UT 3: XML Schema

ÍNDICE

1. Previo: Repaso de espacios de nombre

- 2. Introducción
- 3. Elementos simples y compuestos
- 4. Atributos
- 5. Restricciones
- 6. Extensiones
- 7. Indicadores

Actividad 1: Busca información sobre los espacios de nombre de XML.

Procura contestar <u>al menos</u> estas preguntas:

- ¿Por qué se usan?
- ¿Cómo se usan?
- 2 ejemplos de documentos con espacios de nombre

- Los nombres de XML son definidos por el usuario → Conflictos al mezclar archivos de diferentes aplicaciones XML
 - Ejemplo: Etiqueta de tabla de datos y de mesa (en inglés)

```
\lambda Apples \lambda /td \rangle Apples \lambda /td \rangle \lambda \l
```

• **Espacios de nombre** = Herramienta de XML para evitar los conflictos de nombres

- Para evitar estos conflictos, se usan prefijos:
 - **h**: HTML
 - **f**: furniture (muebles en inglés)

```
<h:table>
  <h:tr>
    <h:td>Apples</h:td>
    <h:td>Bananas</h:td>
 </h:tr>
</h:table>
<f:table>
 <f:name>African Coffee Table</f:name>
 <f:width>80</f:width>
  <f:length>120</f:length>
</f:table>
```

- Para usar prefijos debemos definir un espacio de nombre para cada prefijo
- El espacio de nombre se puede definir con el atributo xmlns (XML NameSpace) en la etiqueta de inicio de un elemento
- El espacio de nombre y el prefijo es válido para todos los elementos hijos
 - Se suelen definir en el elemento raíz

- La URL que se usa para definir el espacio de nombre se utiliza para darle un nombre único y no se usa → No hace falta que exista pero no se puede repetir
- **Espacios de nombre por defecto**: Definimos uno por defecto y nos ahorramos usa el prefijo para cada etiqueta hijo.
 - La notación es similar pero no se especifica prefijo

ÍNDICE

1. Previo: Repaso de espacios de nombre

2. Introducción

- 3. Elementos simples y compuestos
- 4. Atributos
- 5. Restricciones
- 6. Extensiones
- 7. Indicadores

• XML Schema = XSD (XML Schema Definition). Define la estructura del documento XML

¿Por qué usar XSD?

- Gran cantidad de formatos estandarizados. Muchos de ellos se definen con Schemas
- Alternativa basada en XML y más potente que los DTD

XSD tiene tipos de datos

- Más fácil describir el posible contenido del documento → restricciones, patrones...
- Más fácil comprobar si son válidos

Elemento <schema>

- Es la raíz de todos los archivos XSD
- Contiene atributos, normalmente estos:

```
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
targetNamespace="https://www.w3schools.com"
xmlns="https://www.w3schools.com"
elementFormDefault="qualified">
...
</xs:schema>
```

Especifica de donde vienen los elementos y tipos de datos usados y que usan el prefijo **xs**

Se usará siempre

Elemento <schema>

- Es la raíz de todos los archivos XSD
- Contiene atributos, normalmente estos:

```
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
targetNamespace="https://www.w3schools.com"
xmlns="https://www.w3schools.com"
elementFormDefault="qualified">
...
</xs:schema>
```

Especifica de donde tienen que venir los elementos que vamos a definir en el archivo

Su valor dependerá del archivo

Elemento <schema>

- Es la raíz de todos los archivos XSD
- Contiene atributos, normalmente estos:

```
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
targetNamespace="https://www.w3schools.com"
xmlns="https://www.w3schools.com"
elementFormDefault="qualified">
...
</xs:schema>
```

Indica el espacio de nombre por defecto

Enlazando un XML con un Schema

```
<?xml version="1.0"?>
<note xmlns="https://www.w3schools.com"</pre>
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="https://www.w3schools.com note.xsd">
<to>Tove</to>
<from>Jani</from>
<heading>Reminder</heading>
<body>Don't forget me this weekend!</body>
</note>
```

Espacio de nombre por defecto. Si no ponemos el prefijo, pertenecerán a este

Enlazando un XML con un Schema

```
<?xml version="1.0"?>
<note xmlns="https://www.w3schools.com"</pre>
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="https://www.w3schools.com note.xsd">
<to>Tove</to>
<from>Jani</from>
<heading>Reminder</heading>
<body>Don't forget me this weekend!</body>
</note>
```

