# XML知识点总结

# 1817700432 杜欣航

# XML基础

XML被设计用来传输和存储数据

## 1.什么是xml

XML 指可扩展标记语言（EXtensible Markup Language）

XML 是一种标记语言，很类似 HTML

XML 的设计宗旨是传输数据，而非显示数据

XML 是对 HTML 的补充。

XML 不会替代 HTML，理解这一点很重要。在大多数 Web 应用程序中，XML 用于传输数据，而 HTML 用于格式化并显示数据。

对 XML 最好的描述是：

XML 是独立于软件和硬件的信息传输工具。

XML 标签没有被预定义。您需要自行定义标签。

XML 被设计为具有自我描述性。

XML 是 W3C 的推荐标准

## 2. XML 与 HTML 的主要差异

XML 不是 HTML 的替代。

XML 和 HTML 为不同的目的而设计：

XML 被设计为传输和存储数据，其焦点是数据的内容。

HTML 被设计用来显示数据，其焦点是数据的外观。

HTML 旨在显示信息，而 XML 旨在传输信息。

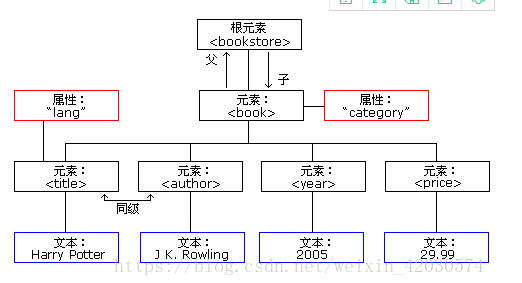
## 3. XML的用途

XML能把数据从HTML分离，可以简化数据共享，简化数据传输，简化平台的变更，使数据更有用，还可用于创建新的 Internet 语言，并且XML可以用来标记数据、定义数据类型，是一种允许用户对自己的标记语言进行定义的源语言。 它非常适合万维网传输，提供统一的方法来描述和交换独立于应用程序或供应商的结构化数据。XML技术已经广泛应用于e-Learning应用系统的开发,大多数的商用e-Learning平台都支持XML标准。一些主要的网络设备制造商，如CISCO、JUNIPER等，生产的网络设备也已提供了对XML的支持，以利于今后基于XML的网络管理。

## 4. XML树结构

XML文档必须包含根元素。该元素是所有其他元素的父元素。

XML文档中的元素形成了一棵文档树。这棵树从根部开始，并扩展到树的最底端。所有的元素都可以有子元素：



**例如：**

<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>

**其中**：

根元素是 <bookstore>。文档中的所有 <book> 元素都被包含在 <bookstore> 中。

<book> 元素有 4 个子元素：<title>、< author>、<year>、<price>

## 4. XML语法规则

**1.所有 XML 元素都须有关闭标签**

<p>This is a paragraph</p>

**2.XML 标签对大小写敏感**

<message>这是正确的。</message>

**3.XML 必须正确地嵌套**

<b><i>This text is bold and italic</i></b>

**4.XML 文档必须有根元素**

<root>

<child>

<subchild>.....</subchild>

</child>

</root>

**5.XML 的属性值须加引号**

<note date="08/08/2008">

<to>George</to>

<from>John</from>

</note>

**6.XML 中的注释**

<!-- This is a comment -->

**7.在 XML 中，空格会被保留**

## 5. XML元素

**1.什么是 XML 元素？**

XML 元素指的是从（且包括）开始标签直到（且包括）结束标签的部分。

元素可包含其他元素、文本或者两者的混合物。元素也可以拥有属性。

**例如：**

<bookstore>

<book category="CHILDREN">

<title>Harry Potter</title>

<author>J K. Rowling</author>

<year>2005</year>

<price>29.99</price>

</book>

</bookstore>

在上例中，<bookstore> 和 <book> 都拥有元素内容，因为它们包含了其他元素。<author> 只有文本内容，因为它仅包含文本。

在上例中，只有 <book> 元素拥有属性

**2.XML 命名规则**

XML 元素必须遵循以下命名规则：

名称可以含字母、数字以及其他的字符

名称不能以数字或者标点符号开始

名称不能以字符 “xml”（或者 XML、Xml）开始

名称不能包含空格

可使用任何名称，没有保留的字词。

**3.XML 元素是可扩展的**

## 6. XML属性

**1.XML 属性必须加引号**

属性值必须被引号包围，不过单引号和双引号均可使用。比如一个人的性别，person 标签可以这样写：

<person sex="female">

**2.避免 XML 属性**

因使用属性而引起的一些问题：

属性无法包含多重的值（元素可以）

属性无法描述树结构（元素可以）

属性不易扩展（为未来的变化）

属性难以阅读和维护

**3.针对元数据的 XML 属性**

有时候会向元素分配 ID 引用。这些 ID 索引可用于标识 XML 元素，它起作用的方式与 HTML 中 ID 属性是一样的。

## 7. XML验证

拥有正确语法的 XML 被称为“形式良好”的 XML。通过 DTD 验证的 XML 是“合法”的 XML。形式良好的 XML 文档，“形式良好”或“结构良好”的XML 文档拥有正确的语法**。**“形式良好”（Well Formed）的 XML 文档会遵守前几章介绍过的 XML 语法规则：

XML 文档必须有根元素

XML 文档必须有关闭标签

XML 标签对大小写敏感

XML 元素必须被正确的嵌套

XML 属性必须加引号