

UT5. XML

UT5. XML

5.8 Validación de documentos XML con esquemas XML.

- 5.8.1. Estructura de un esquema XML. Componentes

- 5.8.2. Componentes básicos de un esquema

- 5.8.3. Tipos de datos

- 5.8.4. Tipos de datos simples vs. complejos

- 5.8.5. Definición de tipos de datos complejos

- 5.8.6. Diferentes declaraciones de elementos

- 5.8.7. Modelos de diseño de esquemas XML

- 5.8.8. Poniendo todo junto

5.8 Validación de documentos XML con esquemas XML

Un esquema XML es un mecanismo para comprobar la validez de un documento XML. Se trata de una forma alternativa a los DTD, pero con ciertas ventajas sobre éstos:

- Es un documento XML, por lo que se puede comprobar si está bien formado.
- Existe un extenso catálogo de tipos de datos predefinidos para elementos y atributos que pueden ser ampliados o restringidos para crear nuevos tipos.
- Permiten concretar con precisión la cardinalidad de un elemento, es decir, las veces que puede aparecer en un documento XML.
- Permite mezclar distintos vocabularios (juegos de etiquetas) gracias a los espacios de nombres.

Como característica negativa, decir que los esquemas XML son más difíciles de interpretar por el ojo humano que los DTD.

5.8 Validación de documentos XML con esquemas XML

En un esquema XML se describe lo que un documento XML puede contener: qué elementos, con qué atributos, de qué tipos de datos son tanto elementos como atributos, en qué orden aparecen los elementos, cuántas ocurrencias puede haber de cada elemento, etc.

Los documentos XML que se validan contra un esquema se llaman **instancias del esquema**.

De hecho, se puede ver el esquema como un molde, y los documentos XML como objetos que tratan de ajustarse a ese molde. También se puede hacer una analogía con la programación orientada a objeto, los esquemas serían como las clases y los documentos XML como los objetos.

5.8.1. Estructura de un esquema XML. Componentes

Un esquema XML es un documento XML que ha de cumplir una serie de reglas:

- Su elemento raíz se llama <schema>.
- Su espacio de nombres debe ser `http://www.w3.org/2001/XMLSchema`. Se podría no definir un prefijo o usar uno de los más comunes: `xs` o `xsd`. Así, el esquema XML más sencillo será:

```
<?xml version="1.0"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
  ...
</xs:schema>
```

5.8.1. Estructura de un esquema XML. Componentes

- Puesto que un documento XML bien formado contiene al menos un elemento raíz, el esquema XML dispondrá de, al menos, un componente de declaración de elemento que defina el elemento raíz del documento XML. Este componente de declaración de elemento se identifica con el elemento **<xs:element>**, que tendrá un atributo name cuyo valor será el nombre del elemento raíz del documento XML.

Ejemplo:

Un documento XML muy simple:

```
<?xml version="1.0"?>
<simple>Es difícil escribir un documento más simple</simple>
```

Un posible esquema que 10 valida:

```
<?xml version="1.0"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="simple" />
</xs:schema>
```

5.8.1. Estructura de un esquema XML. Componentes

- Puede haber un esquema con múltiples elementos raíz.

Ejemplo:

Un esquema con dos elementos raíz declarados, <simple> y <complejo>. Cualquier documento XML cuyo elemento raíz sea bien simple, bien complejo, se considerará conforme con este esquema.

```
<?xml version="1.0"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="simple" />
  <xs:element name="complejo" />
</xs:schema>
```

- Un esquema es un conjunto de declaraciones de elementos y atributos y definición de tipos que sirve para fijar la estructura que deben tener sus documentos instancia XML. **El orden en que se declaran los elementos, llamados componentes, en un esquema no es significativo ni afecta al funcionamiento del mismo.**

5.8.1. Estructura de un esquema XML. Componentes

- La vinculación de un esquema al documento XML se realiza en el documento XML, insertando **xmlns:xsi** y **xsi:noNamespaceSchemaLocation** como atributos del elemento raíz, en el caso de esquemas no asociados a espacios de nombres. Si el esquema estuviera asociado a un espacio de nombres se usaría el atributo **xsi:schemaLocation**.

Ejemplo:

En el documento XML se declara un espacio de nombres de prefijo xsi, propio de las instancias de esquemas XML, y se indica la ubicación del documento del esquema, en este caso de nombre simple.xsd.

```
<?xml version="1.0"?>
<simple xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xsi:noNamespaceSchemaLocation="simple.xsd">
    ...
</simple>
```


5.8.1. Estructura de un esquema XML. Componentes

Atributos de los documentos instancia del esquema o documento XML.

En la terminología de los esquemas, los documentos XML son instancias de los mismos, al fin y al cabo, son objetos salidos del molde que es el esquema.

Por eso, en los documentos XML aparece un espacio de nombres que tiene sentido cuando se trabaja con esquemas asociados. Se ha visto en el elemento raíz del documento XML del ejemplo anterior, la declaración del espacio de nombres **`http://www.w3.org/2001/XMLSchemainstance`** asociada al prefijo **`xsi`**.

A continuación, y en el mismo elemento raíz, se declara un atributo de ese espacio de nombres, **`xsi:noNamespaceSchemaLocation="simple.xsd"`**, que permite vincular el documento XML con un esquema de nombre `simple.xsd`.

5.8.1. Estructura de un esquema XML. Componentes

Existen otros atributos de este espacio de nombres que se usan ocasionalmente en documentos XML (o instancias de esquema). Son:

- **xsi:nil:** asociado a un elemento indica que debe ser vacío de contenido. Puede valer false o true. Por defecto vale false, y para poderlo poner a true, se debe activar el atributo nillable en la definición del elemento en el esquema asociado.

Ejemplo:

En el esquema se declararía un elemento de la forma:

```
<xs:element name="fechaCompra" type="xs:date" nillable="true"/>
```

En el documento instancia XML se explicitaría que el contenido de <fechaCompra> debe ser vacío:

```
<fechaCompra xsi:nil="true"></fechaCompra>
```

5.8.1. Estructura de un esquema XML. Componentes

Existen otros atributos de este espacio de nombres que se usan ocasionalmente en documentos XML (o instancias de esquema). Son:

- **xsi:type:** puede ser la definición de un tipo de un elemento, como `xs:string` o existente en un esquema asociado.
- **xsi:schemaLocation:** localización de un esquema con un espacio de nombres asociado. Se pueden indicar varios esquemas separados por espacios.
- **xsi:noNamespaceSchemaLocation:** localización de un esquema sin un espacio de nombres asociado. Se pueden indicar varios esquemas separados por espacios.

5.8.2. Componentes básicos de un esquema

Se van a introducir los componentes que van a permitir construir un esquema XML que valide documentos XML.

Como se ha comentado, a los documentos XML se les llama instancias del esquema.

Los componentes imprescindibles son:

- `xs:schema`
- `xs:element`
- `xs:attribute`

5.8.2. Componentes básicos de un esquema

xs:schema

Es el componente de declaración de esquema y elemento raíz de todo esquema XML.

Atributos optativos principales:

- **xmlns:** una referencia URI que indica uno o más espacios de nombres a usar en el esquema. Si no se indica ningún prefijo, los componentes del esquema del espacio de nombres pueden usarse de manera no cualificada.

Otros atributos optativos:

- **id:** identificador único para el elemento.
- **targetNamespace:** una referencia URI al espacio de nombres del esquema.
- **elementFormDefault:** formato de los nombres de los elementos en el espacio de nombres del esquema. Puede ser `unqualified`, que indica que los elementos no llevan prefijo, o `qualified`, indica que los elementos deben llevar el prefijo. Posibles valores: `unqualified`, `qualified`
- **attributeFormDefault:** formato de los nombres de los atributos en el espacio de nombres del esquema. Funciona igual `elementFormDefault`. Posibles valores: `unqualified`, `qualified`
- **version:** la versión del esquema.

5.8.2. Componentes básicos de un esquema

xs:schema

Ejemplo:

Estructura de cualquier esquema:

```
<?xml version="1.0"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
  ...
</xs:schema>
```

5.8.2. Componentes básicos de un esquema

xs:element

Es el componente de declaración de elemento y representa la existencia de un elemento en el documento XML.

Atributos optativos principales:

- **name:** indica el nombre del elemento. Es un atributo obligatorio si el elemento padre es `<xs:schema>`.
- **ref:** indica que la descripción del elemento se encuentra en otro lugar del esquema, es decir, referencia a la descripción del elemento. Este atributo no se puede usar si el elemento padre es `<xs:schema>`.
- **type:** indica el tipo del elemento.
- **default:** es el valor que tomará el elemento al ser procesado por alguna aplicación (habitualmente un analizador de XML o un navegador) cuando en el documento instancia XML no había recibido ningún valor. Sólo se puede usar si el contenido del elemento es únicamente textual.
- **fixed:** indica el único valor que puede contener el elemento en el documento instancia XML. Sólo se puede usar si el contenido del elemento es únicamente textual.

