

XML Schema

Базы данных и экспертные системы
кафедра № 36 “Информационные системы и технологии”
НИЯУ “МИФИ”

С. К. Муравьев

23 October 2013

XML Schema

XML Schema — язык описания структуры XML-документа. Спецификация XML Schema является рекомендацией W3C.

Файл, содержащий XML Schema, обычно имеет расширение «.xsd» (XML Schema definition).

```
<?xml version="1.0" encoding="utf-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="country">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="country_name" type="xs:string"/>
        <xs:element name="population" type="xs:decimal"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

xs:all

Все дочерние элементы могут появляться в любом порядке 0 или 1 раз.

Атрибут	Описание
id	Необязательный уникальный идентификатор элемента.
maxOccurs	Необязательный. Определяет максимальное число появлений элемента. Должен равняться 1.
minOccurs	Необязательный. Определяет минимальное число появлений элемента. Может равняться 0 или 1. По умолчанию равен 1.

xs:all

Элементы “firstname” и “lastname” обязаны появиться один раз в произвольном порядке.

```
<xs:element name="person">
  <xs:complexType>
    <xs:all>
      <xs:element name="firstname" type="xs:string"/>
      <xs:element name="lastname" type="xs:string"/>
    </xs:all>
  </xs:complexType>
</xs:element>
```

minOccurs="0"

Элементы “firstname” и “lastname” могут появиться в произвольном порядке только один раз.

```
<xs:element name="person">
  <xs:complexType>
    <xs:all minOccurs="0">
      <xs:element name="firstname" type="xs:string"/>
      <xs:element name="lastname" type="xs:string"/>
    </xs:all>
  </xs:complexType>
</xs:element>
```

xs:annotation

Элемент верхнего уровня с описанием схемы

Атрибут	Описание
id	Необязательный уникальный идентификатор элемента.

```
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">  
  
  <xs:annotation>  
    <xs:appinfo>W3Schools Note</xs:appinfo>  
    <xs:documentation xml:lang="en">  
      This Schema defines a W3Schools note!  
    </xs:documentation>  
  </xs:annotation>  
  
  .  
  .  
  .  
  
</xs:schema>
```

xs:any

Разрешает использовать в XML документе произвольные элементы, не определённые в схеме.

Атрибут	Описание
id	Необязательный уникальный идентификатор элемента.
maxOccurs	Необязательный. Определяет максимальное неотрицательное число появлений произвольного элемента. Неограниченному числу появлений соответствует значение “unbounded”. По умолчанию равен 1.
minOccurs	Необязательный. Определяет минимальное неотрицательное число появлений элемента. По умолчанию равен 1.
namespace	Необязательный. Определяет пространства имён, чьи элементы разрешается использовать. Может принимать следующие значения: <ul style="list-style-type: none">• <code>##any</code> – разрешаются элементы любых пространств имён (по умолчанию);• <code>##other</code> – разрешаются элементы любых пространств имён, кроме пространства имён родительского элемента;• <code>##local</code> – разрешаются элементы без пространства имён;• <code>##targetNamespace</code> – разрешаются элементы из пространства имён родительского элемента;• Список {URI пространств имён, <code>##targetNamespace</code>, <code>##local</code>} – разрешаются элементы из указанных пространств имён.

xs:any

Атрибут	Описание
processContents	Необязательный. Определяет правила проверки соответствующих произвольных элементов. Может принимать следующие значения: <ul style="list-style-type: none">• strict – обработчик XML должен получить схемы требуемых пространств имён и проверить элементы (по умолчанию);• lax – аналогично strict. Если схема не может быть получена, то ошибки не произойдёт;• skip – обработчик XML не пытается получить схемы требуемых пространств имён и проверить элементы

В элементе “person” после дочернего элемента “lastname” может появиться любой дочерний элемент.

