ЛАБОРАТОРНАЯ РАБОТА 10

**Создание XML-документа**

*Цель работы:*формирование практических умений создания XML-документов.

**Теоретические сведения**

Расширяемый язык разметки XML (eXtensible Markup Lan-guage) предназначен для описания данных с помощью разметки, синтаксис которой определен соответствующей спецификацией.

Синтаксис разметки XML-документа, представляя собой со­вокупность тегов и их атрибутов, внешне напоминает HTML. Однако есть существенное отличие: XML является *языком метаразметки*, что означает отсутствие фиксированного набора тегов и атрибутов для описания данных, и позволяет создателям документов определять названия тегов и атрибутов в соответст­вии с предметной областью и содержанием хранимых данных.

Разметка в XML-документе описывает структуру содержа­щихся в нем данных, не указывая при этом, как они должны отображаться. XML-документ может создаваться и редактиро­ваться с помощью любого текстового редактора (например, Блокнота) или специализированного XML-редактора (например, XML Notepad).

Рассмотрим структуру XML-документа, содержащего инфор­мацию об имеющихся в наличии книгах и статьях (пример 1.1).

**Пример 1.1**

*<?xml version="1.0" encoding="Windows-1251"?>*

*<!DOCTYPE documents SYSTEM "documents.dtd">*

*<documents>*

*<books>*

*<book id="Book1">*

*<author>Эрих Мария Ремарк</author>*

*<title>Тритоварища</title>*

*<translate> Иванова И.В. </translate>*

*<year\_publication>2000</year\_publication>*

*</book>*

*<book id="Book2">*

*<author>РэйБрэдбери</author>*

*<title>Вино из одуванчиков</title>*

*<translate>ПетроваА.К. </translate>*

*<year\_publication>2009</year\_publication>*

*</book>*

*</books>*

*<articles>*

*<article id="Article1">*

*<author>Виктор Кожевников</author>*

*<title>Классификация электронных средств обучения</title>*

*<year\_publication>1998</year\_publication >*

*</article>*

*</articles>*

*</documents>*

Первые две строчки называются прологом XML-документа. Непосредственно данные с разметкой являются телом документа, которое должно быть заключено в так называемый корневой элемент (в примере *<documents>*).

В *объявлении типа документа* (строка *<?xml version="1.0" encoding="Windows-1251"?>*) определяются *версия* XML и *кодировка*, используемая в документе.

В разделе *DOCTYPE*указывается имя корневого элемента и DTD-тип документа, описанный в отдельном файле (как в при­мере) или непосредственно в XML-документе в квадратных скобках [ ].

Содержимое XML-документа представляет собой набор элементов, секций CDATA, директив анализатора, комментариев, специальных символов, текстовых данных.

Единица данных вместе с относящейся к ней разметкой на­зывается *элементом*. Любой непустой элемент должен состоять из начального, конечного тэгов и данных, заключенных между ними. Вся информация, располагающаяся между начальным и конечными тэгами, рассматривается в XML как данные, и по­этому учитываются все символы форматирования (пробелы, пе­реводы строк, табуляции не игнорируются, как в HTML). Если элемент не имеет содержимого, то есть данных, которые он должен определять, он называется пустым. Для определения пустого элемента используется один тег с косой чертой перед закрывающей угловой скобкой (например, **<***empty***/>**).

Если при определении элементов необходимо задать какие-ли-бо параметры, уточняющие его характеристики, то можно ис­пользовать *атрибуты* элемента. В примере 1.1 для указания идентификатора книги или статьи используется атрибут *id.*

Комментарии обозначаются последовательностью символов *<!--*и*-->*.

Чтобы задать область документа, которую при разбореXML-анализатор будет рассматривать как простой текст, игно­рируя любые инструкции и специальные символы, но, в отличие от комментариев, иметь возможность использовать их в прило­жении, необходимо использовать тэги *<![CDATA[* и *]]>*. Внутри этого блока можно помещать, например, инструкции JavaScript.

Для того чтобы включить в документ символ, используемый для определения каких-либо конструкций языка (например, символ угловой скобки) и не вызвать при этом ошибок в про­цессе разбора, нужно использовать его *специальный символьный либо числовой идентификатор*. Например, *&amp; &lt; &gt*; *&quot*; или*&#6*0;(десятичная форма записи символа<), &#x3c (шестнадцатеричная форма записи) и т. д.

*Инструкции, предназначенные для анализаторов языка*, описываются при помощи тэгов *<?*и *?>*.

В общем случае XML-документы должны удовлетворять следующим требованиям:

- для непустых элементов обязателен закрывающий тег;

- нужно учитывать регистр символов;

- все значения атрибутов должны быть заключены в ка­вычки;

- элементы не должны перекрываться;

- должен быть только один корневой элемент;

- элемент не может иметь несколько атрибутов с одним именем.

**Задание**

Создайте XML-документ согласно варианту. Документ должен содержать не менее трех записей.

**Вариант 1**

Информация о книгах в библиотеке:

- идентификационный номер;

- стеллаж:

ü номер стеллажа;

ü жанр;

ü тематика;

- автор или авторы (*наличие необязательно*);

- название;

- под редакцией (*наличие необязательно*);

- общее количество экземпляров;

- количество выданных экземпляров.

**Вариант 2**

Информация о сотрудниках фирмы:

- фамилия;

- имя;

- отчество;

- адрес:

ü индекс;

ü тип населенного пункта (город, поселок, деревня);

ü область;

ü район;

ü название населенного пункта;

ü улица;

ü дом;

ü корпус (*наличие необязательно*);

ü квартира;

- телефон:

ü домашний;

ü мобильный (*может быть несколько*);

- должность;

- подразделение.

**Вариант 3**

«Кулинарная книга»:

- тип блюда;

- название блюда;

- мера весов;

- ингредиент 1;

- количество ингредиента 1;

…

- рецепт;

- количество калорий.

**Вариант 4**

Информация об услугах фитнес-центра:

- разновидность услуги;

- стоимость каждой услуги;

- инструктор:

ü фамилия,

ü имя,

ü отчество,

ü образование (уровень квалификации);

- дни недели и время занятий.

**Вариант 5**

Информация о подписке на периодические издания:

- тип издания (газета или журнал);

- тематика;

- название издания;

- периодичность выхода издания;

- цена;

- подписной индекс.

**Вариант 6**

Информация о расписании междугородних автобусов:

- направление;

- остановки;

- дата отправления;

- время отправления (*автобусы могут отправляться не­сколько раз в день*);

- стоимость билетов;

- количество свободных мест.

**Вариант 7**

Информация о передачах на ТВ:

- передача;

- день:

ü число;

ü месяц;

ü год;

ü день недели;

- канал;

- жанр;

- время начала;

- продолжительность.

**Вариант 8**

Информация о графике работы врачей медицинского центра:

- специализация врача;

- врач:

ü фамилия,

ü имя,

ü отчество,

- дни приема;

- время приема по дням;

- номер кабинета.

**Вариант 9**

Информация о предложениях туристического агентства:

- страна;

- тип тура (экскурсия, индивидуальный тур, тур выходного дня);

- описание тура;

- горящий тур (*непарный элемент*);

- стоимость;

- количество.

**Контрольные вопросы**

1. Для чего используется язык XML?

2. Чем XML отличается от HTML?

3. Что такое элемент XML?

4. Для чего используется раздел *DOCTYPE*?

5. Как обозначаются комментарии?

6. Перечислите правила построения XML-элементов.