5.8.2. Componentes básicos de un esquema

xs:element

Atributos optativos principales (continuación):

- **minOccurs:** indica el número mínimo de ocurrencias que del elemento puede haber en el documento XML. No se puede usar este atributo si el componente padre es <xs:schema>. Va desde 0 hasta ilimitado (unbounded). Posibles valores: 0, 1, 2, ... , unbounded.
- **maxOccurs:** indica el número máximo de ocurrencias que del elemento puede haber en el documento XML. No se puede usar este atributo si el componente padre es <xs:schema>. Va desde 0 hasta ilimitado (unbounded). Posibles valores: 0, 1, 2, ... , unbounded.

Otros atributos optativos:

- **id:** identificador único para el elemento.
- **form:** especifica el formato del nombre del atributo. Puede ser cualificado, es decir, con el espacio de nombres como prefijo del nombre, o no cualificado, sin el espacio de nombres. Posibles valores: unqualified y qualified. El valor por defecto es el del atributo elementFormDefault del componente <xs:schema>.

5.8.2. Componentes básicos de un esquema

xs:element

Otros atributos optativos (continuación):

- **substitutionGroup:** indica el nombre de otro elemento que puede ser sustituido por este elemento. Sólo se puede usar este atributo si el componente padre es `<xs:schema>`.
- **nillable:** especifica si puede aparecer el atributo de instancia `xsi:nil`, asociado al elemento en el documento instancia XML. Por defecto es falso, lo que significa que no se puede asignar a un elemento el atributo de instancia `xsi:nil`. Posibles valores: false y true.
- **abstract:** indica si el elemento puede ser usado en un documento XML instancia. Si vale cierto, indica que el elemento no puede aparecer en una instancia. Posibles valores: false y true.
- **final:** indica si el elemento se puede derivar de alguna manera: por extensión o por restricción o ambas. Posibles valores: extension, restriction y #all.

5.8.2. Componentes básicos de un esquema

Ejemplo:

Se quiere indicar que en el documento instancia XML aparecerá un elemento de nombre <autor>. La manera más genérica es:

```
<?xml version="1.0"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="autor"/>
</xs:schema>
```

El problema de esta descripción es que es demasiado general y no es útil, ya que no detalla si el elemento <autor> es de un tipo determinado, si tiene descendientes, contenido textual ...

Para precisar más, indicaremos que el elemento <autor> es del tipo de datos predefinido xs:string (cadena de texto), deberá aparecer un mínimo de una vez y un máximo de tres.

```
<?xml version="1.0"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="autor" type="xs:string" maxOccurs="3"/>
</xs:schema>
```

5.8.2. Componentes básicos de un esquema

Ejemplo:

Se le asigna un valor por defecto, verde, al elemento <semaforo>.

```
<xs:element name="semaforo" type="xs:string" default="verde"/>
```

Lo que representa esto es:

- a. Si no se le da ningún valor al elemento, la aplicación que procese el documento XML le asignará de manera automática el valor por defecto.

```
<semaforo></semaforo>      —————>      <semaforo>verde</semaforo>
```

- b. Si se le da un valor al elemento, que puede ser distinto al valor por defecto, se quedará con ese valor:

```
<semaforo>amarillo</semaforo>
```

5.8.2. Componentes básicos de un esquema

Ejemplo:

Se le asigna un valor fijo, rojo, al elemento `<semaforo>` y no se le podrá asignar ningún otro valor.

```
<xs:element name="semaforo" type="xs:string" fixed="rojo"/>
```

Lo que representa esto es:

- Si no se le da ningún valor al elemento, la aplicación que procese el documento XML le asignará de manera automática el valor fijo.

`<semaforo></semaforo>` \longrightarrow `<semaforo>rojo</semaforo>`

- Se le da un valor al elemento semáforo, que sólo puede ser el valor fijo. Si se le asignara otro valor distinto, produciría un error de documento no válido:

`<semaforo>rojo</semaforo>` ✓

`<semaforo>amarillo</semaforo>` ✗

5.8.2. Componentes básicos de un esquema

xs:attribute

Es el componente de declaración de atributo y representa la existencia de un atributo de un elemento en el documento XML.

Atributos optativos principales:

- **name:** nombre del atributo. Este atributo no puede aparecer simultáneamente que ref.
- **ref:** referencia a la descripción del atributo que se encuentra en otro lugar del esquema. Si aparece este atributo no aparecerán los atributos type, ni form, ni podrá contener un componente `<xs:simpleType>`.
- **type:** tipo del elemento.
- **use:** indica si la existencia del atributo es opcional, obligatoria o prohibida. Posibles valores: optional, required, prohibited.
- **default:** valor que tomará el atributo al ser procesado por alguna aplicación (habitualmente un analizador de XML o un navegador) cuando en el documento XML no había recibido ningún valor. No puede aparecer simultáneamente con fixed.
- **fixed:** único valor que puede contener el atributo en el documento XML. No puede aparecer simultáneamente con default.

5.8.2. Componentes básicos de un esquema

xs:attribute

Otros atributos optativos:

- **id:** indica un identificador único para el atributo.
- **form:** especifica el formato del nombre del atributo. Puede ser cualificado, es decir, con el espacio de nombres como prefijo del nombre, o no cualificado, sin el espacio de nombres. Posibles valores: `unqualified` y `qualified`. El valor por defecto es el del atributo `attributeFormDefault` del componente `<xs:schema>`.

5.8.2. Componentes básicos de un esquema

Ejemplo:

Atributo de nombre moneda, tipo de datos textual y con valor por defecto Euro. Este valor por defecto se le asigna al atributo si no se le ha asignado otro o no aparece.

```
<xs:attribute name="moneda" type="xs:string" default="Euro"/>
```

Ejemplo:

Atributo de nombre unidad, tipo de datos textual y valor fijo Minutos.

```
<xs:attribute name="unidad" type="xs:string" fixed="Minutos"/>
```

Ejemplo:

Atributo de nombre idEmple, tipo de datos entero positivo y existencia obligatoria.

```
<xs:attribute name="idEmple" type="xs:positiveInteger" use="required"/>
```

5.8.3. Tipos de datos

Se han visto ya en algunos ejemplos que en las declaraciones de elementos y atributos, a estos se les puede asignar un tipo de datos como `xs:string`, `xs:positiveInteger` ...

En los esquemas XML, en la declaración de elementos y atributos se concreta el valor que puedan tomar mediante un tipo de datos. Los tipos de datos se organizan según diversos criterios.

Una de las divisiones es en predefinidos y contruidos:

- Los **predefinidos** son los que vienen integrados en la especificación de los esquemas XML.
- Los **contruidos** son tipos generados por el usuario basándose en un tipo predefinido o en un tipo previamente contruidos.

5.8.3. Tipos de datos

Tipos de datos predefinidos

Existen 44 tipos predefinidos (del inglés built-in types, podría traducirse como "tipos que vienen de serie"). Se organizan en forma de relación jerárquica, a la manera de una jerarquía de clases en programación orientada a objeto, de forma que cada tipo será igual que su tipo padre más alguna particularidad que lo distinga.

