Obsah

Obsah **5**
Úvod **12**
Proč tato kniha vznikla 12
Co můžete od knížky očekávat 13
Jak jsou organizovány programy 14
Poděkování 15
Konvence 16
1. XML – základní informace **17**
1.1. Základní charakteristiky 17
1.1.1. Vývoj XML 19
1.1.2. Ověření správnosti XML dokumentu 20
1.1.3. Dvě hlavní oblasti použití XML 21
1.2. Syntaxe a prvky XML 23
1.2.1. Názvy značek v XML 24
1.2.1.1. Praktické doporučení pro datově orientované dok. 24
1.2.2. Obecně platná pravidla pro značky a elementy 24
1.2.3. Atributy 26
1.2.4. Kdy použít elementy a kdy atributy 27
1.2.4.1. Kdy použít atributy 28
1.2.4.2. Kdy použít vnořené elementy 30
1.2.5. Znaková entita 30
1.2.6. Sekce CDATA 31
1.2.7. Komentáře 32
1.2.8. Zpracovávací instrukce 33
1.2.9. Použité kódování a možnosti zápisu netypických znaků 34
1.2.9.1. Nesoulad deklarovaného a použitého kódování 34
1.2.9.2. Výskyt netypických znaků 34
1.3. Ukázka výhod datově orientovaného XML dokumentu 35
1.3.1. Binární soubor v proprietárním formátu 35
1.3.2. Textový soubor 36
1.3.3. XML dokument 37
1.4. Jiné způsoby zápisu XML dokumentu 38
1.4.1. Přísně hierarchické schéma 38
1.4.2. Ploché schéma 40
1.4.3. Porovnání obou způsobů 41
1.5. Kontrola správné strukturovanosti XML dokumentu 41
1.6. Jmenné prostory 43
1.7. Průběžný příklad 49

2. Schémové jazyky, DTD 50
2.1. Význam schémových jazyků 50
2.1.1. Možnosti a způsoby využití schémových jazyků 50
2.2. DTD – Document Type Definition 51
2.2.1. Výhody a nevýhody DTD 52
2.2.2. Spojení souborů DTD a XML 52
2.2.3. Prvky a struktura DTD 53
2.2.3.1. Deklarace elementů 54
2.2.3.1.1. Prázdný element 54
2.2.3.1.2. Modelová skupina 55
2.2.3.1.3. Počty vnořených elementů 59
2.2.3.1.4. Koncový element 60
2.2.3.1.5. Element obsahuje smíšený obsah 60
2.2.3.2. Deklarace atributů 61
2.2.3.2.1. Typ atributu 61
2.2.3.2.2. Povinnost atributu 63
2.2.3.2.3. Více atributů v jednom elementu 65
2.2.3.3. Parametrické entity 66
2.2.4. Větší příklad 66
2.2.5. Průběžný příklad 69
3. Schémový jazyk W3C XML Schema 70
3.1. Základní informace 70
3.2. Základní principy práce s XSD souborem 71
3.2.1. Začátek XSD souboru 72
3.2.2. Jednoduchý úvodní příklad 72
3.3. Výběr z běžně použitelných základních typů 75
3.4. Možnosti základních typů 76
3.4.1. Data a časy 76
3.4.2. Logický datový typ 77
3.4.3. Reálná čísla 78
3.4.4. Desítková čísla 78
3.4.5. Celá čísla 79
3.4.6. Řetězce 83
3.4.7. URI 85
3.5. Možnosti restrikcí 86
3.5.1. Omezení číselného rozsahu 86
3.5.2. Omezení počtu platných míst shora 87
3.5.3. Omezení počtu desetinných míst 87
3.5.4. Omezení počtu znaků 88
3.5.5. `enumeration` – výčet hodnot 89
3.5.6. `pattern` – vzor pomocí regulárního výrazu 90

