**ФЕДЕРАЛЬНОЕ АГЕНТСТВО СВЯЗИ**

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ**

**«САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ТЕЛЕКОММУНИКАЦИЙ ИМ. ПРОФ. М.А. БОНЧ-БРУЕВИЧА»**

**(СПбГУТ)**

Факультет информационных систем и технологий

Кафедра информационных управляющих систем

**Дисциплина:** Корпоративные сетевые технологии

**Лабораторная работа №3**

Действительные документы XML. DTD

Ответственный исполнитель

Студент группы ИСТ–731

\_\_\_\_\_\_\_\_\_\_\_\_ Коханчик И.М.

«\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_\_ 2021 г.

Санкт-Петербург

2021

**Цель лабораторной работы:**

Целью работы является получение сведений о возможных ограничениях для действительных документов XML, способах определения действительной структуры документа, DTD

**Задачи:**

- Изучить возможные ограничения для действительных документов XML и DTD;

- Правильно составить схему DTD.

**1 Возможные ограничения для действительных документов XML**

Документы XML бывают правильно построенными и действительными. Первая группа соответствует всем общим правилам синтаксиса XML. Однако структура – это несколько большее, чем просто синтаксис. Соответствие документа некоторым внешним правилам делает его действительным. Все эти правила содержания можно объединить в виде схемы, которая определяет структуру экземпляра XML. Разработчик создаёт схему для применения к документу, который предназначен для решения определённой проблемы; кроме того, возможен вариант выбора схемы из заранее созданной общедоступной библиотеки.

Схема чётко определяет имя и структуру корневого элемента, включая спецификацию всех его дочерних элементов. Разработчик может задать, какие элементы и в каком количестве обязательны, а какие – необязательны. Схема определяет, какие элементы содержат атрибуты, а также допустимые значения этих атрибутов, в том числе и значения по умолчанию. Как и в случае элементов, атрибуты могут быть обязательными и необязательными.

**2 DTD**

В программировании на XML используется несколько разновидностей схем. Одной из разновидностей является схема DTD (англ. Document Type Definition определение типа документа).

В XML-документах DTD определяет набор действительных элементов, идентифицирует элементы, которые могут находиться в других элементах, и определяет действительные атрибуты для каждого из них.

В DTD для XML используются следующие типы правил: правила для элементов и их атрибутов, описания категорий (макроопределений), описание форматов бинарных данных. Все они описывают основные конструкции языка – элементы, атрибуты, символьные константы внешние файлы бинарных данных.

Для того, чтобы использовать DTD в нашем документе, мы можем или описать его во внешнем файле и при описании DTD просто указать ссылку на этот файл или же непосредственно внутри самого документа выделить область, в которой определить нужные правила.

В том случае, если используются одновременно внутренние и внешние описания, то программой-анализатором будут сначала рассматриваться внутренние, т.е. их приоритет выше. При проверке документа XML-процессор в первую очередь ищет DTD внутри документа.

Элемент в DTD определяется с помощью дескриптора !ELEMENT, в котором указывается название элемента и структура его содержимого. Списки атрибутов элемента определяются с помощью ключевого слова !ATTLIST. Внутри него задаются названия атрибутов, типы их значений и дополнительные параметры. Компонент (entity) представляет собой определения, содержимое которых может быть повторно использовано в документе, создаются DTD-компоненты при помощи инструкции !ENTITY.

**3 Определение DTD схемы**

Определяем DTD схему для xml-документа, который содержит следующий код:

