# XML基础

## Xml基本介绍

问题:什么是xml,它的作用是什么？

XML 指可扩展标记语言

所谓的可扩展标记语言，简单说就是指xml文件中的标签可以任意定义。

XML 被设计用来传输和存储数据。

问题:xml它存储的是什么样数据？

Xml中存储的是有关系的数据。

<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>

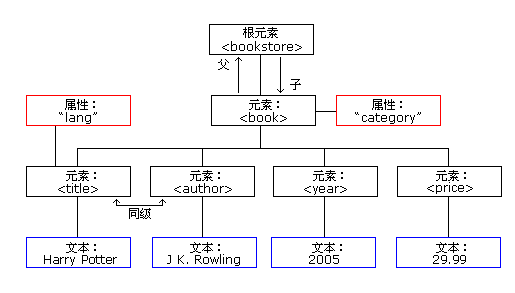
<author>Erik T. Ray</author>

<year>2003</year>

<price>39.95</price>

</book>

</bookstore>

xm 

Xml也可以形成一个树型结构，它里面每一个叫做节点，节点也分成了

元素结点(Element) 属性节点(Attribute) 文本节点(Text) 整体也是一个节点，它叫做document.

Xml在实际开发中，它的真正作用?

1. 将xml做为配置文件
2. 通过xml可以在不同的系统之间传递数据

## Xml基本组成

首先，xml文件它的后缀名就是.xml,我们可以直接使用文本编辑器来对xml文件进行处理。

在xml文件中它有以下几部分

1. 文档声明
2. 属性
3. 元素
4. 注释
5. CDATA
6. PI指令(了解一下)

### 文档声明

它的主要作用就是说明当前文件是一个xml



文档声明有三个属性，而version encoding是常用的。还有一个属性叫做standalone它不常用。

1. version 它是版本 它的值常用的是1.0
2. encoding它的主要作用是声明当前xml文件的字符编码
3. standalone 它的作用是描述文档是否是一个独立文档

### 属性

在xml文件中属性它的值必须使用引号引起来(双引号与单引号都可以)。

在实际开发中，经常会使用子标签来替换属性。

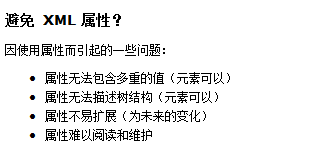
<a type=”xxx”></a>

可以使用子标签来替换属性

<a>

<type>xx</type>

</a>



### 元素

元素就是指标签

在xml中所有的标签必须有结束。

有开始有结束<a></a>，如果标签没有标签体可以自关闭<a/>

格式良好的xml文件，它有且仅有一个根标签。其它所有的标签都是这个根标签的子孙元素

Xml中的元素是可以嵌套的，但是不能交叉嵌套。

在xml文件中空格换行都会被视频内容解析。

关于标签名称的规范:

1. xml区分大小写 <a> <A>这是两个标签。
2. 不能以数字或”\_”(下划线)开头
3. 不能以xm(Xml XML)开头
4. 标签中不能包含空格
5. 名称中间不能包含(:)冒号

### 注释

在xml文件中它的注释与html中一样

<!—注释内容 -->

在xml文件中第一句话只能是文档声明，注释都不可以。

注释不可以嵌套。

### CDATA区域

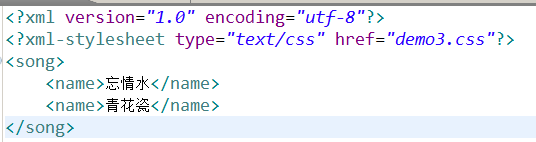
CDATA区域中的内容不会被解析，而是原样展示。

在xml中也存在一些特殊字符.

&lt; <

&gt; >

### PI指令



### Xml语法规则总结

1. 所有的 xml标签必须关闭
2. xml标签对大小写敏感
3. xml标签必须正确嵌套
4. xml文档有且只有一个根元素
5. xml属性必加引号引起。
6. xml中的空格会被保留，对于特殊的字符必须转义才可以使用

## Xml约束介绍

在xml技术中，我们可以编写一个文档来约束一个xml文件的书写的内容规范，这个就称之为xml约束。

简单说，约束就规定了xml文件中可以写什么，不可以写什么。

为什么使用约束?