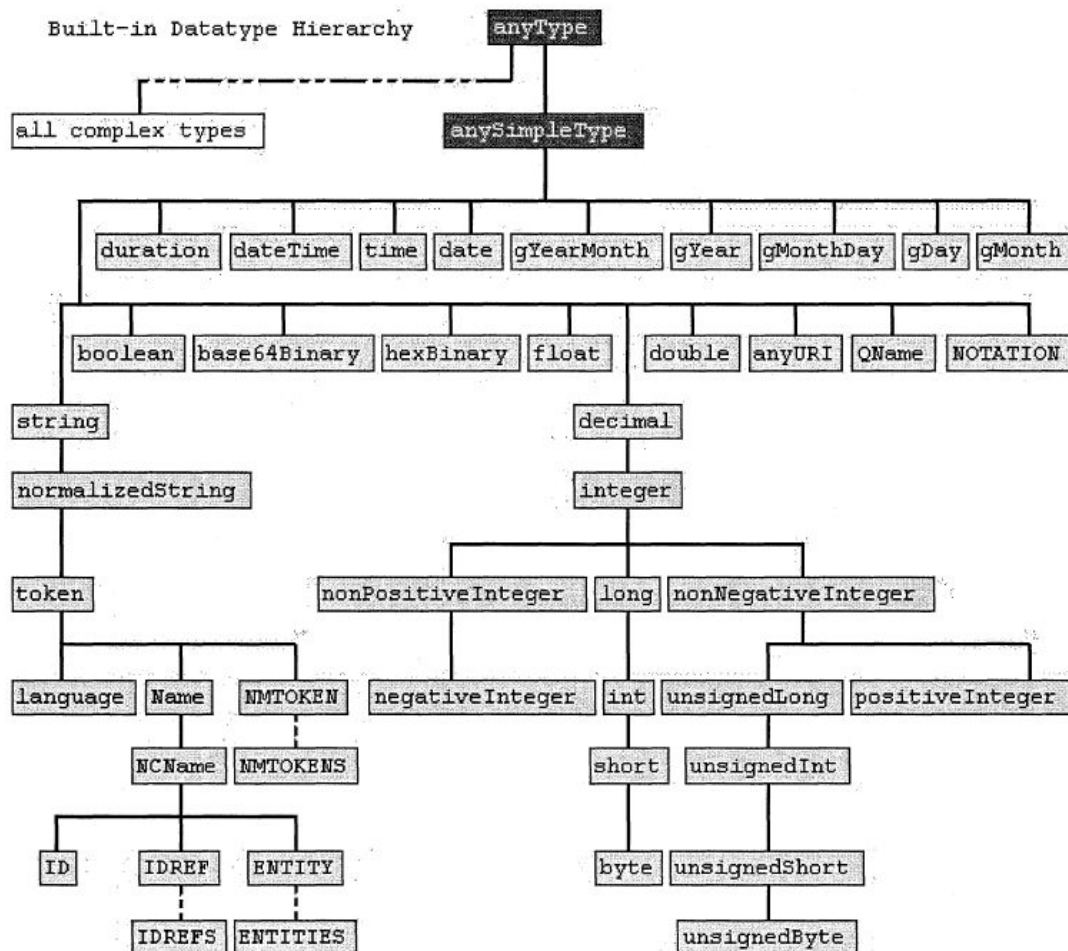
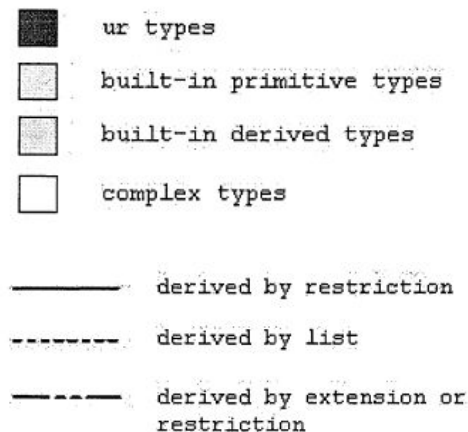
Existe un tipo predefinido especial, **xs:anyType**, que se encuentra en la raíz de la jerarquía de tipos, y del que se derivan todos los demás tipos, tanto predefinidos como complejos.

Este tipo es el más genérico de todos y de cualquier elemento se podría decir que es de ese tipo. Este tipo, en sí mismo, es un tipo complejo, por lo que no se podrían declarar atributos del mismo.

A un elemento del cual no se indica tipo se le asociará el tipo xs:anyType.

Existe otro tipo predefinido especial, xs:anySimpleType, que representa a cualquier tipo simple sin particularizar. Se puede usar para cualquier atributo, y también para elementos que sean de tipo simple pero, de nuevo, es demasiado genérico.

5.8.3. Tipos de datos



5.8.3. Tipos de datos

Tipos de datos predefinidos

Los tipos predefinidos se dividen también en primitivos y no primitivos (derivados de los primitivos).

Los primitivos descienden directamente de `xs: anySimpleType`. Hay 19.

Los no primitivos derivan de alguno de los primitivos. Hay 25.

Los tipos de datos predefinidos se agrupan en cinco categorías, en función de su contenido:

- Numéricos (hay 16)
- De fecha y hora (hay 9)
- De texto (hay 16)
- Binarios (hay 2)
- Booleanos (hay 1)

5.8.3. Tipos de datos

Numéricos

Tipo de datos	Descripción
float	Número en punto flotante de precisión simple (32 bits)
double	Número en punto flotante de precisión doble (64 bits)
decimal	Números reales que pueden ser representados como $i \cdot 10^{-n}$
integer	Números enteros, de $-\infty$ a $+\infty$ ($-\infty \dots -3, -2, -1, 0, 1, 2, 3 \dots +\infty$)
nonPositiveInteger	Números enteros negativos más el 0
negativeInteger	Números enteros menores que 0
nonNegativeInteger	Números enteros positivos más el 0
positiveInteger	Números enteros mayores que 0
unsignedLong	Números enteros positivos desde 0 hasta 18.446.744.073.709.551.615 (64 bits de representación $\rightarrow 2^{64}$ combinaciones)
unsignedInt	Números enteros positivos desde 0 hasta 4.294.967.295 (32 bits de representación $\rightarrow 2^{32}$ combinaciones)
unsignedShort	Números enteros positivos desde 0 hasta 65.535 (16 bits de representación $\rightarrow 2^{16}$ combinaciones)
unsignedByte	Números enteros positivos desde 0 hasta 255 (8 bits de representación $\rightarrow 2^8$ combinaciones)

long	Números enteros representados con 64 bits $\rightarrow 2^{64}$ combinaciones, desde -9.223.372.036.854.775.808 hasta 9.223.372.036.854.775.807
Int	Números enteros representados con 32 bits $\rightarrow 2^{32}$ combinaciones, desde -2.147.483.648 hasta 2.147.483.647
short	Números enteros representados con 16 bits $\rightarrow 2^{16}$ combinaciones, desde -32.768 hasta 32.767
byte	Números enteros representados con 8 bits $\rightarrow 2^8$ combinaciones, desde -128 hasta 127

5.8.3. Tipos de datos

De fecha y hora

Tipo de datos	Descripción
duration	Duración en años + meses + días + horas + minutos + segundos
dateTime	Fecha y hora, en formato aaaa-mm-dd T hh:mm:ss
date	Sólamamente fecha en formato aaaa-mm-dd
time	Sólamamente la hora, en formato hh:mm:ss
gDay	Sólo el día, en formato -dd (la g viene de calendario Gregoriano)
gMonth	Sólo el mes (g → Gregoriano)
gYear	Sólamamente el año, en formato yyyy (g → Gregoriano)
gYearMonth	Sólo el año y el mes, en formato yyyy-mm (g → Gregoriano)
gMonthDay	Sólo el mes y el año, en formato -mm-dd (g → Gregoriano)

El tipo de datos dateTime representa un valor de fecha y hora según las especificaciones de la norma ISO 8601. Es, en sí mismo, la combinación de los otros dos tipos predefinidos time y date.

El patrón de este valor es;

$$\underbrace{(?:\{4\}-\{2\}\{2\}T\{2\}:\{2\}:\{2\}(\backslash\{d+\})?)}_{\text{date}} \underbrace{(([\{+-\}]\{2\}:\{2\})Z)?}_{\text{time}}$$

común a date y time

Patrón	Significado
-?	Un signo negativo (opcional). Para representar fechas anteriores al año 0
\{4\}-\{2\}\{2\}	La parte de la fecha (obligatoria). Equivale a yyyy-mm-dd
\{2\}:\{2\}:\{2\}	La parte de la hora (obligatoria). Equivale a hh:mm:ss
T	Indicador de hora local
(\{d+\})?	La parte decimal de la fracción de segundo (opcional)
(([\{+-\}]\{2\}:\{2\})Z)?	La parte de la zona horaria (opcional) La Z permite expresar las horas en formato UTC (Universal Time Coordinated – Tiempo Universal Coordinado), anteriormente GMT (Greenwich Meridian Time – Hora del Meridiano de Greenwich)

5.8.3. Tipos de datos

Ejemplo:

Para representar una hora que sea las 3:40 PM del horario estándar de la costa este de Estados Unidos, que va 5 horas por detrás del UTC, se escribiría 15:40:00-05:00. Esa misma hora, pero de la zona central de Australia, que va 9 horas y media por delante del UTC, se escribiría como 15:40:00+09:30.

Ejemplo:

Aunque se verá más adelante, se introduce el concepto de faceta, que es una restricción que permite generar nuevos tipos de datos a partir de uno existente, limitando su rango de valores posibles.

