

Символьное содержимое

Используя XML, можно разметить абсолютно любой текст. Текст, вложенный в тег узла, называется символьным содержимым. Для анализатора Flash-плеера символьное содержимое является практически таким же узлом, как и вложенные в тег теги. Содержимое узла может быть смешанным: одновременно он может содержать как теги, так и текст. Если в текст входят теги пустых узлов, то они разрывают его на части, которые рассматриваются анализатором независимо друг от друга.

Текст узла может быть образован любыми символами, кроме двух. В нем не должно быть “<” (так как этот символ обозначает начало тега) и “&” (символ используется в ссылках на сущность). Данные символы должны быть заменены так называемыми ссылками на сущности, напоминающими escape-последовательности ActionScript. Символ “<” заменяется сущностью < символ “&” — &. Для большей читабельности документа закрывающую угловую скобку “>” также стоит обозначать при помощи ссылки на сущность — >. Пример:

```
<info>Для задания тегов используются символы “&lt;” и “&gt;”</info>
```

Полезными могут оказаться ссылки на сущности " и ', задающие, соответственно, двойную и одинарную кавычку. Они используются, если кавычки входят в текст значения атрибута.

Применяя ссылки на сущности, можно ввести в текст любой символ, которого нет на клавиатуре. Для этого нужно использовать запись вида &#N;, где N — номер символа в кодировке в десятичном формате. Чтобы задать символ по его шестнадцатеричному номеру, нужно создать ссылку на сущность по следующей схеме: &#xN;. Например, чтобы ввести греческую букву “σ”, нужно набрать “σ”.

XML-документ вполне может содержать HTML-текст или даже другой XML-документ. При этом количество символов “<” и “&” может быть большим, и заменить их все ссылками “>” и “<” будет довольно сложно. Облегчить работу в таких случаях можно, поместив текст в секцию CDATA. При этом он не будет разбираться анализатором.

Начало секции CDATA обозначается как “<![CDATA[”. Ее конец задается при помощи “]]”. Между этими блоками помещается текст, который нужно сделать не анализируемым. Пример:

```
<text>
```

```
<![CDATA[
```

```
<b>Содержимым <i>XML</i>-документа может быть <i>HTML</i>-текст</b>
```

```
]]>
```

```
</text>
```

Для большинства XML-анализаторов символы с номерами, меньшими 32, не являются полноценными (это так называемые пробельные и служебные символы). Поэтому, если набрать в теге только несколько пробелов, то анализатор посчитает, что такой тег не имеет символьного содержимого. Аналогично, если два слова в тексте разделяет несколько пробелов, то анализатор оставит только один, отбросив остальные.