy of an XML File | 65 |
| The XML declaration | 67 |
| Marking up your content | 68 |
| Playing by the Rules: Well-Formed Documents | 74 |

| | |
|---|----------------|
| Adding Style for the Web | 76 |
| Seeking Validation with DTD and XML Schema | 78 |
| Why describe XML documents? | 79 |
| Choosing between DTD and XML Schema | 80 |
| Chapter 6: Adding Character(s) to XML | 83 |
| About Character Encodings | 84 |
| Introducing Unicode | 85 |
| Character Sets, Fonts, Scripts, and Glyphs | 87 |
| For Each Character, a Code | 88 |
| Key Character Sets | 89 |
| Using Unicode Characters | 91 |
| Finding Character Entity Information | 93 |
| Chapter 7: Handling Formatting with CSS | 95 |
| Viewing XML on the Web with CSS | 96 |
| Basic CSS Formatting: CSS1 | 97 |
| The Icing on the Cake: CSS2 | 98 |
| Building a CSS Stylesheet | 98 |
| Adding CSS to XML | 99 |
| A simple CSS stylesheet for XML | 101 |
| Dissecting a simple CSS stylesheet | 102 |
| Linking CSS and XML | 106 |
| Adding CSS to XSLT | 107 |
| Part III: Building In Validation with DTDs and Schemas | 109 |
| Chapter 8: Understanding and Using DTDs | 111 |
| What's a DTD? | 112 |
| When to use a DTD | 113 |
| When NOT to use a DTD | 113 |
| Inspecting the XML Prolog | 114 |
| Examining the XML declaration | 115 |
| Discovering the DOCTYPE | 116 |
| Understanding comments | 116 |
| Processing instructions | 117 |
| How about that white space? | 117 |
| Reading a DTD | 118 |
| Using Element Declarations | 119 |
| Using the EMPTY element type and the ANY element type | 120 |
| Adding mixed content | 121 |
| Using element content models | 122 |
| Declaring Attributes | 123 |
| Discovering Entities | 125 |
| General entities | 126 |
| Parameter entities | 128 |

| | |
|---|------------|
| Understanding Notations | 130 |
| Calling a DTD | 131 |
| Internal DTDs | 131 |
| External DTDs | 132 |
| When to use an internal or external DTD | 133 |
| Chapter 9: Understanding and Using XML Schema | 135 |
| What's an XML Schema? | 136 |
| So Many Datatypes, So Little Time | 138 |
| XML Prolog | 139 |
| Document Structures | 141 |
| Element declarations | 141 |
| </confirmOrder> Attribute declarations | 144 |
| Attribute groups | 144 |
| What about that white space? | 145 |
| Datatype Declarations | 148 |
| Simple datatypes | 148 |
| Complex datatypes | 149 |
| Defining constraints and value checks | 149 |
| Dealing with Entities, Notations, and More | 150 |
| Annotations | 151 |
| Deciding When to Use a Schema | 152 |
| Referencing XML Schema Documents | 153 |
| The inside view: Referencing a schema in an XML document | 153 |
| Calling for outside support: Referencing external schemas in your schema | 153 |
| Double-Checking Your Schemas and Documents | 155 |
| Chapter 10: Building a Custom XML Schema | 157 |
| Doing the Validity Rag | 157 |
| Step 1: Understanding Your Data | 159 |
| Step 2: Being the Root of All Structure: Elements | 159 |
| Step 3: Building Content Models | 161 |
| Step 4: Using Attributes to Shed Light on Data Structure | 163 |
| Step 5: Using Datatype Declarations to Define What's What | 164 |
| Tricks of the Trade | 167 |
| Creating a Simple Schema | 168 |
| Using a Schema with an XML File in Word 2003 | 170 |
| Chapter 11: Modifying an Existing Schema | 173 |
| Trading Control for Flexibility | 174 |
| Eliciting Markup from an XML Schema | 174 |
| Modifying a Schema | 176 |
| Using Datatypes Effectively | 177 |
| Using datatypes with data-intensive content | 177 |
| Using datatypes with text-intensive content | 179 |

| | |
|--|-----|
| Making Elements Work Wisely and Well | 180 |
| Creating crafty content models | 180 |
| A matter of selection | 181 |
| Mixing up the order | 183 |
| Using Complex Datatypes | 183 |
| When XML Schemas Collide: Namespaces | 185 |
| Including External Data | 188 |
| Including/Excluding Document Content | 188 |
| Converting DTDs to Schemas | 190 |

Part IV: Transforming and Processing XML 195

Chapter 12: Handling Transformations with XSL 197

| | |
|---|-----|
| The Two Faces of XSL | 198 |
| XSLT | 198 |
| XSL-FO | 200 |
| XSL Stylesheets Are XML Documents | 201 |
| A Simple Transformation Using XSLT | 202 |
| An XSLT Stylesheet for Converting XML to HTML | 202 |
| The pieces of the stylesheet puzzle | 205 |
| Processing element content | 207 |
| Dealing with repeating elements | 209 |
| Creating an XSLT Stylesheet with XSLT Editors | 210 |

