



## Predgovor

U nekoliko posljednjih godina, XML je počeo da se upotrebljava u najrazličitijim oblastima, među kojima su: pravo, aeronautika, finansije, osiguranje, robotika, multimedija, ugostiteljstvo, turizam, likovne umetnosti, graditeljstvo, telekomunikacije, softver, poljoprivreda, fizika, novinarstvo, teologija, maloprodaja i stripovi. XML je postao najpoželjnija sintaksa za nove formate dokumenata u gotovo svim računarskim aplikacijama. Upotrebljava se na Linuxu, Windowsu, Macintoshu i mnogim drugim računarskim platformama. Veliki centralni računari u njujorškom Volstritu međusobno trguju deonicama tako što razmenjuju odgovarajuće XML dokumente. Deca koja igraju igrice na kućnim računarima snimaju svoje dokumente u XML-u. Navijači na ekranima svojih mobilnih telefona u realnom vremenu primaju rezultate utakmica, opet u formatu XML. XML je jednostavno najrobusnija, najpouzdanija i najfleksibilnija dosad izumljena sintaksa za dokumente.

*XML za programere* je sveobuhvatan vodič kroz brzo rastući svet XML-a. Obrađuje sve vidove XML-a – od osnovnih pravila sintakse, preko pojedinosti pravljenja DTD-ova i šema, do interfejsa za programiranje aplikacija (API-ja) pomoću kojih možete čitati i pisati XML dokumente na raznim programskim jezicima.

### Šta obuhvata ova knjiga

Postoje hiljade XML aplikacija koje su formalno ustanovili konzorcijum W3C i druga tela za standardizaciju, kao što su OASIS i Object Management Group. Jos je više nestandardizovanih aplikacija, poput Microsoftove Channel Definition Format i Mind Reading Markup Language Johna Guajarda. Sve njih nije moguće opisati u jednoj knjizi, kao što se u jednoj knjizi ne mogu opisati ni svi programi koji se mogu napisati ili su već napisani na Javi. Ova knjiga se usredsređuje na sam XML. Obuhvaćena su osnovna pravila kojih se moraju pridržavati svi XML dokumenti i njihovi autori – od Web dizajnera koji upotrebljavaju SMIL za dodavanje animacija na Web stranice, do C++ programera koji za razmenu serijalizovanih objekata sa udaljenom bazom podataka koriste SOAP.

U knjizi su opisane i sledeće opšte tehnologije za podršku XML-a, koje se koriste u raznim XML aplikacijama:

#### *XLink*

Atributska sintaksa za hiperveze između XML i ne-XML dokumenata. One mogu biti: proste jednosmerne veze poznate iz HTML-a, višesmerne veze između više dokumenata i veze između dokumenata u koje ne možete upisivati.

#### *XSLT*

XML aplikacija koja opisuje transformacije jednog dokumenta u drugi, pri čemu se koriste isti ili različiti XML rečnici.

#### *XPointer*

Sintaksa za URI identifikatore odlomaka, kojima se biraju određeni delovi XML dokumenta na koji upućuje URI identifikator; često se koristi zajedno sa XLinkom.

#### *XPath*

Ne-XML sintaksa koju XPointer i XSLT upotrebljavaju za identifikovanje određenih delova XML dokumenata. Primera radi, pomoću XPatha može se izdvojiti treći element adrese dokumenta ili svi elementi sa atributom `email` čija je vrednost `elharo@metalab.unc.edu`.

#### *XInclude*

Sredstvo za pravljenje velikih XML dokumenata kombinovanjem drugih gotovih dokumenata i njihovih odlomaka.

#### *Prostori imena*

Sredstvo za uočavanje razlika između elemenata i atributa iz različitih, ali istoi-  
menih XML rečnika; na primer, za razlikovanje naslova knjige od naslova Web  
stranice koja je sastavni deo Web lokacije o knjigama.

#### *Šeme*

XML rečnik za opisivanje dozvoljenog sadržaja XML dokumenata napravlje-  
nih pomoću drugih XML rečnika.