3.6. Složené (komplexní) datové typy ... 91
3.6.1. Skládání pomocí sequence ... 92
3.6.1.1. Ohraničené opakování ... 93
3.6.1.2. Neohraničené opakování ... 93
3.6.2. Skládání pomocí all ... 94
3.6.2.1. Ohraničené opakování ... 95
3.6.3. Skládání pomocí choice ... 96
3.6.3.1. Ohraničení ... 97
3.6.4. Smíšený obsah ... 97
3.6.5. Prázdný element ... 98
3.7. Atributy ... 99
3.7.1. Atribut je součástí složeného elementu ... 99
3.7.2. Atribut je součástí elementu obsahujícího hodnotu ... 100
3.7.3. Atribut je součástí prázdného elementu ... 101
3.7.4. Více atributů než jeden ... 102
3.8. Závěrečná definice kořenového elementu ... 102
3.9. Připojení schématu ke XML dokumentu ... 103
3.10. Validace pomocí Xerces ... 104
3.11. Jmenné prostory ... 104
3.11.1. Příprava XSD souboru ... 104
3.11.2. XML soubor využívá pouze jeden XSD soubor ... 105
3.11.3. XML soubor využívá více XSD souborů ... 106
3.12. Speciální dovednosti ... 111
3.12.1. Použití prázdné hodnoty ... 111
3.12.2. Práce s okrajovými bílými znaky ... 112
3.13. Průběžný příklad ... 115
4. Java a XML, SAX ... 117
4.1. Java a XML ... 117
4.1.1. Obecné vlastnosti parseru ... 117
4.1.1.1. Proudové čtení ... 118
4.1.1.2. Práce se stromovou reprezentací dokumentu ... 119
4.1.2. Rozhraní parserů pro Javu ... 120
4.1.2.1. JAXP – Java API for XML Processing ... 120
4.1.2.2. JDOM ... 120
4.1.3. Stručný přehled parserů ... 121
4.1.4. Konkrétní parsery ... 122
4.2. SAX – Simple API for XML ... 122
4.2.1. Úvodní informace ... 122
4.2.2. Základní postup při zpracování ... 123
4.2.2.1. Hlavní program ... 123
4.2.2.2. Zpracování výjimek v hlavním programu ... 125
4.2.2.3. Pomocná třída pro zpracování chyb ... 126

4.2.3. Zpracování parsovaného XML dokumentu 127
4.2.3.1. Práce pouze s obsahem elementů 129
4.2.3.2. Práce s atributy 132
4.2.3.3. Převod infosetu na seznam objektů 133
4.2.4. Zpracování složitějšího XML dokumentu 136
4.2.5. Problematika různého kódování 137
4.2.6. Nastavení vlastností parseru 138
4.2.7. Validace oproti DTD nebo XSD 140
4.2.8. Práce se jmennými prostory 143
5. DOM – Document Object Model 146
5.1. Základní informace 146
5.1.1. Kolekce uzlů 149
5.2. Základní použití DOM 149
5.3. Zpracování parsovaného XML dokumentu 151
5.3.1. Práce pouze s obsahem elementů 151
5.4. Metody rozhraní `Node` 153
5.4.1. Metody pro získání informace o uzlu 154
5.4.2. Metoda pro pohyb v hierarchii nahoru (na rodiče) 154
5.4.3. Metody pro horizontální pohyb (na sourozence) 154
5.4.4. Metody pro pohyb v hierarchii dolů (na potomky) 155
5.4.5. Metody pro práci s atributy 155
5.5. Bezpečné vyhledávání uzlu a práce s atributy 156
5.6. Problém vkládaných elementů (odřádkování) 158
5.7. Automatické odstranění komentářů 160
5.8. Převod infosetu na seznam objektů 161
5.9. Sofistikovaný průchod stromem dokumentu 163
5.10. Automatické odstranění odřádkovávacích uzlů 169
5.11. Zápis dokumentu 170
5.11.1. Ovlivnění práce transformační třídy 171
5.11.2. Ukázka zápisu do souboru 173
5.11.3. Problematika odřádkování a odsazování 177
5.11.3.1. Změna odsazení XML dokumentu 181
5.12. Modifikace dokumentu 182
5.12.1. Změna hodnoty již existujícího elementu či atributu 182
5.12.2. Odstranění uzlu 183
5.12.3. Vkládání nových uzlů 183
5.12.4. Příklad změny XML dokumentu a jeho zápisu 184
5.13. Vytváření nového dokumentu 186
5.13.1. Klonování uzlů 186
5.14. Validace nově vytvořeného nebo měněného dokumentu 188
5.14.1. Validace v JDK 1.6 191