```
<xs:element name="person">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="firstname" type="xs:string"/>
      <xs:element name="lastname" type="xs:string"/>
      <xs:any minOccurs="0"/>
    </xs:sequence>
  </xs:complexType>
</xs:element>
```

xs:anyAttribute

Разрешает использовать в XML документе произвольные атрибуты, не определённые в схеме.

Атрибут	Описание
id	Необязательный уникальный идентификатор элемента.
namespace	Необязательный. Определяет пространства имён, чьи атрибуты разрешается использовать. Может принимать следующие значения: <ul style="list-style-type: none">• <code>##any</code> – разрешаются атрибуты любых пространств имён (по умолчанию);• <code>##other</code> – разрешаются атрибуты любых пространств имён, кроме пространства имён родительского элемента;• <code>##local</code> – разрешаются атрибуты без пространства имён;• <code>##targetNamespace</code> – разрешаются атрибуты из пространства имён родительского элемента;• Список {URI пространств имён, <code>##targetNamespace</code>, <code>##local</code>} – разрешаются атрибуты из указанных пространств имён.

xs:anyAttribute

Атрибут	Описание
processContents	Необязательный. Определяет правила проверки соответствующих произвольных элементов. Может принимать следующие значения: <ul style="list-style-type: none">• strict – обработчик XML должен получить схемы требуемых пространств имён и проверить элементы атрибуты (по умолчанию);• lax – аналогично strict. Если схема не может быть получена, то ошибки не произойдёт;• skip – обработчик XML не пытается получить схемы требуемых пространств имён и проверить элементы

В элементе “person” могут содержаться произвольные атрибуты.

```
<xs:element name="person">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="firstname" type="xs:string"/>
      <xs:element name="lastname" type="xs:string"/>
      <xs:anyAttribute/>
    </xs:sequence>
  </xs:complexType>
</xs:element>
```

xs:anyAttribute

Атрибут	Описание
processContents	Необязательный. Определяет правила проверки соответствующих произвольных элементов. Может принимать следующие значения: <ul style="list-style-type: none">• strict – обработчик XML должен получить схемы требуемых пространств имён и проверить элементы атрибуты (по умолчанию);• lax – аналогично strict. Если схема не может быть получена, то ошибки не произойдёт;• skip – обработчик XML не пытается получить схемы требуемых пространств имён и проверить элементы

В элементе “person” могут содержаться произвольные атрибуты.

```
<xs:element name="person">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="firstname" type="xs:string"/>
      <xs:element name="lastname" type="xs:string"/>
      <xs:anyAttribute/>
    </xs:sequence>
  </xs:complexType>
</xs:element>
```

xs:appinfo

Определяет информацию для использования приложениями.

Атрибут	Описание
source	Необязательный. Определяет источник получения информации.

```
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">

  <xs:annotation>
    <xs:appinfo>W3Schools Note</xs:appinfo>
    <xs:documentation xml:lang="en">
      This Schema defines a W3Schools note!
    </xs:documentation>
  </xs:annotation>

  .
  .
  .

</xs:schema>
```

xs:attribute

Определение атрибута

Атрибут	Описание
default	Необязательный. Определяет значение атрибута по умолчанию. Не может использоваться вместе с атрибутом fixed.
fixed	Необязательный. Определяет фиксированное значение атрибута. Не может использоваться вместе с атрибутом default.
form	Необязательный. Определяет форму атрибута. Значение по умолчанию берётся из атрибута attributeFormDefault родительского элемента. Может принимать следующие значения: <ul style="list-style-type: none">• qualified – имя атрибута должно начинаться с префикса пространства имён и не должно содержать двоеточий (NCName);• unqualified – на имя атрибута не налагается ограничений.
id	Необязательный уникальный идентификатор элемента.
name	Необязательный. Определяет имя атрибута. Не может использоваться вместе с атрибутом ref.
ref	Необязательный. Определяет ссылку на именованный атрибут. Не может использоваться вместе с атрибутами name, form, type и элементом simpleType

xs:attribute

Атрибут	Описание
type	Необязательный. Определяет тип данных. Не может использоваться вместе элементом <code>simpleType</code> .
use	Необязательный. Определяет порядок использования атрибута. Может принимать следующие значения: <ul style="list-style-type: none">• <code>optional</code> – атрибут не является обязательным (значение по умолчанию);• <code>prohibited</code> – использование атрибута запрещено;• <code>required</code> – атрибут является обязательным.

Атрибут “code” может принимать значения из двух прописных букв от A до Z.