#### *SAX*

Akronim od *Simple API for XML*, što je naziv interfejsa za programiranje apli-  
kacija zasnovanih na događajima; implementiraju ga mnogi analizatori  
(raščlanjivači) XML teksta.

#### *DOM*

Akronim od *Document Object Model* (objektni model dokumenta), što je ime  
jezički nezavisnog interfejsa za programiranje aplikacija, koji XML dokument  
tretira kao stablo ugnežđenih objekata različitih svojstava.

#### *XHTML*

XML-izovana verzija HTML-a, koja se može proširiti drugim XML aplikacija-  
ma, kao što su MathML i SVG.

#### *RDDL*

Akronim od *Resource Directory Description Language* (jezik za opisivanje kata-  
loga resursa), što je naziv XML aplikacije zasnovane na jeziku XHTML, koja je  
namenjena dokumentima smeštenim na kraj URL adrese prostora imena.

Navedene tehnologije se koriste u više različitih XML aplikacija, bez obzira na to da li su definisane u XML-u (kao XLinks, XSLT, XInclude, prostori imena, šeme, XHTML i RDDL) ili u nekoj drugoj sintaksi (kao XPointers, XPath, SAX i DOM).

U ovoj knjizi nisu detaljno objašnjene XML aplikacije koje su važne samo manjem delu korisnika XML-a. Među takvim aplikacijama su:

#### SVG

*Scalable Vector Graphics* – standard za XML kodiranje vektorskih crteža, koji je usvojio W3C.

#### MathML

*Mathematical Markup Language* –XML aplikacija za ugradnju jednačina u Web stranice i druge dokumente, koju je W3C usvojio kao standard.

#### RDF

*Resource Description Framework* –XML aplikacija za opisivanje resursa, naročito metapodataka koji se mogu naći u kartičnom katalogu biblioteka. Prihvaćena u W3C kao standard.

Ponekad ćemo neku od tih aplikacija upotrebiti u primerima, ali nećemo detaljno objašnjavati sve aspekte odgovarajućeg rečnika. Te aplikacije su zanimljive i važne, ali su (kao i hiljade drugih) prvenstveno namenjene upotrebi u specijalnom softveru koji poznaje sve pojedinosti njihovih formata. Na primer, većina grafičkih dizajnera ne radi neposredno sa SVG-om, nego SVG dokumente pravi pomoću svojih uobičajenih alatki, kao što je Adobe Illustrator. Oni ne moraju ni znati da koriste XML.

Ova knjiga se usredsređuje na standarde koji su važni gotovo svim programerima što rade u XML-u. Istražićemo XML tehnologije zajedničke velikom broju primena XML-a, a ne one koje su važne u samo nekoliko ograničenih domena.

## Šta je novo u trećem izdanju

XML nije stajao u mestu tokom dve godine protekle od drugog izdanja knjige *XML in a Nutshell*. Najvidljivija promena u ovom izdanju jeste to što ono opisuje XML 1.1. Međutim, izmene u verziji 1.1 nisu toliko velike kao što bi novi broj verzije (.1) implicirao. Zapravo, ako ne govorite mongolski, burmanski, amharski, kambodžanski ili još nekoliko manjih jezika, u XML-u 1.1 ima veoma malo novog materijala zanimljivog za vas. XML 1.0 i 1.1 isti su u gotovo svakom praktičnom pogledu. Razlika između njih je sigurno mnogo manja od razlike između Java 1.0 i 1.1. Zato ćemo u ovoj knjizi razmatrati XML kao jedan entitet, a na XML 1.1 eksplicitno upućivati samo u retkim prilikama kada se te verzije zaista razlikuju. Verovatno se otprilike 98% knjige jednako dobro odnosi i na XML 1.0 i na XML 1.1.