Especifica de donde vienen los elementos y tipos de datos usados y que usan el prefijo **xsi**

Se usará siempre

Enlazando un XML con un Schema

```
<?xml version="1.0"?>
<note xmlns="https://www.w3schools.com"</pre>
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="https://www.w3schools.com note.xsd">
<to>Tove</to>
<from>Jani</from>
<heading>Reminder</heading>
<body>Don't forget me this weekend!</body>
</note>
```

Especifica el espacio de nombres que vamos a validar + archivo usado

Crea un archivo XML y un Schema en NetBeans

• ¿Qué pasa si no usas "targetNamespace" en el .xsd?

- Dependiendo de dónde coloquemos la definición de los elementos del esquema, varía su ámbito de aplicación:
 - <u>Ámbito global</u>: se colocan dentro de la etiqueta raíz schema y no están dentro de ninguna otra. Estos elementos se pueden utilizar en cualquier parte del esquema.
 - <u>Ámbito local</u>: están definidos dentro de otros elementos. Se pueden utilizar sólo dentro del elemento en el que están definidos y no en todo el documento

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"</pre>
           xmlns:doc="http://www.educacion.navarra.es"
           targetNamespace="http://www.educacion.navarra.es">
 <xs:element...> <!--Definición global -->
    <xs:simpleType...> <!--Definición local -->
    </xs:simpleType>
  </xs:element >
  <xs:simpleType...> <!--Definición global -->
  </xs:simpleType>
</xs:schema>
```

ÍNDICE

- 1. Previo: Repaso de espacios de nombre
- 2. Introducción
- 3. Elementos simples y compuestos
- 4. Atributos
- 5. Restricciones
- 6. Extensiones
- 7. Indicadores

Elementos simples

- Son elementos que solo pueden contener texto. No pueden contener otros elementos ni atributos
- Solo texto = datos en alguno de los tipos de XSD con restricciones o patrones
- Sintaxis <xs:element name="xxx" type="yyy"/>
- Tipos básicos: xs:string, xs:decimal, xs:integer, xs:boolean, xs:date, xs:time
- Valores por defecto y fixed: atributos default y fixed

```
Elementos simples <xs:element name="xxx" type="yyy"/>
```

```
<lastname>Refsnes<age>36</age><dateborn>1970-03-27
```



```
<xs:element name="lastname" type="xs:string"/>
<xs:element name="age" type="xs:integer"/>
<xs:element name="dateborn" type="xs:date"/>
```

Elementos compuestos

- Es un elemento que contiene otros elementos y/o atributos
- Hay 4 tipos:
 - 1. Elementos que contienen otros elementos
 - 2. Elementos que contienen solo texto
 - 3. Elementos que contienen elementos y texto (mixtos)
 - 4. Elementos vacíos
- Cada uno de estos tipos puede tener atributos

Elementos compuestos: <u>elementos que contienen otros elementos</u>

• **Opción 1**: definición local

```
<employee>
    <firstname>John</firstname>
    <lastname>Smith</lastname>
</employee>
_
```

Los elementos complejos se definen como un tipo complejo

Elementos compuestos: <u>elementos que contienen otros elementos</u>

• **Opción 1**: definición local

Con xs:sequence marcamos los elementos que componen a otro

Elementos compuestos: <u>elementos que contienen otros elementos</u>

• **Opción 1**: definición local

```
<employee>
    <firstname>John</firstname>
    <lastname>Smith</lastname>
</employee>
_
```

Definición de los elementos internos de <employee>

Elementos compuestos: <u>elementos que contienen otros elementos</u>

Opción 2: definición global

Elementos compuestos: <u>elementos que contienen otros elementos</u>

- **Opción 2**: definición global:
 - Reutilizar tipo

Definición (**global**) del tipo de datos
Le hemos dado el nombre personinfo

```
<xs:element name="employee" type="personinfo"/>
<xs:element name="student" type="personinfo"/>
<xs:element name="member" type="personinfo"/>
```

Elementos compuestos: <u>elementos que contienen otros elementos</u>

- **Opción 2**: definición global:
 - Reutilizar tipo

Elementos del tipo definido

```
<xs:element name="employee" type="personinfo"/>
<xs:element name="student" type="personinfo"/>
<xs:element name="member" type="personinfo"/>
```

Elementos compuestos

- Opción 2: definición global:
 - Reutilizar tipo
 - Usar como base

Nuevo tipo fullpersoninfo

Elementos compuestos

- **Opción 2**: definición global:
 - Reutilizar tipo
 - Usar como base

Toma como base personinfo...