Chapter 13: The XML Path Language 215

| | |
|--|-----|
| Why Do You Need Directions? | 216 |
| XPath document trees | 217 |
| Understanding XPath nodes | 218 |
| XPath Directions and Destinations | 220 |
| XPath Syntax | 221 |
| Some simple location paths | 222 |
| Adding expressions | 223 |
| Taking steps along the XPath | 223 |
| Looking at attributes | 224 |
| Going backward | 224 |
| Reversing direction | 225 |
| Null results | 225 |
| Getting back to your roots | 226 |
| XPath functions | 226 |
| Using XPath with XMLSpy | 226 |
| The Short Version | 228 |
| Child-axis abbreviations | 229 |
| Attribute-axis abbreviation | 229 |
| Predicate and expression abbreviations | 229 |
| Some more abbreviations | 230 |
| What's New in XPath 2.0? | 231 |
| Where to Now? | 233 |

| | |
|--|------------|
| Chapter 14: Processing XML | 235 |
| Frankly, My Dear, I Don't Give a DOM | 235 |
| Keeping in touch with the family | 238 |
| Understanding DOM structure | 238 |
| What Goes In Must Come Out: Processing XML | 240 |
| So many processors, so little time | 242 |
| Which processor is right for you? | 243 |

Part V: XML Application Development

245

| | |
|---|------------|
| Chapter 15: Using XML with Web Services | 247 |
| What's Up with Web Services? | 248 |
| A Web Services Architecture | 251 |
| Transport: Moving XML messages | 252 |
| Packaging/Extensions: Managing information exchange | 253 |
| Description: Specifying services and related components | 254 |
| Discovery: Finding what's available | 255 |
| Where Will Web Services Lead? | 256 |

| | |
|---|------------|
| Chapter 16: XML and Forms | 259 |
| Collecting Information with Forms: The Basics | 260 |
| HTML Forms | 260 |
| XML Forms | 261 |
| XForms | 261 |
| InfoPath | 267 |

| | |
|---|------------|
| Chapter 17: Serving Up the Data: XML and Databases | 271 |
| Using Databases with XML | 272 |
| Text-intensive XML | 272 |
| Data-intensive XML | 273 |
| Creating XML from Database Files | 273 |
| Using Word 2003 | 274 |
| Using InfoPath | 275 |
| Using XMLSpy | 278 |
| Using XML with Access 2003 | 281 |

| | |
|--------------------------------------|------------|
| Chapter 18: XML and RSS | 285 |
| Introducing RSS | 286 |
| Sorting Out the Versions | 286 |
| RSS 0.9x | 287 |
| RSS 2.0/2.01 | 290 |
| RSS 1.0 | 291 |
| Validating an RSS Feed | 295 |
| Creating RSS Feeds | 296 |
| Get Syndicated! | 297 |
| Using an RSS Reader | 298 |

Part VI: The Part of Tens*299*Chapter 19: XML Tools and Technologies301**

| | |
|--|-----|
| Creating Documents with Authoring Tools | 301 |
| Epic Editor | 302 |
| Turbo XML v2.4.1 | 303 |
| XMetaL Author 4.5 | 303 |
| XML Pro v2.0.1 | 303 |
| XML Spy 2005 | 304 |
| Checking Documents with Parser Tools | 304 |
| Ælfred | 305 |
| expat | 306 |
| Lark | 306 |
| Viewing with XML Browsers | 307 |
| Amaya | 307 |
| Internet Explorer 6 | 307 |
| Mozilla | 308 |
| Firefox 1.0 | 308 |
| Opera | 308 |
| Using XML Parsers and Engines | 309 |
| XML C Library for Gnome | 309 |
| Java XML Pack | 310 |
| Xerces | 310 |
| Employing Conversion Tools | 311 |
| HTML Tidy | 311 |
| Extensible Programming Script (XPS) | 311 |
| The Ultimate XML Grab Bag and Goodie Box | 312 |
| Microsoft does XML, too! | 312 |
| webMethods automates XML excellence | 312 |

Chapter 20: Ten Top XML Applications313

| | |
|---|-----|
| XHTML = XML + HTML | 314 |
| XML Style Is a Matter of Application | 314 |
| Wireless Markup Language (WML) | 314 |
| DocBook, Anyone? | 315 |
| Mathematical Markup Language (MathML) | 315 |
| Scalable Vector Graphics (SVG) | 316 |
| Resource Description Framework (RDF) | 316 |
| Synchronized Multimedia Integration Language (SMIL) | 317 |
| Servin' Up Web Services | 317 |
| XQuery | 318 |
| Create XML Applications with Zope | 319 |

Chapter 21: Ten Ultimate XML Resources321

| | |
|--------------------------------------|-----|
| XML's Many and Marvelous Specs | 321 |
| An XML Nonpareil | 322 |
| Top XML Tutorial Sites | 322 |

| | |
|--|------------|
| XML in the Mail | 323 |
| Excellent XML Examples at zvon.org | 323 |
| XML News and Information | 323 |
| XML Training Options | 324 |
| Building a Bodacious XML Bookshelf | 325 |
| Studying XML for Certification | 326 |
| Serious Searches Lead to Success | 327 |
| <i>Glossary</i> | 329 |
| <i>Index</i> | 347 |