

XML and DTD

Prof. Dr. Raphael Volz

Hochschule Pforzheim

Contents

1	XML	5
1.1	eXtensible Markup Language	5
1.2	XML Example	5
1.3	XML Elements	6
1.4	XML Example 2	6
1.5	XML Attributes	7
1.6	Empty elements	7
2	DTD	9
2.1	Document Type Definition	9

Chapter 1

XML

1.1 eXtensible Markup Language

- World-Wide Web Consortium (W3C) Standard
 - Recommendation since February 1998
- Simple, very flexible, text-based data format
- Original purpose: structuring texts
 - XHTML web pages are XML documents
 - Unlike HTML elements are not predefined
- XML is one of the most widely-used formats for sharing structured information today, used for
 - data exchange between applications
 - storing data (of an application)
- omnipresent in enterprise applications

1.2 XML Example

```
<?xml version="1.0" encoding="UTF-8"?>
<author>
  <title>Sir</title>
  <lastname>Barrie</lastname>
  <firstname>James</firstname>
  <firstname>Matthew</firstname>
  Best known for his book Peter Pan,
  or The Boy Who Wouldn't Grow Up
</author>
```

1.3 XML Elements

```
<?xml version="1.0" encoding="UTF-8"?>
<author>
  <title>Sir</title>
  <lastname>Barrie</lastname>
  <firstname>James</firstname>
  <firstname>Matthew</firstname>
  Best known for his book Peter Pan,
  or The Boy Who Wouldn't Grow Up
</author>
```

- *author*, *title*, *lastname*, and *firstname* are called **elements**
- *author* is called **root** (element)
- Elements can be nested
- An element starts with opening tag, contains content (possibly other elements) and closes with closing tag repeating the same name as in the opening tag

1.4 XML Example 2

```
<?xml version="1.0" encoding="UTF-8"?>
<bookstore>
  <book category="cooking">
    <title lang="en">Everyday Italian</title>
    <author>Giada De Laurentiis</author>
    <year>2005</year>
    <price>30.00</price>
  </book>
  <book category="children">
    <title lang="en">Harry Potter</title>
    <author>J K. Rowling</author>
    <year>2005</year>
    <price>29.99</price>
  </book>
  <book category="web">
    <title lang="en">Learning XML</title>
    <author>Erik T. Ray</author>
    <year>2003</year>
    <price>39.95</price>
  </book>
</bookstore>
```

Source: XML tutorial

1.5 XML Attributes

1.6 Empty elements

`{xml}<firstname/>` is the same as `{xml}<firstname></firstname>`

- Empty elements do not have content
- Empty elements can have attributes
- Examples from HTML
 - `
` (line break)
 - `<hr/>` (horizontal line)
 - `` (Image) ## XML example (2) ## Vocabulary (XML)

A vocabulary refers to the collection of all names of elements and attributes in an XML document

- Many Standard vocabularies exist
 - for industries
 - for specific purposes
- Examples
 - financial reporting (XRML)
 - geographic data (GPX / KML)
 - Universal Plug and Play (UPnP)
 - vector graphics (SVG)
 - XHTML
 - ...

Chapter 2

DTD

2.1 Document Type Definition

```
twelve = "1", "2";  
non-zero-number = "1" | "2" | "3" | "4" | "5" | "6" | "7" | "8" | "9" ;  
digit = "0" | non-zero-number ;  
natural-number = non-zero-number, { Digit } ;  
integer = "0" | [ "-" ], natural-number ;
```