```
<xs:attribute name="code">  
  
  <xs:simpleType>  
    <xs:restriction base="xs:string">  
      <xs:pattern value="[A-Z][A-Z]"/>  
    </xs:restriction>  
  </xs:simpleType>  
  
</xs:attribute>
```

xs:attributeGroup

Определение группы атрибутов, которую можно использовать в определении сложных типов

Атрибут	Описание
	attributeGroup

xs:choice

Определение перечня элементов, один из которых может появиться в соответствующей позиции XML документа

Атрибут	Описание

xs:complexContent

Определяет расширения или ограничения для сложного типа, содержащего только элементы или смешанное содержимое.

Атрибут	Описание

xs:complexType

Определение сложного типа данных. Элемент XML документа, имеющий сложный тип, содержит другие элементы и/или атрибуты.

Атрибут	Описание

xs:documentation

Текстовое описание схемы

Атрибут	Описание
source	Необязательный. Определяет источник получения описания.
xml:lang	Необязательный. Определяет язык, на котором сделано описание.

```
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">

  <xs:annotation>
    <xs:appinfo>W3Schools Note</xs:appinfo>
    <xs:documentation xml:lang="en">
      This Schema defines a W3Schools note!
    </xs:documentation>
  </xs:annotation>
  .
  .
  .
</xs:schema>
```

xs:element

Определение элемент.

Атрибут	Описание

xs:extension

Расширяет существующее определение простого или сложного типа

Атрибут	Описание

xs:field

Задаёт выражение XPath, результат которого используется для определения ограничения идентичности.

Атрибут	Описание

xs:group

Определение группы элементов для использования в объявлении сложных типов данных.

Атрибут	Описание

xs:import

Позволяет включать в документ несколько схем с разными пространствами имён.

Атрибут	Описание

xs:include

Позволяет включать в документ несколько схем в одном пространстве имён.

Атрибут	Описание

xs:key

Позволяет указать, что атрибут или элемент является ключом.

Атрибут	Описание

xs:keyref

Позволяет указать, что атрибут или элемент содержит ключ другого элемента.

Атрибут	Описание

xs:list

Определяет список допустимых значений простого типа данных.

Атрибут	Описание

xs:notation

Определяет формат не-XML данных внутри XML документа.

Атрибут	Описание

xs:redefine

Переопределяет простые и сложные типы данных или группы во внешней схеме.

Атрибут	Описание

xs:restriction

Задаёт ограничения для элементов `simpleType`, `simpleContent`, или `complexContent`

Атрибут	Описание

xs:Schema

Определяет корневой элемент схемы

Атрибут	Описание

xs:selector

Задаёт выражение XPath, которое задаёт множество элементов для ограничения идентичности.

Атрибут	Описание

xs:sequence

Задаёт перечень допустимых дочерних элементов.

Атрибут	Описание

xs:simpleContent

Задаёт расширения или ограничения на текстовое содержимое сложного типа данных или на простой тип данных.

Атрибут	Описание

xs:simpleType

Определяет простой тип данных.

Атрибут	Описание

xs:union

Определяет простой тип данных как набор значений.

Атрибут	Описание

xs:unique

Задаёт требование уникальности элемента или атрибута.

Атрибут	Описание

Ограничения и уточнения XSD

- ▶ Ограничения значений
- ▶ Ограничения множеств значений
- ▶ Ограничения списков значений
- ▶ Ограничения на пробельные символы
- ▶ Ограничения длины
- ▶ Ограничения типов данных

Практика

- ▶ Разработать схему XML-документа со структурой университетов
- ▶ Выложить результат на Github
- ▶ Проверить соответствие XML-документа вашей схеме