Un tipo derivado de `xs:date` que representa la fecha en la que se ha entregado una determinada práctica, siendo el valor de la faceta `xs:maxInclusive` la fecha límite.

```
<!-- Se usa maxInclusive para indicar una límite final de fecha -->
<xs:simpleType name="TipoFechaEntregaPractica">
  <xs:restriction base="xs:date">
    <xs:maxInclusive value="2009-12-31"/>
  </xs:restriction>
</xs:simpleType>
```

5.8.3. Tipos de datos

Ejemplo:

Un tipo derivado de `xs:time` que representa la hora en la que un pasajero ha embarcado en un determinado vuelo, siendo los valores de las facetas `xs:minInclusive` y `xs:maxInclusive` la hora de comienzo y fin del embarque.

```
<!-- Se usa minInclusive y maxInclusive para indicar límites  
      iniciales y finales de horas -->  
<xs:simpleType name="TipoHorarioEmbarque">  
  <xs:restriction base="xs:time">  
    <xs:minInclusive value="05:00:00"/>  
    <xs:maxInclusive value="05:30:00"/>  
  </xs:restriction>  
</xs:simpleType>
```

5.8.3. Tipos de datos

De texto

Tipo de datos	Descripción
string	Cadenas de texto
normalizedString	Cadenas de texto en las que se convierten los caracteres tabulador, nueva línea, y retorno de carro en espacios simples
token	Cadenas de texto sin los caracteres tabulador, ni nueva línea, ni retorno de carro, sin espacios por delante o por detrás, y con espacios simples en su interior
language	Valores válidos para xml:lang (según la especificación de XML)

Ejemplos:

Language	en-US
Name	horaSalida
QName	cliente:nombre
NCName	Nombre
anyURI	http://www.w3c.com
NMTOKENS	SP FR PR IT

NMTOKEN	Tipo de datos para atributo según XML 1.0, compatible con DTD
NMTOKENS	Lista separada por espacios de NMTOKEN, compatible con DTD
Name	Tipo de nombre según XML 1.0
QName	Nombre cualificado de espacio de nombres XML
NCName	QName sin el prefijo ni los dos puntos
anyURI	Cualquier URI
ID	Tipo de datos para atributo según XML 1.0, compatible con DTD. Han de ser valores únicos en el documento XML
IDREF	Tipo de datos para atributo según XML 1.0, compatible con DTD
IDREFS	Lista separada por espacios de IDREF, compatible con DTD
ENTITY	Tipo de datos para atributo en XML 1.0, compatible con DTD
ENTITIES	Lista separada por espacios de ENTITY, compatible con DTD
NOTATION	Tipo de datos para atributo en XML 1.0, compatible con DTD

5.8.3. Tipos de datos

Binarios

Tipo de datos	Descripción
hexBinary	Secuencia de dígitos hexadecimales (0..9,A,B,C,D,E,F)
base64Binary	Secuencia de dígitos en base 64

Booleano

Tipo de datos	Descripción
boolean	Puede tener 4 valores: 0, 1, true y false

5.8.3. Tipos de datos

Todos los tipos de datos predefinidos:

- Pueden ser asignados tanto a elementos como a tributos.
- Pertenecen al espacio de nombres <http://www.w3.org/2001/XMLSchema>.

Ejemplo:

Se declararán tres elementos cuyos tipos de datos, supuestamente predefinidos, están incorrectamente declarados. Finalmente, se declara un elemento con un tipo de datos válido:

```
<xs:element name="elementoA" type="date"/> ❶  
<xs:element name="elementoB" type="w3c:date"/> ❷  
<xs:element name="elementoC" type="xsd:date"/> ❸  
<xs:element name="elementoD" type="xs:date"/> ❹
```

1. El tipo de datos `date` no se puede resolver como ninguno de los componentes de definición de tipos del esquema.
2. El tipo de datos `w3c:date` tiene un prefijo, `w3c`, que no ha sido declarado.
3. El tipo de datos `xsd:date` se encuentra en el espacio de nombres <http://www.w3.org/2001/XMLSchema-datatypes>, pero los componentes existentes en este espacio de nombres no son referenciables desde el esquema. Se ha indicado un espacio de nombres erróneo.
4. El tipo de datos `xs:date` se encuentra en el espacio de nombres <http://www.w3.org/2001/XMLSchema>, que es el adecuado para los tipos de datos predefinidos.

5.8.4. Tipos de datos simples vs. complejos

Aparece aquí una nueva catalogación de tipos de datos, en simples y complejos.

- Tipos de datos simples:
 - Todos los tipos de datos predefinidos son simples.
 - En general representan valores atómicos (el número 7, el texto "Hola"), no listas.
 - Excepcionalmente, se construyen no atómicos, como una lista de números primos (1, 2, 3, 5, 7, 11 ...). Todos los tipos predefinidos son atómicos.
 - Se pueden asignar tanto a elementos que sólo tengan contenido textual como a atributos.
 - También se pueden construir, derivando de un tipo base predefinido al que se le introducen restricciones.

Ejemplo:

Un tipo de datos simple derivado es el que represente un correo electrónico. Se basa en el tipo de datos predefinido `xs: string`, que representa cualquier cadena de texto, pero fuerza a que cumpla una serie de reglas (contener el carácter "@", el carácter "." ...).

5.8.4. Tipos de datos simples vs. complejos

Aparece aquí una nueva catalogación de tipos de datos, en simples y complejos.

- Tipos de datos complejos:
 - Sólo se pueden asignar a elementos que tengan elementos descendientes y/o atributos.
 - Estos elementos pueden tener contenido textual.
 - Por defecto, un elemento con un tipo de datos complejo contiene contenido complejo, lo que significa que tiene elementos descendientes.
 - Un tipo de datos complejo se puede limitar a tener contenido simple, lo que significa que sólo contiene texto y atributos. Se diferencia de los tipos de datos simples precisamente en que tiene atributos.
 - Se pueden limitar para que no tengan contenido, es decir, que sean vacíos, aunque pueden tener atributos.
 - Pueden tener contenido mixto: combinación de contenido textual con elementos descendientes.

5.8.4. Tipos de datos simples vs. complejos

Definición de tipos simples

Se usará el componente de definición de tipos simples, `<xs: simpleType>`. Se podrán limitar el rango de valores con el componente `<xs : restriction>`.

Se pueden asignar a elementos y a atributos.

xs:simpleType

Es el componente de definición de tipos simples.

Atributos optativos principales:

- **name:** nombre del tipo. Este atributo es obligatorio si el componente es hijo de `<xs:schema>`, de lo contrario no está permitido.
- **id:** identificador único para el componente.

Hasta ahora sólo se ha visto cómo asignar un tipo de datos predefinido (que siempre es simple) a un elemento.

5.8.4. Tipos de datos simples vs. complejos

Ejemplo:

Asignar el tipo de datos predefinido `xs:float` al elemento `<ingresosAnuales>`.

```
<xs:element name="ingresosAnuales" type="xs:float"/>
```

La declaración anterior de elemento con un tipo simple asignado es una forma simplificada de la que se muestra a continuación. Se recomienda usar la simplificada.

```
<xs:element name="ingresosAnuales">
  <xs:simpleType>
    <xs:restriction base="xs:float" />
  </xs:simpleType>
</xs:element>
```

5.8.4. Tipos de datos simples vs. complejos

Ejemplo (continuación):

Ahora se puede construir un tipo de datos simple, al que se le asignará un nombre, y que será un sinónimo del tipo de datos predefinido `xs: float`. Se usará el componente de restricción de valores de un tipo simple, `xs: restriction`. Este tipo no aporta nada nuevo `xs:float` y se muestra a modo de ejemplo. No se recomienda definir tipos de datos que sean sinónimos literales de predefinidos.

```
<xs:simpleType name="TipoIngresosAnuales"> ❶  
  <xs:restriction base="xs:float" />  
</xs:simpleType>
```

```
<xs:element name="ingresosAnuales" type="TipoIngresosAnuales"/> ❷
```

1. Se define el tipo simple de nombre `TipoIngresosAnuales`. A partir de ahora se podrá referenciar en todo el esquema y es absolutamente equivalente al predefinido `xs:float`.
2. Se le asigna ese tipo al elemento `<ingresosAnuales>`.

5.8.4. Tipos de datos simples vs. complejos

xs:restriction

Es un componente que define restricciones en los valores de un tipo.

Atributos obligatorios:

- **base:** nombre del tipo base a partir del cual se construirá el nuevo tipo, sea predefinido o construido.

Atributos optativos principales:

- **id:** identificador único para el componente.

5.8.4. Tipos de datos simples vs. complejos

Ejemplo:

Se declara un tipo simple que se utilizará para elementos que sean una contraseña. Será de tipo base `xs:string`, con unas facetas (se comentan a continuación) que fuercen su tamaño mínimo a 6 y máximo a 12.

```
<xs:simpleType name="TipoContrasenia">
  <xs:restriction base="xs:string"> ❶
    <xs:minLength value="6" /> ❷
    <xs:maxLength value="12" /> ❸
  </xs:restriction>
</xs:simpleType>
```

1. Se está construyendo un nuevo tipo simple a base de restringir el rango de valores permitidos del tipo base.
2. Se fuerza a que el nuevo tipo sea de cadenas de longitud mínima de 6. En una faceta.
3. Se fuerza a que el nuevo tipo sea de cadenas de longitud máxima de 6. Es una faceta.