```
<xs:element name="employee" type="fullpersoninfo"/>
<xs:complexType name="personinfo">
  <xs:sequence>
    <xs:element name="firstname" type="xs:string"/>
    <xs:element name="lastname" type="xs:string"/>
  </xs:sequence>
</xs:complexType>
<xs:complexType name="fullpersoninfo">
  <xs:complexContent>
    <xs:extension base="personinfo">
      <xs:sequence>
        <xs:element name="address" type="xs:string"/>
        <xs:element name="city" type="xs:string"/>
        <xs:element name="country" type="xs:string"/>
      </xs:sequence>
    </xs:extension>
  </xs:complexContent>
</xs:complexType>
```

Elementos compuestos

- **Opción 2**: definición global:
 - Reutilizar tipo
 - Usar como base

...y AÑADE estos elementos

```
<xs:element name="employee" type="fullpersoninfo"/>
<xs:complexType name="personinfo">
  <xs:sequence>
    <xs:element name="firstname" type="xs:string"/>
    <xs:element name="lastname" type="xs:string"/>
  </xs:sequence>
</xs:complexType>
<xs:complexType name="fullpersoninfo">
  <xs:complexContent>
    <xs:extension base="personinfo">
      <xs:sequence>
        <xs:element name="address" type="xs:string"/>
        <xs:element name="city" type="xs:string"/>
        <xs:element name="country" type="xs:string"/>
      </xs:sequence>
    </xs:extension>
  </xs:complexContent>
</xs:complexType>
```

Elementos compuestos: <u>texto (+ atributos)</u>

- Se declaran con un tipo de datos complejo (<xs:complexType>) con contenido simple
 (<xs:simpleContent>) ya que es texto y atributos
- Se toma un tipo de datos y se <u>extiende</u> (con < xs:extension>) se le añaden los atributos,
 convirtiendo al tipo en complejo
- <u>Ejemplo:</u> El elemento <textarea> de HTML acepta texto y atributos (name, cols, rows,...)

Elementos compuestos: <u>texto (+ atributos)</u>

• <u>Ejemplo:</u> El elemento <textarea > de HTML acepta texto y atributos (name, cols, rows,...)

```
<textarea name="comentarios" cols="10" rows="5">
Introduzca aquí sus comentarios
</textarea>
```

<xs:simpleContent> especifica que el elemento complejo no tiene elementos descendientes

Elementos compuestos: <u>texto (+ atributos)</u>

<zapatossize country="france">35</zapatossize>

```
<xs:element name="zapatossize">
    <xs:complexType>
    <xs:simpleContent>
        <xs:extension base="xs:integer">
             <xs:attribute name="country" type="xs:string" />
             </xs:extension>
        </xs:simpleContent>
        </xs:complexType>
</xs:element>
```

```
<xs:element name="zapatosize" type="zapatostype"/>
<xs:complexType name="zapatostype">
  <xs:complexType name="zapatostype">
  <xs:simpleContent>
  <xs:extension base="xs:integer">
   <xs:attribute name="country" type="xs:string" />
  </xs:extension>
  </xs:simpleContent>
  </xs:complexType>
```

Definición LOCAL

Definición GLOBAL

Elementos compuestos: contenido mixto

- Contienen texto y otros elementos (y puede que atributos)
- Se usa el atributo mixed="true" en <xs:complexType>
- El valor por defecto de mixed es false, por lo que no hace falta ponerlo en el resto de elementos

Elementos compuestos: contenido mixto

```
<letter>
  Dear Mr.<name>John Smith</name>.
  Your order <orderid>1032</orderid>
  will be shipped on <shipdate>2001-07-13</shipdate>.
</letter>
```

Definición LOCAL

Definición GLOBAL

ELEMENTOS SIMPLES Y COMPUESTOS

Elementos compuestos: elementos vacíos

- No puede tener contenido, solo atributos
- Definimos un tipo que acepte elementos dentro, pero no definimos estos elementos

ELEMENTOS SIMPLES Y COMPUESTOS

Elementos compuestos: elementos vacíos

cproduct prodid="1345" />

