

XML CDATA

XML 文档中的所有文本均会被解析器解析。

只有 CDATA 区段中的文本会被解析器忽略。

PCDATA - 被解析的字符数据

XML 解析器通常会解析 XML 文档中所有的文本。

当某个 XML 元素被解析时，其标签之间的文本也会被解析：

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

解析器之所以这么做是因为 XML 元素可包含其他元素，就像这个实例中，其中的 <name> 元素包含着另外的两个元素（first 和 last）：

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

而解析器会把它分解为像这样的子元素：

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

解析字符数据（PCDATA）是 XML 解析器解析的文本数据使用的一个术语。

CDATA - （未解析）字符数据

术语 CDATA 是不应该由 XML 解析器解析的文本数据。

像 "<" 和 "&" 字符在 XML 元素中都是非法的。

"<" 会产生错误，因为解析器会把该字符解释为新元素的开始。

"&" 会产生错误，因为解析器会把该字符解释为字符实体的开始。

某些文本，比如 JavaScript 代码，包含大量 "<" 或 "&" 字符。为了避免错误，可以将脚本代码定义为 CDATA。

CDATA 部分中的所有内容都会被解析器忽略。

CDATA 部分由 "<![CDATA[" 开始，由 "]]>" 结束：

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

在上面的实例中，解析器会忽略 CDATA 部分中的所有内容。

关于 CDATA 部分的注释：

CDATA 部分不能包含字符串 "]]>"。也不允许嵌套的 CDATA 部分。

标记 CDATA 部分结尾的 "]]>" 不能包含空格或换行。

[← XML 命名空间](#)

[XML 编码 →](#)

[✎ 点我分享笔记](#)