<?xml version="1.0" encoding="UTF-8"?>

<film>

<DVD жанр="боевик">

<название>Гладиатор</название>

<год знач="2000"/>

<страна>США, Великобритания</страна>

<режиссер>Ридли Скотт</режиссер>

</DVD>

<DVD жанр="фантастика">

<название>Зеленая миля</название>

<год знач="1999"/>

<страна>США</страна>

<режиссер>Фрэнк Дарабонт</режиссер>

</DVD>

<BD жанр="комедия">

<название>Зависнуть в Палм-Спрингс</название>

<год знач="2020"/>

<страна>США</страна>

<режиссер>Макс Барбаков</режиссер>

</BD>

<BD жанр="триллер">

<название>Суд над чикагской семеркой</название>

<год знач="2020"/>

<страна>США</страна>

<режиссер>Аарон Соркин</режиссер>

</BD>

<DVD жанр="фантастика">

<название>Довод</название>

<год знач="2020"/>

<страна>США</страна>

<режиссер>Кристофер Нолан</режиссер>

</DVD>

<DVD жанр="драма">

<название>Маленькие женщины</название>

<год знач="2019"/>

<страна>США</страна>

<режиссер>Грета Гервиг</режиссер>

</DVD>

<DVD жанр="криминал">

<название>Джентльмены</название>

<год знач="2019"/>

<страна>США</страна>

<режиссер>Гай Ричи</режиссер>

</DVD>

<DVD жанр="биография">

<название>Манк</название>

<год знач="2020"/>

<страна>США</страна>

<режиссер>Девид Финчер</режиссер>

</DVD>

<DVD жанр="военный">

<название>1917</название>

<год знач="2019"/>

<страна>США</страна>

<режиссер>Сэм Мендес</режиссер>

</DVD>

<DVD жанр="драма">

<название>Еще по одной</название>

<год знач="2020"/>

<страна>Дания</страна>

<режиссер>Томас Винтерберг</режиссер>

</DVD>

<DVD жанр="ужасы">

<название>Маяк</название>

<год знач="2019"/>

<страна>США</страна>

<режиссер>Роберт Эггерс</режиссер>

</DVD>

</film>

Внешняя DTD схема представлена ниже:

<!ELEMENT film (DVD, BD)>

<!ELEMENT DVD (название, год, страна, режиссер)>

<!ELEMENT BD (название, год, страна, режиссер)>

<!ELEMENT название (#PCDATA)>

<!ELEMENT страна (#PCDATA)>

<!ELEMENT режиссер (#PCDATA)>

<!ATTLIST DVD жанр CDATA #REQUIRED>

<!ATTLIST DVD год CDATA #REQUIRED>

На рисунке один продемонстрирован тест валидности документа с использованием программы XMLPAD.

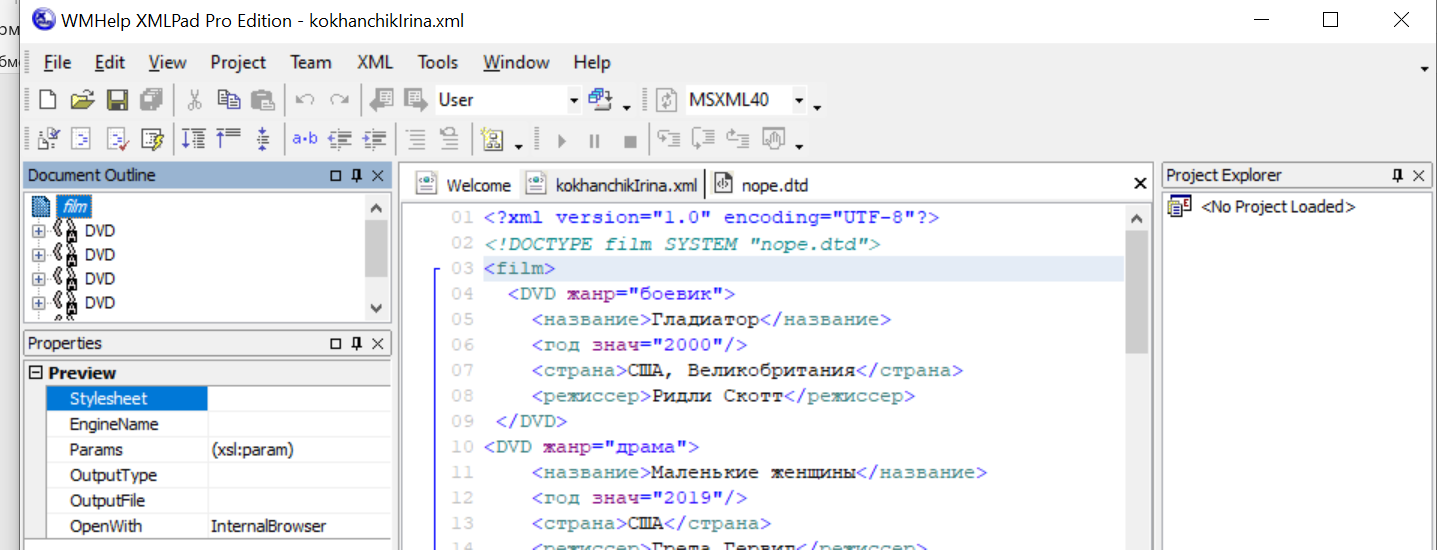


Рисунок 1

Ошибки отсутствуют (рисунок 2).

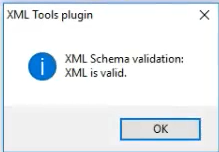


Рисунок 2