Dodali smo i poglavlje posvećeno XIncludeu, novijem pronalasku konzorcijuma W3C; koristi se za pravljenje velikih XML dokumenata kombinovanjem manjih dokumenata i njihovih delova. Elliotte je napisao gotovo polovinu prvih realizacija XIncludea, a i prvu knjigu napravljenu pomoću XIncludea, pa nam je ova tema posebno značajna. Ostala poglavlja knjige prerađena su koliko je zahtevao uticaj verzije 1.1 XML-a na obrađene teme, kao i uticaj promena datih tehnologija u pomenute dve godine. Mnoge teme su ažurirane tako da odgovaraju poslednjoj verziji odgovarajuće specifikacije, i to:

- SAX 2.01
- Prostor imena 1.1
- DOM Level 3
- XPointer 1.0
- Unicode 4.01

I najzad, u knjizi su ispravljene mnoge male greške i propusti.

## Organizacija knjige

Deo I, *Koncepti XML-a*, uvodi osnovne standarde koji čine jezgro XML-a i kojih se moraju pridržavati sve XML aplikacije i softver. U njemu ćete ukratko saznati šta je dobro oblikovan XML dokument, i šta su definicija tipa dokumenta (DTD), prostor imena i Unicode.

Deo II, *Narativni dokumenti*, istražuje tehnologije koje se uglavnom upotrebljavaju za XML dokumente narativnog tipa, kao što su Web stranice, knjige, članci, dnevni- ci i drame. Saznaćete šta su i kako rade XSLT, CSS, XSL-FO, XLinks, XPointers, XPath, XInclude i RDDL.

Među najneočekivanim događajima u istoriji XML-a bilo je to što je široko prihvaćen za izradu strukturiranih dokumenata punih podataka, kao što su pro- računske tabele (engl. *spreadsheets*), finansijski statistički podaci, matematičke tabele i formati datoteka. Deo III, *Dokumenti slični zapisima baze podataka*, istražuje upotrebu XML-a u takvim primenama. U njemu ćemo se usredsrediti na alatke i interfejs za programiranje aplikacija koji su potrebni za pisanje softvera za obradu XML-a – a to su SAX, DOM i šeme.

Najzad, deo IV, *Referentni pregled osnovnih XML tehnologija*, predstavlja niz referent- nih poglavlja koja čine jezgro svake knjige iz ove edicije. U njima su data detaljna sintaktička pravila za osnovne XML tehnologije, među kojima su XML, DTD-ovi, šeme, XPath, XSLT, SAX i DOM. Pređite u taj deo knjige kada hoćete da brzo pronađete preciznu sintaksu nečega što znate da možete da uradite, ali ne znate kako.

## Konvencije korišćene u knjizi

Tekst konstantne širine koristi se za:

- Sve što se može pojaviti u XML dokumentu, kao što su imena elemenata, ozna- ke (engl. *tags*), vrednosti atributa, reference entiteta i instrukcije za obradu.
- Sve što se može pojaviti u programu, kao što su rezervisane reči, operatori, ime- na metoda, imena klasa i literali.

**Polucrni tekst konstantne širine** koristi se za:

- Ono što unosi korisnik.
- Naglašavanje u primerima i odlomcima koda.

*Kurziv konstantne širine* upotrebljen je za:

- Zamenljive elemente u programskom kodu.

*Kurziv* se upotrebljava za:

- Nove termine na mestu definicije i termine na engleskom.
- Naglašavanje u telu teksta.
- Putanje, imena datoteka i programa. (Međutim, ako je ime programa ujedno i ime Javine klase, napisano je fontom konstantne širine, kao i ostala imena klasa.)
- Imena domena i mrežnih računara (*cafeconleche.org*).



Ova sličica ukazuje na savet, predlog ili opštu napomenu.



Ova sličica ukazuje na upozorenje ili opomenu.

Značajni delovi koda, celi programi i dokumenti obično su ovako smešteni u zaseban pasus:

```
<?xml version="1.0"?>
<?xml-stylesheet href="person.css" type="text/css" ?>
<osoba>
  Alen Tjuring
</osoba>
```

XML uzima u obzir razliku između velikih i malih slova. Element `OSOBA` nije jednak elementu `osoba`, niti elementu `Osoba`. Pisci na jezicima koji uzimaju u obzir razliku između velikih i malih slova ne mogu uvek da se drže gramatičkih pravila standardnog engleskog jezika. Obično se rečenica može preformulisati tako da se poštuju pravila i XML-a i engleskog, i mi smo se potrudili da tako uradimo kad god je bilo moguće. Međutim, u retkim slučajevima problem nije bilo moguće zaobići, i tada je engleski lošije prolazio.