因为xml书写太随意，在对信息进行描述时，如果没有约束，我们是不知道，xml文件如果定义信息。所以必须规定xml文件中可以写什么，也就是定义其规范，这样才可以通过xml文件进行传输信息。

在xml技术中有哪些xml约束?

1. DTD约束

文档类型定义（DTD）可定义合法的XML文档构建模块

Document type definition

Dtd文件后缀名就是dtd

1. SCHEMA约束

Schema文件它就是一个xml文件，只不过它的文件后缀名是xsd.

XML Schema 是基于 XML 的 DTD 替代者。

XML Schema 描述 XML 文档的结构。

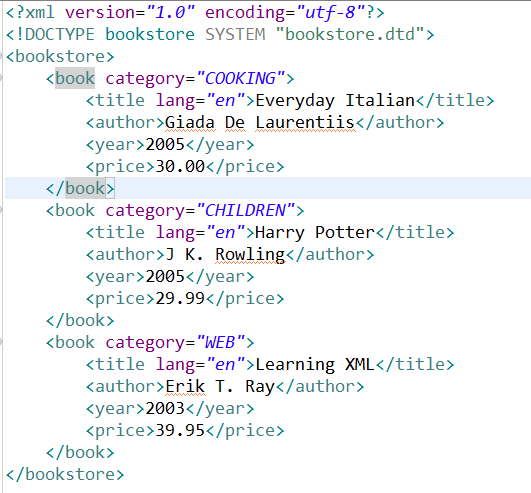
XML Schema 语言也称作 XML Schema 定义（XML Schema Definition，XSD）。

## DTD约束

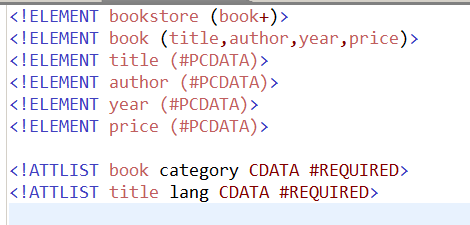
Dtd约束文件的后缀名是.dtd.

Dtd快速入门

1. 创建一个文件



1. 创建一个dtd文件



## DTD约束详解

### dtd元素

dtd元素声明格式

方式1：<!ELEMENT 元素名称 类别>

方式2：<!ELEMENT 元素名称 (元素内容)>

在dtd中的特殊符号：

+ 表示可以出现一次或多次

\*表示可以出现0次或多次

? 表示可以出现，但只能出现一次

（）表示对元素进行分组

｜ 表示在列出的元素中选择一个

， 表示对象必须是按照指定的顺序出现

元素的类别

#PCDATA 表示可以包含任意字符，但不能是子元素

ANY 表示可以是任意内容

EMPTY 表示是一个空内容，不可以包含任何内容

### dtd属性

属性声明使用下列语法：

<!ATTLIST 元素名称 属性名称 属性类型 默认值>

**以下是属性类型的选项：**

|  |  |
| --- | --- |
| **类型** | **描述** |
| CDATA | 值为字符数据 (character data) |
| (*en1*|*en2*|..) | 此值是枚举列表中的一个值 |
| ID | 值为唯一的 id |
| IDREF | 值为另外一个元素的 id |
| IDREFS | 值为其他 id 的列表 |
| NMTOKEN | 值为合法的 XML 名称 |
| NMTOKENS | 值为合法的 XML 名称的列表 |
| ENTITY | 值是一个实体 |
| ENTITIES | 值是一个实体列表 |
| NOTATION | 此值是符号的名称 |
| xml: | 值是一个预定义的 XML 值 |
|  |  |
|  |  |

**默认值参数可使用下列值：**

|  |  |
| --- | --- |
| **值** | **解释** |
| 值 | 属性的默认值 |
| #REQUIRED | 属性值是必需的 |
| #IMPLIED | 属性不是必需的 |
| #FIXED value | 属性值是固定的 |

类型主要介绍三个：

1. CDATA 就代表一个字符串

2.(*en1*|*en2*|..) 枚举类型，简单说就是可以从其中任意选择一个值

3.ID 唯一值