**New Perspectives XML Comprehensive, 3e**

**Tutorial 1: Creating an XML Document**

**Key Terms**

**Aggregator** *See* Feed reader. (XML 9)

**Attribute** Stores additional information about an element and has a value surrounded by a matching pair of single or double quotes. (XML 22)

**Cascading Style Sheets (CSS)** The style sheet language developed for use with HTML on the web. *Also called* CSS. (XML 45)

**CDATA section** A large block of text that XML treats as character data only; CDATA is not processed, it is treated as pure data content. (XML 22)

**Character data** A short memorable name used in place of a numeric character reference. *Also called* entity reference. (XML 39)

**Character entity reference** A short memorable name used in place of a numeric character reference. *Also called* entity reference. (XML 36)

**Character reference** *See* numeric character reference. (XML 23)

**Chemical Markup Language (CML)** An XML vocabulary that codes molecular information. *Also called* CML. (XML 8)

**Child element** An element that is nested completely within another element, known as its parent element. (XML 22)

**Closing tag** The XML code that follows the content of an element. (XML 24)

**CML** *See* Chemical Markup Language. (XML 8)

**Comment lines** Optional parts of the prolog of an XML document thatprovide additional information about the document contents. (XML 11)

**Compound document** An XML document composed of elements from other vocabularies or schemas, such as from HTML and XML vocabularies. (XML 50)

**CSS** *See* Cascading Style Sheets (CSS). (XML 45)

**Default namespace** The namespace that applies to any descendant element or attribute unless a different namespace is declared within one of the child elements. (XML 49)

**Document body** The part of an XML filethat contains the document content in a hierarchical tree structure. (XML 2)

**Document element** *See* Root element. (XML 26)

**Document type declaration (DTD)** The statement included in the prolog of an XML document thatprovides information about the rules used in the XML document’s vocabulary. *Also called* DTD. (XML 2)

**Element** A building block of an XML file that contains data to be stored in the document. (XML 24)

**Empty element** *See* open element. (XML 25)

**Entity reference** *See* character entity reference. (XML 36)

**Epilog** The part of an XML file that contains any final comment lines and processing instructions. (XML 2)

**Extensible Hypertext Markup Language (XHTML)** A style sheet language developed specifically for XML; an XSL style sheet must follow the rules of well-formed XML because XSL is an XML vocabulary. *Also called* XSL. (XML 1)

**Feed reader** A software program used to provide periodic updates to subscribers of an RSS feed. *Also called* aggregator. (XML 9)

**ISO/IEC character set** An international numbering system for referencing characters from virtually any language. (XML 36)

**Namespace** A defined collection of element and attribute names. (XML 23)

**Nested** Describes an element that is contained within another element. (XML 25)

**Numeric character reference** The numeric version of a reference used to enter a character symbol. *Also called* character reference. (XML 35)

**One-sided tag** A tag used to enter an open element. (XML 25)

**Open element** An element with no content. (XML 25)

**Opening tag** The XML code that precedes the content of an element. (XML 24)

**Parent element** An element that has another element, known as its child element, nested completely within it. (XML 22)

**Parsed character data (PCDATA)** All those characters that XML treats as parts of the code of an XML document. *Also called* PCDATA. (XML 22)

**Parser** *See* XML parser. (XML 16)

**PCDATA** *See* Parsed character data (PCDATA). (XML 22)

**Processing instructions** The optional statement(s) included in the prolog of an XML document thatprovide additional instructions to be run by programs that read the XML document. (XML 11)

**Processor** *See* XML parser. (XML 16)

**Prolog** The part of an XML document that contains the XML declaration, optional comment lines, optional processing instructions, and an optional document type declaration. (XML 2)

**Really Simple Syndication (RSS)** The language used for distributing news articles and any content that changes on a regular basis. (XML 9)

**Root element** The statement in an XML document that contains all the other elements in the document. *Also called* document element. (XML 22)

**Schema** Rules that specifically control what code and content a document may include. (XML 2)

S**GML** *See* Standard Generalized Markup Language (SGML). (XML 4)

**Sibling elements** Elements that are children to the same parent element. (XML 22)

**Standard Generalized Markup Language (SGML)** A language introduced in the 1980s that describes the structure and content of any machine-readable information. SGML is device-independent and system-independent. *Also called* SGML. (XML 4)

**Standard vocabulary** A set of XML tags for a particular industry or business function. (XML 8)

**Unicode** A computing industry standard for the consistent encoding, representation, and handling of text expressed in most of the world’s written languages; can be implemented using different character encodings, with the most commonly used encodings being UTF-8 and UTF-16. (XML 12)

**Uniform Resource Identifier (URI)** A text string that uniquely identifies a resource; one version of a URI is the Uniform Resource Locator (URL), which is used to identify a location on the Web, and the other is a Uniform Resource Name (URN), which provides a persistent name for a resource, independent of that resource’s location. *Also called* URI. (XML 23)

**Uniform Resource Locator (URL)** A URI used to identify the location of a resource on the web. *Also called* URL. (XML 23)

**Uniform Resource Name (URN)** A URI that provides a persistent name for a resource, independent of that resource’s location. *Also called* URN. (XML 23)

**URI** *See* Uniform Resource Identifier. (XML 23)

**URL** *See* Uniform Resource Locator. (XML 23)

**URN** *See* Uniform Resource Name. (XML 23)

**Valid document** An XML document that is well formed and satisfies the rules of a DTD or schema. (XML 11)

**W3C** *See* World Wide Web Consortium (W3C). (XML 4)

**Well-formed document** An XML document that has no syntax errors and satisfies the general specifications for XML code defined by the World Wide Web Consortium (W3C). (XML 3)

**White space** Nonprintable characters such as spaces, new line characters, or tab characters. (XML 40)

**White space stripping** A step in the process of parsing HTML pages that results in consecutive occurrences of white space being treated as a single space. (XML 40)

**World Wide Web Consortium (W3C)** The organization that develops and maintains the standards for XML. *Also called* W3C. (XML 4)

**XHTML** *See* Extensible Hypertext Markup Language (XHTML). (XML 10)

**XML** *See* Extensible Markup Language. (XML 1)

**XML application** *See* XML vocabulary. (XML 7)

**XML declaration** The statement at the start of the prolog, which signals to the program reading the file that the document is written in XML and provides information about how that code is to be interpreted by the program. (XML 2)

**XML parser** A program that interprets an XML document’s code and verifies that it satisfies the W3C specifications. (XML 3)

**XML processor** *See* XML parser. (XML 3)

X**ML vocabulary** A markup language tailored to contain specific pieces of information. (XML 7)

**XSL** *See* Extensible Stylesheet Language (XSL). (XML 48)