```
<xs:element name="product">
  <xs:complexType>
    <xs:attribute name="prodid" type="xs:integer"/>
    </xs:complexType>
</xs:element>
```

Definición LOCAL

```
<xs:element name="product" type="prodtype"/>
<xs:complexType name="prodtype">
  <xs:attribute name="prodid" type="xs:integer"/>
  </xs:complexType>
```

Definición GLOBAL

ÍNDICE

- 1. Previo: Repaso de espacios de nombre
- 2. Introducción
- 3. Elementos simples y compuestos

4. Atributos

- 5. Restricciones
- 6. Extensiones
- 7. Indicadores

ATRIBUTOS

Atributos

- Solo aparecen en los elementos compuestos
- Se definen con un tipo simple <xs:attribute name="xxx" type="yyy"/>
- Por defecto, son opcionales → Atributo use para obligatorio: use="required"
 - Para especificar que es opcional (aunque es el valor por defecto): use="optional"
- También admiten los atributos fixed y default para valores fijos y por defecto

ÍNDICE

- 1. Previo: Repaso de espacios de nombre
- 2. Introducción
- 3. Elementos simples y compuestos
- 4. Atributos
- 5. Restricciones
- 6. Extensiones
- 7. Indicadores

Restricciones o facetas (facets)

- Sirven para determinar los valores aceptados para los elementos y atributos
- Se escriben dentro del tipo (<xs:SimpleType> o <xs:ComplexType>)
- Algunos ejemplos:
 - Longitud de un atributo (máxima, mínima o exacta)
 - Enumeración de posibles valores
 - Valores que encajen en un determinado patrón (expresiones regulares)
 - •

Faceta	Descripción
xs:maxInclusive	El valor debe ser menor o igual que el indicado.
xs:maxExclusive	El valor debe ser menor que el indicado.
xs:minExclusive	El valor debe ser mayor que el indicado.
xs:minInclusive	El valor debe ser mayor o igual que el indicado.
xs:enumeration	Lista de valores admitidos.
xs:pattern	Patrón de caracteres admitidos.
xs:length	Longitud fija.
xs:minLength	Longitud mínima.
xs:maxLength	Longitud máxima.
xs:totalDigits	El número máximo de dígitos que puede tener un número.
xs:fractionDigits	El número máximo de decimales que puede tener un número.
xs:whiteSpace	Cómo se debe tratar a los posibles espacios en blanco, las tabulaciones, los saltos de línea y los retornos de carro que puedan aparecer.

Restricciones o facetas (facets) – min/max Inclusive/Exclusive

- 4 combinaciones: minInclusive, minExclusive, maxInclusive, maxExclusive
 - Inclusive: incluye el igual // Exclusive: no lo incluye

Definición de un elemento **mes**Su valor es un entero entre 1 y 12
(ambos incluidos)

Restricciones o facetas (facets) – enumeration

Sirve para expresar los únicos valores posibles del elemento

Definición de un elemento **color** Su contenido es de tipo string y solo puede vale verde, amarillo o rojo

Restricciones o facetas (facets) – pattern

- Sirve para limitar el contenido de un elemento a aquellos que tengan una determinada forma
- Ejemplos:
 - Usa solo letras minúsculas
 - Usa solo 3 letras, mayúsculas o minúsculas, seguidas de un número del 0 al 3
 - •
- Expresiones regulares

Restricciones o facetas (facets) – pattern

- Los corchetes [] definen rangos o conjuntos → [A-Z] [a-z] [0-9]
- Símbolos *, +, | se usan igual que en DTD
- $[0-9]{10} \rightarrow el$ elemento [0-9] se repite 10 veces \rightarrow 10 cifras del 0 al 9

Definición de un elemento **letra** Su contenido es una única letra minúscula

Restricciones o facetas (facets) – length, minLength, maxLength

• Sirven para restringir los valores de un string a aquellos de una determinada longitud (o fijar un mínimo o máximo a estos)

Definición de un elemento **clave** Su contenido es de tipo string y tiene una longitud de 12 caracteres

Restricciones o facetas (facets) – totalDigits, fractionDigits

- Parecidos a los length
- Se usan en números
- totalDigits: número <u>máximo</u> de cifras <u>totales</u>