Premda je većina primera upotrebljenih u knjizi pojednostavljena da bi se mogla još negde upotrebiti, nekoliko njih ima stvarnu vrednost. Slobodno ih upotrebite u svom kodu – cele ili samo delove. Za to vam nije potrebna specijalna dozvola. Što se nas tiče, oni su javno vlasništvo (što svakako ne važi za tekst objašnjenja).

## Komentari čitalaca

Biće nam drago da čujemo opšte komentare čitalaca o tome kako bi se knjiga mogla poboljšati – konkretnim ispravkama ili temama koje bi trebalo obuhvatiti. Adrese elektronske pošte autora knjige jesu [elharo@metabolab.unc.edu](mailto:elharo@metabolab.unc.edu) i [smeans@ewm.biz](mailto:smeans@ewm.biz). Molimo vas, uzmite u obzir da nas dvojica svakodnevno primamo više stotina elektronskih poruka i da ne možemo odgovoriti na svaku. Najveće šanse da dobiju odgovor imaju oni koji se predstavljaju kao čitaoci ove knjige. Takođe vas molimo da svoje poruke šaljete s naloga e-pošte na koji treba poslati odgovor i u kome je adresa za odgovor tačno zadata. Ništa nas toliko ne obeshrabruje kao kada nam se, nakon što utrošimo sat i više vremena na proučavanje i pisanje detaljnog odgovora na zanimljivo pitanje, odgovor vrati zato što je sagovornik svoju poruku poslao s javnog

terminala, a prethodno se nije potrudio da upiše svoju pravu adresu e-pošte među parametre čitača Weba.

Proverili smo i testirali informacije koje navodimo u knjizi, ali se ipak može desiti da se mogućnosti opisanih tehnologija promene (ili ćete pronaći grešku u našem opisu). Pristalice smo stare izreke „Ako vam se knjiga sviđa, recite to svojim prijateljima. Ako vam se ne sviđa, recite to nama“. Naročito bismo hteli da nam skrenete pažnju na eventualne greške. Koliko god da su se pisci i redakcija potrudili oko ove knjige, mora biti da nam je ipak promaklo nekoliko grešaka. Ako pronađete neku, molimo vas da nas obavestite kako bismo mogli da je ispravimo u sledećem izdanju. Sve podatke o greškama molimo da pošaljete neposredno autorima ove knjige na prethodno navedene adrese e-pošte.\*

Otvorena je i Web lokacija originala ove knjige, na kojoj su navedene pronađene greške, primeri i dodatne informacije. Adresa te lokacije je:

<http://www.cafeconleche.org/books/xian3/>

Pre nego što utrošite vreme na prijavu greške, molimo vas da posetite ovu lokaciju i pogledate da li je tamo već objavljena ispravka. Tehnička pitanja i komentare šalјite e-poštom neposredno autorima knjige ili izdavačima, na adresu:

[bookquestions@oreilly.com](mailto:bookquestions@oreilly.com), odnosno, [redakcija@mikroknjiga.co.yu](mailto:redakcija@mikroknjiga.co.yu)

Više informacija o drugim izdanjima, softveru i ostalim resursima koje O'Reilly stavlja na raspolaganje javnosti, potražite na adresama:

<http://www.oreilly.com>

<http://xml.oreilly.com>

<http://www.xml.com>

Više informacija o drugim izdanjima i aktivnostima Mikro knjige potražite na adresi:

[www.mikroknjiga.co.yu](http://www.mikroknjiga.co.yu)

---

\* Većina primera je prevedena, pa je moguće da su se i tu potkrale greške. Ako primetite grešku u prevedenom izdanju, molimo vas da pošaljete poruku na adresu [redakcija@mikroknjiga.co.yu](mailto:redakcija@mikroknjiga.co.yu).