La declaración de un elemento de nombre `<clave>` como del tipo recién creado sería:

```
<xs:element name="clave" type="TipoContrasenia">
```

Un par de ejemplos de uso, uno válido y uno inválido, en un documento instancia XML son:

```
<clave>claveValida!</clave> ✓
```

```
<clave>mala</clave> ✗
```

5.8.4. Tipos de datos simples vs. complejos

Facetas para restringir rangos de valores

Se usan para limitar el rango de valores que tiene un tipo base, lo que produce la aparición de un tipo limitado o facetado.

En función del tipo base, existen unas posibles facetas que se pueden aplicar.

Faceta	Uso	Tipos de datos para donde se usa
xs:minInclusive	Especifica el límite inferior del rango de valores aceptable. El propio valor está incluido.	Numéricos y de fecha/horas
xs:maxInclusive	Especifica el límite superior del rango de valores aceptable. El propio valor está incluido.	Numéricos y de fecha/horas
xs:minExclusive	Especifica el límite inferior del rango de valores aceptable. El propio valor no está incluido.	Numéricos y de fecha/horas
xs:maxExclusive	Especifica el límite superior del rango de valores aceptable. El propio valor no está incluido.	Numéricos y de fecha/horas

5.8.4. Tipos de datos simples vs. complejos

<code>xs:enumeration</code>	Especifica una lista de valores aceptables.	Todos
<code>xs:pattern</code>	Especifica un patrón o expresión regular que deben cumplir los valores válidos.	Texto
<code>xs:whiteSpace</code>	Especifica cómo se tratan los espacios en blanco, entendiéndose como tales los saltos de línea, tabuladores y los propios espacios. Sus posibles valores son <i>preserve</i> , <i>replace</i> y <i>collapse</i> .	Texto
<code>xs:length</code>	Especifica el número exacto de caracteres (o elementos de una lista) permitidos. Ha de ser mayor o igual que 0.	Texto
<code>xs:minLength</code>	Especifica el mínimo número de caracteres (o elementos de una lista) permitidos. Ha de ser mayor o igual que 0.	Texto
<code>xs:maxLength</code>	Especifica el máximo número de caracteres (o elementos de una lista) permitidos. Ha de ser mayor o igual que 0.	Texto
<code>xs:fractionDigits</code>	Especifica el número máximo de posiciones decimales permitidas en números reales. Ha de ser mayor o igual que 0.	Números con parte decimal
<code>xs:totalDigits</code>	Especifica el número exacto de dígitos permitidos en números. Ha de ser mayor que 0.	Numéricos

5.8.4. Tipos de datos simples vs. complejos

Ejemplo:

El elemento `<edadLaboral>` es un entero no negativo, que debe tener un valor mínimo de 16, incluido, y máximo 70, no incluido. En notación matemática se emplean los corchetes para identificar los intervalos cerrados (los extremos están incluidos) y paréntesis para identificar los intervalos abiertos (los extremos no están incluidos). En nuestro caso sería `[16, 70)`

```
<xs:element name="edadLaboral">
  <xs:simpleType>
    <xs:restriction base="xs:nonNegativeInteger">
      <xs:minInclusive value="16"/>
      <xs:maxExclusive value="70"/>
    </xs:restriction>
  </xs:simpleType>
</xs:element>
```

Si quisiéramos definir un tipo de datos al que poder referenciar repetidas veces y declarar un elemento `<edadLaboral>` de ese tipo, sería:

```
<xs:simpleType name="TipoEdadLaboral">
  <xs:restriction base="xs:nonNegativeInteger">
    <xs:minInclusive value="16"/>
    <xs:maxExclusive value="70"/>
  </xs:restriction>
</xs:simpleType>

<xs:element name="edadLaboral" type="TipoEdadLaboral">
```

5.8.4. Tipos de datos simples vs. complejos

Ejemplo:

Se quiere definir un tipo de dato simple, TipoEstaciones, basado en el tipo predefinido xs:token, y que sólomente permita como valores los nombres de las cuatro estaciones.

```
<xs:simpleType name="TipoEstaciones">
  <xs:restriction base="xs:token">
    <xs:enumeration value="Primavera" />
    <xs:enumeration value="Verano" />
    <xs:enumeration value="Otoño" />
    <xs:enumeration value="Invierno" />
  </xs:restriction>
</xs:simpleType>
```

Ejemplo:

Se quiere un tipo de datos, TipoCantidad, basado en xs:decimal, que permita números con 2 dígitos decimales y 11 dígitos totales. El rango de valores implícito sería desde -999.999.999,99 a 999.999.999,99.

```
<xs:simpleType name="TipoCantidad">
  <xs:restriction base="xs:decimal">
    <xs:totalDigits value="11"/>
    <xs:fractionDigits value="2"/>
  </xs:restriction>
</xs:simpleType>
```

5.8.4. Tipos de datos simples vs. complejos

Ejemplo:

Se define un tipo basado en `xs:string`, usado para registrar direcciones en el que se quiere que todos los caracteres de espaciado (espacio, tabulador, nueva línea y retorno de carro) se colapsen, lo que significa que si estos caracteres aparecen al principio o final de la cadena se eliminan, y si lo hacen en su interior, formando una hilera, se sustituyen en un solo espacio.

La definición de este tipo coincide con el tipo predefinido `xs:token`, derivado de `xs:string`.

```
<xs:simpleType name="TipoDireccionFormateada">
  <xs:restriction base="xs:string">
    <xs:whiteSpace value="replace"/>
  </xs:restriction>
</xs:simpleType>
```

5.8.4. Tipos de datos simples vs. complejos

Una faceta que permite usos muy precisos es `xs:pattern`, que se analiza en detalle.

xs:pattern

Representa un patrón o expresión regular que debe de cumplir la cadena de texto a la que se aplique.

Ejemplo: Se quiere definir un tipo de datos, basado en `xs:string`, que se caracterice por cadenas de tres letras mayúsculas.

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

La sintaxis usada para estos patrones se basa en expresiones regulares. Las más usadas son:

Patrón	Significado	Ejemplo
.	Cualquier carácter	;
\w	Cualquier letra	M
\d	Un dígito	7
\D	Cualquier carácter no dígito	j
\s	Cualquier carácter de espaciado (tabulador, espacio...)	
\S	Cualquier carácter de no espaciado	C
\d{n}	n dígitos exactamente (Ej. cuatro dígitos exactamente)	3856

5.8.4. Tipos de datos simples vs. complejos

<code>\d{n,m}</code>	De n a m dígitos (Ej. de dos o tres dígitos)	91
<code>\d{n,}</code>	n o más dígitos (Ej. Tres o más dígitos)	3500
<code>[xyz]</code>	Uno de los caracteres x, y o z (en minúscula)	Y
<code>[A-Z]</code>	Uno de los caracteres de la A a la Z (en mayúscula)	
<code>[^abc]</code>	Negación de un grupo de caracteres	D
<code>[F-J-[H]]</code>	Sustracción de un carácter de un rango	G
<code>(a b)</code>	Alternativa entre dos expresiones	b
<code>b?</code>	Sucesión de 0 o una ocurrencias de una cadena	B
<code>1*</code>	Sucesión de 0 o más ocurrencias de una cadena	111
<code>(cd)+</code>	Sucesión de 1 o más ocurrencias de una cadena	cdcd

Como se ha visto en los patrones anteriores, en las facetas se utilizan cuantificadores:

Cuantificador	Significado
<code>?</code>	0 ó 1 ocurrencias
<code>*</code>	0 o más ocurrencias
<code>+</code>	1 o más ocurrencias
<code>{n}</code>	n ocurrencias exactas
<code>{n,m}</code>	De n a m ocurrencias (siendo $n < m$)
<code>{n,}</code>	n o más ocurrencias

5.8.4. Tipos de datos simples vs. complejos

Hay ciertos caracteres que tienen un significado propio en las expresiones regulares: {, }, [,], (,), ?, *, +, -, |, ^, ., \

Para indicar en una expresión regular la aparición literal de uno de estos caracteres, hay que anteponerle el carácter de escape \.

Ejemplo:

La expresión regular para una cadena que contenga las letras a o b o el carácter + sería:

```
[ab\+]
```

Ejemplo:

Un número de teléfono que se quiere se represente como 3 dígitos, un punto, 3 dígitos, un punto, tres dígitos y un punto. Por ejemplo 912.345.678:

```
<xs:simpleType name="TipoTelefonoPunteado">  
  <xs:restriction base="xs:string">  
    <xs:pattern value="\d{3}\.\d{3}\.\d{3}"/>  
  </xs:restriction>  
</xs:simpleType>
```

5.8.4. Tipos de datos simples vs. complejos

Actividad: Define diferentes tipos simples a partir de las siguientes especificaciones.

1. Una cadena de texto que represente las matrículas españolas anteriores a la normalización del año 2000. El formato era: una o dos letras mayúsculas + un guión + cuatro dígitos + un guión + una o dos letras mayúsculas. (ej. Z-1234-AB)
2. Una cadena de texto que represente las cuatro posibles formas de pago: con tarjeta VISA, MasterCard, American Express y en Efectivo.