- **fractionDigits**: número <u>máximo</u> de cifras decimales

Restricciones o facetas (facets) – whiteSpace

- Sirven para definir qué hacer con los espacios en blanco, saltos de línea y tabulaciones que haya dentro de un valor.
- Posibles valores:
 - preserve: mantiene los espacios en blanco (valor por defecto)
 - **replace**: sustituye todos los saltos de línea y tabulaciones por espacios en blanco
 - collapse: elimina los espacios al principio y final, elimina si hay varios (como en HTML)

ÍNDICE

- 1. Previo: Repaso de espacios de nombre
- 2. Introducción
- 3. Elementos simples y compuestos
- 4. Atributos
- 5. Restricciones
- 6. Extensiones
- 7. Indicadores

EXTENSIONES

- Sirven para agregar características a un elemento simple o compuesto
- Ya las hemos visto 2 veces:
 - 1. Para añadir atributos a un elemento con solo texto → Dentro de un simpleContent

EXTENSIONES

- Sirven para agregar características a un elemento simple o compuesto
- Ya las hemos visto 2 veces. Nos basta con conocer esos dos usos:
 - 1. Para añadir atributos a un elemento con solo texto → Dentro de un simpleContent
 - **2. Para añadir elementos a un tipo declarado como global** → Dentro de un complexContent

ÍNDICE

- 1. Previo: Repaso de espacios de nombre
- 2. Introducción
- 3. Elementos simples y compuestos
- 4. Atributos
- 5. Restricciones
- 6. Extensiones
- 7. Indicadores

Indicadores

- Establecen cómo se van a utilizar los elementos en un documento XML
- Corresponden a lo que en DTD era *, +, | , etc
- Son 7, que se dividen en 3 tipos:
 - <u>De orden</u>: todo (**all**), secuencia (**sequence**) y elección (**choice**)
 - <u>De ocurrencia</u>: **maxOccurs** y **minOccurs**
 - <u>De grupo</u>: de elementos (**group**) y de atributos (**attributeGroup**)

Indicadores de orden

- Secuencia (**sequence**) → Indica los elementos que van a aparecer y su orden
- All → Indica los elementos que van a aparecer pero no el orden (pueden hacerlo en cualquiera)
 - Intentaremos evitar usarlo
- Elección (**choice**) → Indica que solo se puede escribir uno de los elementos hijos

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

Indicadores de ocurrencia

- maxOccurs

 Indica el número máximo de ocurrencias de un elemento
 - Usaremos maxOccurs = "unbounded" para indicar que se puede usar un número ilimitado de veces
- minOccurs

 Indica el número mínimo de ocurrencias de un elemento

</xs:element>

Indicadores de grupo

- Son dos: <xs:group> y <xs:attributeGroup>
- Ambos sirven para agrupar un conjunto de declaraciones relacionadas, pudiendo referirnos a ellas con el nombre asignado
 - <xs:group> agrupa elementos
 - <xs:attributeGroup> agrupa atributos
- Facilitan la escritura de archivos XSD complejos

Indicadores de grupo

<xs:group>

```
<xs:complexType name="datosDePersona">
  <xs:sequence>
   <xs:group ref="datosBasicos"/>
    <xs:element name="telefono" type="xs:string"/>
  </xs:sequence>
</xs:complexType>
<xs:group name="datosBasicos">
  <xs:sequence>
    <xs:element name="nombre" type="xs:string"/>
    <xs:element name="edad" type="xs:positiveInteger"/>
    <xs:element name="pais" type="xs:string"/>
  </xs:sequence>
</xs:group>
```

Indicadores de grupo

< xs:attributeGroup >

```
<xs:element name="personas">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="persona" maxOccurs="unbounded">
        <xs:complexType>
          <xs:attributeGroup ref="datosDePersona"/>
        </xs:complexType>
      </xs:element>
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:attributeGroup name="datosDePersona">
  <xs:attribute name="nombre" type="xs:string"/>
  <xs:attribute name="edad" type="xs:positiveInteger"/>
  <xs:attribute name="pais" type="xs:string"/>
</xs:attributeGroup>
```