

## XML CDATA的作用 - smodi - 博客园

操作XML文件时，如果允许用户输入内容，例如：“< ”、“>”、“/”、“”等，当生成XML时，会破坏了XML结构，使数据中断。

这就要用XML CDATA

在XML文档中的所有文本都会被解析器解析。

只有在CDATA部件之内的文本会被解析器忽略。

---

### 解析数据

XML 解析器通常情况下会处理XML文档中的所有文本。

当XML元素被解析的时候，XML元素内部的文本也会被解析：

```
< message>This text is also parsed< /message>
```

XML解析器这样做的原因是XML元素内部可能还包含了别的元素，象下面的例子，name元素内部包含了first和last两个元素：

```
< name>< first>Bill< /first>< last>Gates< /last>< /name>
```

解析器会认为上面的代码是这样的：

```
< name>
< first>Bill< /first>
< last>Gates< /last>
< /name>
```

---

### 转义字符

不合法的XML字符必须被替换为相应的实体。

如果在XML文档中使用类似“< ”的字符，那么解析器将会出现错误，因为解析器会认为这是一个新元素的开始。所以不应该象下面那样书写代码：

```
< message>if salary < 1000 then< /message>
```

为了避免出现这种情况，必须将字符“< ”转换成实体，象下面这样：

```
< message>if salary < 1000 then< /message>
```

下面是五个在XML文档中预定义好的实体：

< < 小于号

> > 大于号

& & 和

' ' 单引号

" " 双引号

实体必须以符号“&”开头，以符号“;”结尾。

注意：只有“< ”字符和“&”字符对于XML来说是严格禁止使用的。剩下的都是合法的，为了减少出错，使用实体是一个好习惯。

---

### CDATA部件

在CDATA内部的所有内容都会被解析器忽略。

如果文本包含了很多的“< ”字符和“&”字符——就象程序代码一样，那么最好把他们都放到CDATA部件中。

一个 CDATA 部件以“< ![CDATA[” 标记开始，以“]]>”标记结束：

```
< script>
< ![CDATA[
function matchwo(a,b)
{
if (a < b && a < 0) then
{
return 1
}
else
{
return 0
}
}
```

```
}  
]]>  
< /script>
```

在前面的例子中，所有在CDATA部件之间的文本都会被解析器忽略。

CDATA注意事项：

CDATA部件之间不能再包含CDATA部件（不能嵌套）。如果CDATA部件包含了字符“]]>” 或者“<![CDATA[”，将很有可能出错哦。  
同样要注意在字符串“]]>”之间没有空格或者换行符