

Introduzione

Una delle caratteristiche che maggiormente colpiscono di XML è il fatto che, nonostante venga definito un “linguaggio”, in realtà non lo è, almeno non nell’accezione generale del termine.

Si tratta invece di qualcosa di più, di un insieme di regole e convenzioni che consentono di descrivere qualunque linguaggio di markup, basato quindi su marcatori, per le esigenze del momento, siano esse descrittive e progettuali oppure prettamente operative. L’idea non è certo nuova ed è in parte derivata dai concetti di base di SGML (*Standard Generalized Markup Language*), che aveva esattamente questi obiettivi. Il vantaggio di XML su SGML è dato dalla semplicità con la quale in XML è possibile definire una struttura gerarchica per le informazioni, semplicità che si traduce in velocità di apprendimento da parte dell’utente e quindi in maggiore possibilità di diffusione.

Con XML è possibile scrivere il proprio linguaggio definendo, anche se ciò non è obbligatorio, le relazioni di gerarchia tra le varie componenti, fino ad arrivare alla determinazione di vere e proprie grammatiche che consentono di verificare la correttezza semantica dei documenti XML.

La possibilità di creare linguaggi personalizzati si sposa inoltre con la caratteristica propria di XML di poter essere trasformato in qualunque altra cosa: una differente applicazione XML, una pagina web attraverso (X)HTML, un documento PDF o PostScript e, in generale, ogni formato per il quale esista un filtro di trasformazione.

Ecco perché XML viene utilizzato sia facendo riferimento ai tanti linguaggi che sono stati messi a disposizione dagli organismi che definiscono gli standard internazionali (come per esempio il W3C, che ha scritto le specifiche XHTML), sia come linguaggio di interscambio di informazioni tra applicazioni diverse per le quali sono state definite strutture grammaticali proprietarie.

Scrivere in XML significa da un lato garantirsi la possibilità di creare ed estendere il proprio linguaggio, dall’altro avere a disposizione il supporto di strumenti sempre più evoluti per l’editing, la validazione della grammatica e la trasformazione dei documenti in formati diversi.

Organizzazione dell'opera

La successione dei capitoli guida il lettore nella comprensione di tutte le caratteristiche sintattiche di XML, nella definizione di linguaggi personalizzati e nella trasformazione dei documenti XML in file di tipo differente.

- Il Capitolo 1 è una panoramica delle caratteristiche di XML: dalle motivazioni che sono alla base della sua creazione, alle regole e alla sintassi elementare. Si accenna alla differenza tra documenti *ben formati* e documenti *validi*, concetti che saranno ripresi più avanti nel testo.
- Il Capitolo 2 illustra le regole sintattiche fondamentali di XML, dalla struttura del documento a tutte le sue componenti obbligatorie e facoltative.
- Il Capitolo 3 tratta delle grammatiche che è possibile creare per regolamentare la presenza e la gerarchia degli elementi. Viene anche affrontato il processo di validazione che verifica la coerenza di un documento XML rispetto a una determinata grammatica.
- Il Capitolo 4 descrive come inserire in una grammatica DTD la gestione delle entità che consentono a XML di governare tutte le forme sintattiche, anche quelle che prevedono l'utilizzo di caratteri speciali.
- Il Capitolo 5 completa il discorso sulle grammatiche spiegando come sia possibile stabilire delle regole sintattiche non soltanto per gli elementi XML, ma anche per i loro attributi.
- Il Capitolo 6 affronta il tema dei *namespace*, i dizionari che consentono di utilizzare in un documento XML elementi che provengono da librerie diverse senza conflitti, nemmeno nel caso in cui si tratti di elementi identici.
- Il Capitolo 7 è dedicato alle trasformazioni XSL che permettono di applicare a un documento XML un foglio di stile in grado di trasformare il documento in un testo differente.
- Il Capitolo 8 affronta il tema della formattazione dei documenti XML utilizzando opportuni fogli di stile e strumenti di trasformazione che consentono di generare, per esempio, documenti in formato PDF a partire da documenti XML.
- Il Capitolo 9 descrive XPath, lo strumento sintattico in grado di navigare all'interno di un documento XML alla ricerca di particolari elementi o attributi.
- Il Capitolo 10 affronta XLink e XPointer, che costituiscono la modalità per collegare tra loro documenti XML diversi.
- Il Capitolo 11 tratta di XHTML, forse la più importante e diffusa applicazione di XML, che permette la conversione in XML delle varie grammatiche HTML esistenti.
- Il Capitolo 12 è dedicato a TEI-Lite, un'applicazione di XML composta da una grammatica corredata da una serie di strumenti per esprimere in XML l'organizzazione semantica e strutturale di un testo scritto generico.
- Il Capitolo 13 descrive altre due interessanti applicazioni XML: si tratta di MathML, per la definizione di espressioni e formule matematiche visualizzabili con un browser web, e di RSS, un'applicazione di XML in grado di trasmettere notizie ed estratti di contenuti da un'applicazione a un'altra.
- L'Appendice descrive infine l'integrazione fra PHP e XML.

Finalità

Questo libro affronta XML partendo dai concetti sintattici di base per poi passare ad argomenti più complessi come le grammatiche DTD, i processi di validazione e le funzionalità di trasformazione e presentazione.

Ogni capitolo è corredato da esempi concreti e riferimenti agli strumenti utilizzati che consentono al lettore di verificare e toccare con mano il tema illustrato.

La seconda parte del libro, in particolare, introduce il lettore alla conoscenza di alcuni utilizzi pratici di XML, mostrando come sia possibile utilizzare gli elementi sintattici e semantici visti in precedenza per creare applicazioni e implementazioni robuste basate su XML.

Il libro è dedicato a chi vuole avere una panoramica di XML che spazi dagli aspetti sintattici all'espressività semantica, fino ad arrivare alle applicazioni reali che si possono realizzare.

Documentazione online

Di seguito sono elencati gli indirizzi Internet dove consultare le specifiche complete di XML e delle applicazioni XML presentate nel corso del libro.

XML 1.0, <https://www.w3.org/TR/REC-xml>

XSL-FO 1.0, <https://www.w3.org/TR/xsl>

XSLT 1.0, <https://www.w3.org/TR/xslt>

XHTML 1.0, <https://www.w3.org/TR/xhtml1>

XPath 1.0, <https://www.w3.org/TR/xpath>

XLink 1.0, <https://www.w3.org/TR/xlink>

XPointer, <https://www.w3.org/TR/xptr>

TEI-Lite, https://www.tei-c.org/Lite/teiu5_it.htm

MathML 2.0, <https://www.w3.org/TR/MathML2>

RSS 1.0, <https://web.resource.org/rss/1.0>

Convenzioni utilizzate nel libro

Nel libro sono presenti alcune etichette che identificano tipi particolari di informazioni relative agli argomenti trattati.

ATTENZIONE

Indicazioni da non trascurare per evitare le difficoltà in cui gli utenti, soprattutto alle prime armi, possono imbattersi.

NOTA

Una nota contiene informazioni interessanti, talvolta tecniche, relative all'argomento trattato. Talvolta riporta approfondimenti e curiosità.