A continuación, crea un esquema en el que sólo haya un elemento al que se le asigne un tipo simple de los definidos.

Por último, escribe un documento instancia XML que sea válido con respecto a ese esquema.

Repítelo para los distintos tipos definidos.

5.8.4. Tipos de datos simples vs. complejos

Más sobre derivación de tipos simples: uniones y listas

Otra forma de derivar un tipo de datos consiste en definir uniones o listas de tipos de datos base:

- **Unión:** es un tipo de datos creado a partir de una colección de tipos de datos base. Un valor es válido para una unión de tipos de datos si es válido para al menos uno de los tipos de datos que forman la unión. Se construye con el componente `<xs:union>`.

Ejemplo: Se quiere reflejar que las tallas de ropa se pueden identificar por números (38, 40, 42) o por letras, que son las iniciales en inglés (S, M, L).

```
<xs:element name="talla">
  <xs:simpleType>
    <xs:union memberTypes="TipoTallaNumerica TipoTallaTextual" />
  </xs:simpleType>
</xs:element>

<xs:simpleType name="TipoTallaNumerica">
  <xs:restriction base="xs:positiveInteger">
    <xs:enumeration value="38"/>
    <xs:enumeration value="40"/>
    <xs:enumeration value="42"/>
  </xs:restriction>
</xs:simpleType>

<xs:simpleType name="TipoTallaTextual">
  <xs:restriction base="xs:string">
    <xs:enumeration value="S"/>
    <xs:enumeration value="M"/>
    <xs:enumeration value="L"/>
  </xs:restriction>
</xs:simpleType>
```

Cualquiera de estos dos fragmentos XML serían válidos:

```
<talla>42</talla>
```

Y también:

```
<talla>M</talla>
```

5.8.4. Tipos de datos simples vs. complejos

- **Lista:** es un tipo de datos compuesto por listas de valores de un tipo de datos base. La lista de valores debe aparecer separada por espacios. Se construye con el componente `<xs:list>`.

Ejemplo:

Se define un tipo `TipoListaNumerosBingo` como una lista de valores enteros positivos que empieza en el 1 y acaba en el 90.

```
<xs:simpleType name="TipoListaNumerosBingo">
  <xs:list>
    <xs:restriction base="xs:positiveInteger">
      <xs:maxInclusive value="90"/>
    </xs:restriction>
  </xs:list>
</xs:simpleType>

<xs:element name="numerosGanan" type="TipoListaNumerosBingo"/>
```

Un fragmento XML que sería válido para la definición precedente será:

```
<numerosGanadores>1 7 19 34 42 60 63 73</numerosGanadores>
```

5.8.4. Tipos de datos simples vs. complejos

Actividad: Define diferentes tipos simples a partir de las siguientes especificaciones. A continuación, crea un esquema en el que sólo haya un elemento al que se le asigne un tipo simple de los definidos. Por último, escribe un documento instancia XML que sea válido con respecto a ese esquema. Repítelo para los distintos tipos definidos:

1. Un número real con tres decimales que represente las temperaturas posibles en la Tierra, suponiendo que van desde -75° a 75° , ambas inclusive.
2. Un xs: token que sólo pueda valer las siglas de los países vecinos de España, incluyendo a la propia España: ES, PR, FR, AN.
3. Un número real que represente salarios, con 5 dígitos enteros y 2 decimales.
4. Un mensaje de la red social twitter, conocido como tweet (trino), es una cadena de texto de una longitud máxima de 140 caracteres.

5.8.4. Tipos de datos simples vs. complejos

Actividad: Define diferentes tipos simples a partir de las siguientes especificaciones.

1. Una cadena de texto que represente las matrículas españolas anteriores a la normalización del año 2000. El formato era: una o dos letras mayúsculas + un guión + cuatro dígitos + un guión + una o dos letras mayúsculas. (ej. Z-1234-AB)
2. Una cadena de texto que represente las cuatro posibles formas de pago: con tarjeta VISA, MasterCard, American Express y en Efectivo.

A continuación, crea un esquema en el que sólo haya un elemento al que se le asigne un tipo simple de los definidos.

Por último, escribe un documento instancia XML que sea válido con respecto a ese esquema.

Repítelo para los distintos tipos definidos.

5.8.4. Tipos de datos simples vs. complejos

Actividad: Dado un tipo simple, `TipoColoresSemaforo`, basado en `xs:string` y que sólo pueda tomar los valores Rojo, Amarillo y Verde, define un nuevo tipo `ListaColoresSemaforo` que sea una lista cuyos elementos sean del tipo `TipoColoresSemaforo`.

Por último, crea otro tipo que sea una lista de sólo tres colores de semáforo, `Tipo3ColoresSemaforo`.

Crea un esquema que contenga estos tipos y declarar un elemento del tipo recién construido `Tipo3ColoresSemaforo`.

Crea un documento XML instancia del esquema, que contenga un único elemento de ese tipo.

Comprueba que es válido respecto al esquema.

5.8.5. Definición de tipos de datos complejos

Se usará el componente de definición de tipos complejos, `<xs:complexType>`. Sólo se pueden asignar a elementos. Un elemento se considera de tipo complejo si tiene atributos y/o descendientes.

El contenido de un elemento es aquello que va entre sus etiquetas de apertura y de cierre. Un elemento puede tener el siguiente contenido:

- **Contenido simple** (`xs:simpleContent`): el elemento declarado sólo tiene contenido textual, sin elementos descendientes.
- **Contenido complejo** (`xs:complexContent`): el elemento declarado tiene elementos descendientes. Puede tener o no contenido textual.

5.8.5. Definición de tipos de datos complejos

Modelos de contenido

En el momento en el que se habla de elementos descendientes, aparece otro elemento en los esquemas que son los compositores o modelos de contenido.

Indican la distribución que tienen entre sí los elementos descendientes de uno dado. Siempre aparecen conformando un tipo de datos complejo. Hay tres posibilidades distintas, cada una de ellas declarada con un componente:

- **Secuencia:**

Los elementos aparecen en fila, unos detrás de otros, en un orden determinado. Se declara con el componente `<xs:sequence>`. Al igual que sucedía con los elementos, la secuencia puede aparecer un número variable de veces, configurable con los atributos `minOccurs` y `maxOccurs` (ambos valen 1 por defecto).

5.8.5. Definición de tipos de datos complejos

- **Secuencia:**

Ejemplo:

El <mensaje> con elementos descendientes <emisor>, <receptor> y <contenido>.

```
<xs:sequence>
  <xs:element name="emisor" type="xs:string"/>
  <xs:element name="receptor" type="xs:string"/>
  <xs:element name="contenido" type="xs:string"/>
</xs:sequence>
```

5.8.5. Definición de tipos de datos complejos

- **Alternativa:**

Los elementos aparecen como alternativa unos de los otros. Sólo se elige uno. Se declara con el componente `<xs:choice>`. La alternativa puede aparecer un número variable de veces, configurable con los atributos `minOccurs` y `maxOccurs` (ambos valen 1 por defecto).

Ejemplo:

Al comienzo de una novela podrá haber un elemento `<prologo>`, 0 un elemento `<prefacio>` o un elemento `<introduccion>`, sólo uno de los tres, todos textuales.

```
<xs:choice>
  <xs:element name="prologo" type="xs:string"/>
  <xs:element name="prefacio" type="xs:string"/>
  <xs:element name="introduccion" type="xs:string"/>
</xs:choice>
```

5.8.5. Definición de tipos de datos complejos

- **Todos:**

Los elementos aparecen en cualquier orden. Se declara con el componente `<xs:all>`. El uso de este modelo de contenido no se recomienda por la pérdida de control sobre el orden de los elementos. La aparición de todos los elementos puede aparecer 0 o 1 vez, configurable con los atributos `minOccurs` y `maxOccurs` (ambos valen 1 por defecto).

Ejemplo:

Supóngase que existen dos elementos `<alfa>` y `<omega>`, que tienen que aparecer como descendientes de `<griego>`, pero el orden es indiferente. Sería:

```
<xs:all>
  <xs:element name="alfa" type="xs:string"/>
  <xs:element name="omega" type="xs:string"/>
</xs:all>
```

5.8.5. Definición de tipos de datos complejos

xs:complexType

Es el componente de definición de tipos complejos.

Atributos optativos principales:

- **name:** nombre del tipo. Este atributo es obligatorio SI el componente es hijo de <xs:schema>, de lo contrario no está permitido.
- **mixed:** indica si se intercala contenido textual con los elementos descendientes: Posibles valores: false, true.
- **id:** identificador único para el componente.
- **abstract:** indica si se puede usar directamente como tipo de un elemento de un documento instancia XML. Si vale cierto, significa que no puede usarse directamente y el tipo debe derivarse para generar otro tipo que sí se pueda usar. Posibles valores: false, true.
- **final:** indica si el tipo puede ser derivable por extensión por restricción o por ambas. Posibles valores: extension, restriction, #all.

