

Лабораторная работа № 2. Создание валидных XML-документов

Валидные XML документы

Валидный XML документ не то же самое, что и синтаксически верный XML документ.

Первое правило для валидного XML документа то, что он должен быть синтаксически верным.

Второе правило – валидный XML документ должен соответствовать определенному типу документов.

Правила, определяющие допустимые элементы и атрибуты для XML документа, называются определениями документа или схемами документа.

С XML можно использовать различные типы определений документа:

- оригинальное определение типа документа (DTD)
- более новый тип определений, основанный на XML, – XML схема.

XML DTD

Цель DTD состоит в том, чтобы определить структуру XML документа. Это делается путем определения списка допустимых элементов.

В приведенном ниже примере

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE note SYSTEM "Note.dtd">
<note>
  <to>Анна</to>
  <from>Дмитрий</from>
  <heading>Напоминание</heading>
  <body>Не забудь обо мне в эти выходные!</body>
</note>
```

декларация DOCTYPE является ссылкой на внешний файл определений типов документа (DTD). Содержимое этого файла показано ниже.

```
<!DOCTYPE note [
  <!ELEMENT note (to,from,heading,body)>
  <!ELEMENT to (#PCDATA)>
  <!ELEMENT from (#PCDATA)>
  <!ELEMENT heading (#PCDATA)>
  <!ELEMENT body (#PCDATA)>
]>
```

Приведенное выше DTD интерпретируется следующим образом:

- !DOCTYPE note определяет, что корневым элементом документа является *note*;
- !ELEMENT note определяет, что элемент *note* содержит четыре элемента: *to*, *from*, *heading*, *body*;
- !ELEMENT to определяет, что элемент *to* должен быть типа "#PCDATA";

– !ELEMENT from определяет, что элемент *from* должен быть типа "#PCDATA"

– !ELEMENT heading определяет, что элемент *heading* должен быть типа "#PCDATA"

– !ELEMENT body определяет, что элемент *body* должен быть типа "#PCDATA"

#PCDATA означает разбираемые текстовые данные.

С DTD XML файл может нести собственный формат и можно быть уверенным, что получаемые из внешних источников данные будут корректными.

XML схема

Также, как и DTD, XML схемы описывают структуру XML документа. XML документ, прошедший проверку по XML схеме, является "синтаксически верным" и "валидным".

Пример XML схемы

```
<?xml version="1.0"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="note">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="to" type="xs:string"/>
        <xs:element name="from" type="xs:string"/>
        <xs:element name="heading" type="xs:string"/>
        <xs:element name="body" type="xs:string"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

Приведенная выше схема интерпретируется следующим образом:

- `<xs:element name="note">` определяет элемент "note";
- `<xs:complexType>` у элемента "note" комплексный тип;
- `<xs:sequence>` комплексный тип – эпоследовательность элементов;
- `<xs:element name="to" type="xs:string">` у элемента "to" строковый тип (текст);
- `<xs:element name="from" type="xs:string">` у элемента "from" строковый тип;
- `<xs:element name="heading" type="xs:string">` у элемента "heading" строковый тип;
- `<xs:element name="body" type="xs:string">` у элемента "body" строковый тип.

Как видно из примера, каждая XML схема состоит с корневого элемента «schema» и обязательного пространства имен «http://www.w3.org/2001/XMLSchema». Далее идет описание схемы и

собственно сама схема. При этом очень часто в очень качественных схемах описание бывает куда большим, чем сама XML Schema.

XML схема мощнее DTD, так как:

- XML схема пишется на XML;
- XML схема легко расширяется;
- XML схема поддерживает типы данных;
- XML схема поддерживает пространства имен;
- XML схема поддерживает типы данных.

Задание к лабораторной работе № 2

На основе прошлой лабораторной работы составьте валидный XML-документ, используя 2 способа: с помощью DTD, а также XML схемы.