6. JWSDP, StAX 195
6.1. JWSDP 195
6.2. StAX 196
6.2.1. Základní informace 197
6.2.2. Základní postup při zpracování 198
6.2.3. Přehled základních možností čtení 200
6.2.3.1. Práce pouze s obsahem elementů 200
6.2.4. Zpracování atributů 201
6.2.4.1. Výpočet celkové ceny 201
6.2.4.2. Převod infosetu na seznam objektů 202
6.2.5. Čtení na žádost 204
6.2.6. Zápis do XML dokumentu 206
6.2.6.1. Generování souboru pomocí transformace 208
7. JAXB 212
7.1. Základní informace 212
7.1.1. Rozdíl ve verzích JDK 213
7.1.1.1. Společné věci 214
7.1.1.2. Odlišnosti pro JDK 1.5 214
7.1.2. Podpora z Ant 214
7.2. Generování souborů z XSD 215
7.2.1. Princip bindingu 217
7.3. Čtení XML dokumentu 219
7.3.1. Práce pouze s obsahem elementů 220
7.4. Čtení dokumentu včetně zpracování atributů 222
7.4.1. Všechny objekty v paměti 223
7.5. Změna hodnot a vytvoření nových elementů 225
7.6. Zápis do XML dokumentu 226
7.7. Validace 229
7.8. Příprava kompletně nového dokumentu 231
8. Ant 234
8.1. Základní informace 234
8.2. Jak Ant získat 235
8.3. Základní použití 236
8.3.1. Pojmenování a spuštění souboru projektu 236
8.3.2. Target 237
8.3.3. Atributy `<project>` 239
8.4. Použití `<property>` 240
8.4.1. Nastavení obecné hodnoty 240
8.4.2. Nastavení hodnoty ve smyslu „jméno souboru“ 241
8.4.2.1. Pozor na skládání jmen adresářů 242

8.4.3. Přednastavené konstanty....243
8.4.3.1. Konstanta file....243
8.4.3.2. Konstanta environment....244
8.5. Nastavování cest a opětovné použití jejich nastavení....245
8.6. Nastavení jmen souborů....247
8.7. Kombinování <path> a <fileset>....248
8.8. Možnosti <target>....249
8.8.1. Seznam <target>....249
8.8.2. Vzájemné závislosti....250
8.8.3. Podmíněné provádění <target>....252
8.8.3.1. Kombinování podmínek a závislostí....254
8.9. Když je něco špatně....255
8.10. Výkonné příkazy – úlohy....256
8.11. Přehled a použití často používaných úloh....257
8.11.1. <copy> – kopírování souborů a adresářů....257
8.11.2. <delete> – mazání....259
8.11.3. <mkdir> – vytvoření adresářů....260
8.11.4. <move> – přesunutí souborů....260
8.11.5. <javac> – Java překladač....260
8.11.6. <java> – Java interpret....262
8.11.7. <javadoc> – generování dokumentace....263
8.11.8. <jar> – vytvoření archivu....264
8.11.9. Časová známka <tstamp>....265
8.11.10. <native2ascii> – optional task....267
8.11.11. Použití vlastní úlohy – XJC....268
8.11.11.1. XJC pro JDK 1.6....269
8.12. Ukázka komplexního projektu použitelného v praxi....270
9. Kódování znaků....277
9.1. Základní terminologie....277
9.1.1. Coded Character Set....277
9.1.2. Character Encoding Form....278
9.1.3. Character Encoding Scheme....279
9.2. Historie....281
9.2.1. Osmibitové znakové sady....281
9.3. Současnost....283
9.3.1. Historie Unicode....284
9.3.1.1. Překročení hranice 16 bitů....285

9.4. Problémy v Unicode a jejich řešení 286
9.4.1. Základem je US-ASCII 286
9.4.2. Problém pořadí bajtů 286
9.4.2.1. Little Endian 287
9.4.2.2. Big Endian 287
9.4.3. Problém kódovacích schémat 287
9.4.3.1. ISO 10646 288
9.4.3.2. Unicode 288
9.4.4. Značka bajtového pořadí 290
9.5. Kódování používaná pro Unicode 291
9.5.1. UTF-8 292
9.5.1.1. Princip kódování 294
9.5.1.2. Problém značky bajtového pořadí 296
9.5.1.3. Kódovací schéma UTF-7 298
9.5.2. UTF-16 298
9.5.2.1. Podrobný princip 299
9.5.3. UTF-32 300
9.6. Problémy pojmenování kódování 300
9.7. Praktický dopad na uživatele počítačů v České republice 301
9.7.1. US-ASCII 301
9.7.2. ISO-8859-2 – Latin Alphabet No. 2 302
9.7.3. windows-1250 – Windows Eastern European 302
9.7.4. IBM852 – MS-DOS Latin-2 302
9.7.5. x-MacCentralEurope – Macintosh Latin-2 302
9.7.6. UTF-8 – Eight-bit UCS Transformation Format 303
9.7.7. UTF-16 – Sixteen-bit UCS Transformation Format 303
9.7.8. UTF-16BE 303
9.7.9. UTF-16LE 303
9.7.10. Závěrečný přehled 304
Literatura 305
Rejstřík 306