5.8.5. Definición de tipos de datos complejos

Ejemplo:

Un tipo complejo para un elemento <persona> que contenga una secuencia de elementos descendientes <primerApellido>, <segundoApellido> y <fechaNacimiento>, de los tipos esperables.

```
<xs:element name="persona">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="primerApellido" type="xs:string" />
      <xs:element name="primerApellido" type="xs:string" />
      <xs:element name="fechaNacimiento" type="xs:date" />
    </xs:sequence>
  </xs:complexType>
</xs:element>
```

Como ya se ha comentado previamente, se podría definir un TipoPersona y declarar un elemento de ese tipo:

```
<xs:complexType name="TipoPersona">
  <xs:sequence>
    <xs:element name="primerApellido" type="xs:string" />
    <xs:element name="primerApellido" type="xs:string" />
    <xs:element name="fechaNacimiento" type="xs:date" />
  </xs:sequence>
</xs:complexType>

<xs:element name="persona" type="TipoPersona">
```

5.8.6. Diferentes declaraciones de elementos

Para tipos simples y complejos, combinando las posibilidades de contenido (simple y complejo) y los elementos que condicionan el tipo y el contenido (atributos, elementos descendientes, contenido textual) tendríamos la siguiente tabla:

Tipo	Contenido	Elementos descendientes	Atributos	Contenido textual
Simple	Simple			✓
Complejo	Simple		✓	✓
Complejo	Complejo			
Complejo	Complejo		✓	
Complejo	Complejo	✓		
Complejo	Complejo	✓	✓	
Complejo	Complejo mixto	✓		✓
Complejo	Complejo mixto	✓	✓	✓

5.8.6. Diferentes declaraciones de elementos

Declaración de elementos vacíos

Un elemento es vacío (EMPTY en los DTD) si no tiene contenido textual, ni elementos descendientes (aunque podría tener atributos).

Hay diversas maneras de representarlo:

- Como un tipo simple derivado de `xs:string` con longitud 0.
- Como un tipo complejo sin contenido, que no contiene elementos descendientes. Se aconseja esta.

Ejemplo:

Para el elemento `
` de HTML, se declara un tipo complejo sin contenido.

```
<xs:element name="br">
  <xs:complexType />
</xs:element>
```


5.8.6. Diferentes declaraciones de elementos

Declaración de elementos con contenido textual y atributos

Se declarara con un tipo de datos complejo con contenido simple que contenga algún atributo.

Lo que se hace realmente es tomar un tipo de datos simple y, mediante un componente de declaración de extensión, `<xs:extension>`, se le añade uno o más atributos, lo que hace que el tipo se convierta en complejo.

Ejemplo:

El elemento `<textarea>` de HTML, que acepta contenido textual y tiene atributos (name, cols, rows ...). Un fragmento de un documento HTML sería:

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

5.8.6. Diferentes declaraciones de elementos

La definición del tipo complejo para un elemento <textarea> sería:

```
<xs:element name="textarea">
  <xs:complexType>
    <xs:simpleContent> ❶
      <xs:extension base="xs:string"> ❷
        <xs:attribute name="name" type="xs:string"/>
        <xs:attribute name="cols" type="xs:positiveInteger"/>
        <xs:attribute name="rows" type="xs:positiveInteger"/>
      </xs:extension>
    </xs:simpleContent>
  </xs:complexType>
</xs:element>
```

1. El componente de definición de contenido simple especifica que el elemento complejo no debe tener elementos descendientes.
2. El componente de definición de extensión, <xs: extension>, extiende el tipo de datos indicado como base para que el contenido incluya declaración de atributos.

5.8.6. Diferentes declaraciones de elementos

Declaración de elementos sólo con atributos

Se declara con un tipo de datos complejo con contenido complejo que contenga algún atributo.

Ejemplo:

El elemento `` de HTML, que tiene atributos (`src`, `width`, `height` ...) pero no tiene elementos descendientes ni contenido textual.

```

```

La declaración sería:

```
<xs:element name="img">
  <xs:complexType>
    <xs:attribute name="src" type="xs:string"/>
    <xs:attribute name="alt" type="xs:string"/>
  </xs:complexType>
</xs:element>
```

5.8.6. Diferentes declaraciones de elementos

Declaración de elementos sólo con elementos descendientes

Se declarara con un tipo de datos complejo con contenido complejo que contenga algún elemento.

Ejemplo:

El elemento `` de HTML, que tiene elementos descendientes `` pero no tiene atributos ni contenido textual. Un fragmento en un documento HTML sería:

```
<ul>
  <li>Primer elemento</li>
  <li>Segundo elemento</li>
</ul>
```

La declaración del elemento es:

```
<xs:element name="ul">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="li" type="xs:string" maxOccurs="unbounded"/>
    </xs:sequence>
  </xs:complexType>
</xs:element>
```

5.8.6. Diferentes declaraciones de elementos

Declaración de elementos con atributos y elementos descendientes

Se declara con un tipo de datos complejo con contenido complejo que contenga algún elemento descendiente y algún atributo.

Ejemplo:

El elemento `<table>` de HTML, que tiene atributos, como `border`, y elementos descendientes, como `<tr>`, pero no tiene contenido textual. Un fragmento de un documento HTML podría ser:

```
<table border="1">
  <tr>
    <td>Primera celda</td>
    ...
  </tr>
  ...
</table>
```

```
<xs:element name="table">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="tr" maxOccurs="unbounded"/>
    </xs:sequence>
    <xs:attribute name="border" type="xs:string"/>
  </xs:complexType>
</xs:element>
<xs:element name="tr">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="td" type="xs:string"/>
    </xs:sequence>
  </xs:complexType>
</xs:element>
```

5.8.6. Diferentes declaraciones de elementos

Declaración de elementos con contenido textual y elementos descendientes

Se declarara con un tipo de datos complejo con contenido complejo que contenga algún elemento descendiente y contenido textual.

Ejemplo:

El elemento `<p>` de HTML, suponiendo que no tuviese atributos, puede contener texto directamente, o texto en negrita ``, cursiva `<i></i>`, etc.

```
<p>
  El perro de <b>San Roque</b> no tiene rabo porque <i>Ramón
Rodríguez</i> se lo ha robado.
</p>
```

Para indicar que el contenido es mixto, textual y elementos descendientes, se fija atributo `mixed="true"`. La declaración será:

```
<xs:element name="p">
  <xs:complexType mixed="true"> ❶
    <xs:choice minOccurs="0" maxOccurs="unbounded"> ❷
      <xs:element name="b" type="xs:string" />
      <xs:element name="i" type="xs:string" />
    </xs:choice>
  </xs:complexType>
</xs:element>
```

1. Aparece el atributo `mixed="true"` para indicar un contenido mixto.
2. Aparece el modelo de contenido `<xs:choice>`, fijando el valor de sus atributos `minOccurs="0"` y `maxOccurs="unbounded"`, para permitir un número indefinido de ocurrencias del elemento `` o del elemento `<i>`, en el orden que sea.

5.8.6. Diferentes declaraciones de elementos

Declaración de elementos con atributos, elementos descendientes y contenido textual

Se declara con un tipo de datos complejo con contenido complejo que contenga elementos, atributos y contenido textual.

Ejemplo:

El elemento `<form>` de HTML, que tiene atributos (`name`, `method` ...), elementos descendientes como `<input>` y contenido textual.

```
<form name="f1" method="post">
  Apellido: <input type="text" name="apellido" />
  ...
</form>
```

Para indicar que el contenido es mixto, textual y elementos descendientes, se fija el atributo `mixed="true"`. La declaración será:

```
<xs:element name="form">
  <xs:complexType mixed="true">
    <xs:choice minOccurs="0" maxOccurs="unbounded">
      <xs:element ref="input"/>
    </xs:choice>
    <xs:attribute name="name" type="xs:string"/>
    <xs:attribute name="method" type="xs:string"/>
  </xs:complexType>
</xs:element>
<xs:element name="input">
  <xs:complexType>
    <xs:attribute name="type" type="xs:string" use="required"/>
    <xs:attribute name="name" type="xs:string" use="required"/>
  </xs:complexType>
</xs:element>
```

5.8.6. Diferentes declaraciones de elementos

Extensión de un tipo complejo

La extensión de tipos complejos se asemeja a la herencia en la programación orientada a objeto: se pueden añadir nuevos elementos y atributos a tipos complejos ya existentes. Los elementos añadidos aparecerán al final de todos los elementos del tipo base.

Ejemplo:

Se dispone de un tipo complejo TipoAlumno con la siguiente definición:

```
<xs:complexType name="TipoAlumno">
  <xs:sequence>
    <xs:element name="Apellido" type="xs:string" />
    <xs:element name="Ciclo" type="xs:string" />
  </xs:sequence>
</xs:complexType>
```

Se le quiere añadir un elemento numérico <Curso> y un atributo que sea el número de Matrícula.

5.8.6. Diferentes declaraciones de elementos

Extensión de un tipo complejo

La definición completa del tipo base y el tipo extendido es:

```
<xs:complexType name="TipoAlumno">
  <xs:sequence>
    <xs:element name="Apellido" type="xs:string" />
    <xs:element name="Ciclo" type="xs:string" />
  </xs:sequence>
</xs:complexType>

<xs:complexType name="TipoAlumnoExtendido">
  <xs:complexContent>
    <xs:extension base="TipoAlumno">
      <xs:sequence>
        <xs:element name="Curso" type="xs:positiveInteger" />
      </xs:sequence>
      <xs:attribute name="Matrícula" type="xs:positiveInteger"
        use="required" />
    </xs:extension>
  </xs:complexContent>
</xs:complexType>
```

5.8.7. Modelos de diseño de esquemas XML

Existen varios modelos de estructuración de las declaraciones al construir esquemas, si bien conviene recordar que el orden en que aparecen los componentes en un esquema no es representativo.

- **Diseño anidado o de muñecas rusas:** se llama así porque se anidan declaraciones unas dentro de otras, como las muñecas Matrioskas. Se describe cada elemento y atributo en el mismo lugar donde se declaran. Esto produce duplicidades en la descripción de elementos con tipos iguales y puede haber elementos con igual nombre y distintas descripciones. Con esta técnica los esquemas son más cortos pero su lectura es más compleja para el ojo humano.

```
<?xml version="1.0" encoding="utf-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="libro">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="titulo" type="xs:string"/>
        <xs:element name="autor" type="xs:string"/>
        <xs:element name="personaje" minOccurs="0"
                      maxOccurs="unbounded">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="nombre" type="xs:string"/>
              <xs:element name="amigoDe" type="xs:string"
                            minOccurs="0" maxOccurs="unbounded"/>
              <xs:element name="desde" type="xs:date"/>
              <xs:element name="calificacion" type="xs:string"/>
            </xs:sequence>
          </xs:complexType>
        </xs:element>
      </xs:sequence>
      <xs:attribute name="isbn" type="xs:string"/>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

5.8.7. Modelos de diseño de esquemas XML

- **Diseño plano:** se declaran los elementos y los atributos y se indica una referencia a su definición, que se realiza en otro lugar del documento. Es una técnica que recuerda a los DTD.

```
<?xml version="1.0" encoding="io-8859-1"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">

  <!-- definición de elementos de tipo simple -->
  <xs:element name="titulo" type="xs:string"/>
  <xs:element name="autor" type="xs:string"/>
  <xs:element name="nombre" type="xs:string"/>
  <xs:element name="amigoDe" type="xs:string"/>
  <xs:element name="desde" type="xs:date"/>
  <xs:element name="calificacion" type="xs:string"/>

  <!-- definición de atributos -->
  <xs:attribute name="isbn" type="xs:string"/>

  <!-- definición de elementos de tipo complejo -->
  <xs:element name="personaje">
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="nombre"/>
        <xs:element ref="amigoDe" minOccurs="0"
                      maxOccurs="unbounded"/>
        <xs:element ref="desde"/>
        <xs:element ref="calificacion"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

```
    <!-- los elementos de tipo simple se referencian con el
         atributo ref -->
    <!-- la definición de la cardinalidad se hace cuando el
         elemento es referenciado -->
  </xs:sequence>
</xs:complexType>
</xs:element>
<xs:element name="libro">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="titulo"/>
      <xs:element ref="autor"/>
      <xs:element ref="personaje" minOccurs="0"
                                maxOccurs="unbounded"/>
    </xs:sequence>
    <xs:attribute ref="isbn"/>
  </xs:complexType>
</xs:element>
</xs:schema>
```

5.8.7. Modelos de diseño de esquemas XML

- **Diseño con tipos con nombre reutilizables:** se definen tipos de datos simples o complejos a los que se identifica con un nombre (son plantillas, como las clases en la programación orientada a objeto). Al declarar elementos y atributos, se indica que son de alguno de los tipos con nombre previamente definidos (aquí elementos y atributos son instancias de las plantillas ya definidas, como objetos en la programación orientada a objeto).

```
<?xml version="1.0" encoding="iso-8859-1"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">

  <!-- definición de tipos simples -->
  <xs:simpleType name="TipoNombre">
    <xs:restriction base="xs:string">
      <xs:maxLength value="32"/>
    </xs:restriction>
  </xs:simpleType>
  <xs:simpleType name="TipoDesde">
    <xs:restriction base="xs:date"/>
  </xs:simpleType>
  <xs:simpleType name="TipoDescripcion">
    <xs:restriction base="xs:string"/>
  </xs:simpleType>
  <xs:simpleType name="TipoISBN">
    <xs:restriction base="xs:string">
      <xs:pattern value="[0-9]{10}"/>
    </xs:restriction>
  </xs:simpleType>

  <!-- definición of tipos complejos -->
  <xs:complexType name="TipoPersonaje">
    <xs:sequence>
      <xs:element name="nombre" type="TipoNombre"/>
      <xs:element name="amigoDe" type="TipoNombre"
        minOccurs="0"
        maxOccurs="unbounded"/>
      <xs:element name="desde" type="TipoDesde"/>
    </xs:sequence>
  </xs:complexType>

  <xs:complexType name="TipoLibro">
    <xs:sequence>
      <xs:element name="calificación" type="TipoDescripcion"/>
      <xs:element name="titulo" type="TipoNombre"/>
      <xs:element name="autor" type="TipoNombre"/>
    </xs:sequence>
    <xs:attribute name="isbn" type="TipoISBN" use="required"/>
  </xs:complexType>

  <!-- definición del elemento personaje, usando el tipo
  complejo TipoPersonaje -->
  <xs:element name="personaje" type="TipoPersonaje"
    minOccurs="0"/>

  <!-- definición del elemento libro usando el tipo complejo
  TipoLibro -->
  <xs:element name="libro" type="TipoLibro"/>
</xs:schema>
```

5.8.8. Poniendo todo junto

Se va a desarrollar un ejemplo de esquema para recopilar los elementos más usados.

Ejemplo:

Se dispone del siguiente documento XML, persona.xml:

```
<?xml version="1.0" encoding="iso-8859-1"?>
<persona dni="12345678-L" xsi:noNamespaceSchemaLocation="persona.xsd"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <nombre>Juan Antonio</nombre>
  <apellido>Abascal</apellido>
  <estadoCivil>Soltero</estadoCivil>
  <edad>60</edad>
  <enActivo/>
</persona>
```

El esquema XML asociado se quiere que cumpla ciertas restricciones semánticas:

- El atributo dni es obligatorio y su formato es de 8 dígitos, un guion y una letra mayúscula.
- El estado civil puede ser: Soltero, Casado, Divorciado o Viudo. Por defecto es Soltero.
- La edad debe ser de 0 a 150.
- El elemento <enActivo> es optativo.

5.8.8. Poniendo todo junto

El esquema que cumple con esto, ubicado en el archivo persona.xsd, es:

```
<?xml version="1.0" encoding="utf-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">

  <!-- Declaración de elementos -->
  <xs:element name="persona">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="nombre" type="xs:string"/>
        <xs:element name="apellido" type="xs:string"/>
        <xs:element name="estadoCivil" type="TipoEstadoCivil"
          default="Soltero"/>
        <xs:element name="edad" type="TipoEdadHumana"/>
        <xs:element name="enActivo" minOccurs="0">
          <xs:complexType/>
        </xs:element>
      </xs:sequence>
      <xs:attribute name="dni" type="TipoDni" use="required"/>
    </xs:complexType>
  </xs:element>
```

```
<!-- Definición de tipos -->
<xs:simpleType name="TipoDni">
  <xs:restriction base="xs:string">
    <xs:pattern value="[0-9]{8}\-[A-Z]"/>
  </xs:restriction>
</xs:simpleType>

<xs:simpleType name="TipoEstadoCivil">
  <xs:restriction base="xs:string">
    <xs:enumeration value="Soltero"/>
    <xs:enumeration value="Casado"/>
    <xs:enumeration value="Divorciado"/>
    <xs:enumeration value="Viudo"/>
  </xs:restriction>
</xs:simpleType>

<xs:simpleType name="TipoEdadHumana">
  <xs:restriction base="xs:nonNegativeInteger">
    <xs:maxInclusive value="150"/>
  </xs:restriction>
</xs:simpleType>

</xs:schema>
```