Tema 04 - DTD

Parte 2

9. DTD's

En una empresa tenemos dos programadores que programan así:

Programador 1

Programador 2

9. DTD's

Está claro, que ninguno de los dos puede leer los archivos del otro. Por eso, resulta crítico ponerse de acuerdo en lo que se puede hacer, lo que puede aparecer y en qué orden debe hacerlo. Esto se hará mediante las DTD.

DTD significa Declaración de Tipo de Documento, y es un conjunto de reglas sintácticas para definir etiquetas. Nos indica qué etiquetas se pueden usar en un documento, en qué orden deben aparecer, cuáles pueden aparecer dentro de otras, cuáles tienen atributos, etc. Toda esta información se guarda en un archivo de texto con extensión .dtd.

9. DTD's

Un DTD:

- No es un documento XML y por lo tanto no necesita prólogo, salvo en ciertos casos especiales.
- Crear un DTD es como crear nuestro propio lenguaje de marcado para una aplicación específica. Pueden ser parte del documento XML, pero se suele colocar aparte para poder reutilizarlo.
- Puesto que XML es un sistema para definir lenguajes, quien necesite usar XML para intercambio de datos debe definir su propio DTD.
- Un problema que presentan los DTDs es que no siguen la sintaxis XML, sino una propia.

- Supongamos que en nuestros ficheros deseamos indicar que el elemento raíz es < listaclientes > .
- Dentro de deseamos permitir uno o más elementos <cli><cliente>.
- Dentro de <cliente> todos deberán tener <cif> y <nombre> y en ese orden.
- Dentro de <cliente> puede aparecer o no un elemento <diasentrega> para indicar que ese cliente exige un máximo de plazos.
- Como no todo el mundo usa plazos el <diasentrega> es optativo.
- Se puede indicar si un elemento aparece o no de forma opcional (usando?)

¿Cuáles son válidos y cuáles no?

```
<listaclientes>
</listaclientes>
```

Crear el DTD

Para ello creamos un fichero de texto con extensión .dtd.

Vamos describiendo cada elemento uno a uno desde el nodo raíz hasta el último elemento.

Declaración DOCTYPE

Esta declaración es necesaria en el documento XML. Sirve para especificar a que DTD está unido el documento. Siempre debe ir colocado después del prólogo y antes del nodo raíz.

Hay tres formas de especificar un DTD:

- **SYSTEM:** Se utiliza en el caso de DTDs "personales". Para especificar un fichero local:

```
<!DOCTYPE nodo_raiz SYSTEM "fichero.dtd">
```

- **PUBLIC:** Se utiliza en el caso de una DTD pública.

```
<!DOCTYPE nodo_raiz PUBLIC "-//OASIS//DTD Libro XML//EN" "../dtds/cap.dtd">
```

También es posible incrustar la DTD internamente en el documento XML:

DTD incrustado en el documento XML:

Declaración de elementos

 Los elementos permitidos se especifican con ELEMENT, seguido del nombre y el tipo de elemento:

<!ELEMENT nombre_elemento tipo_elemento>

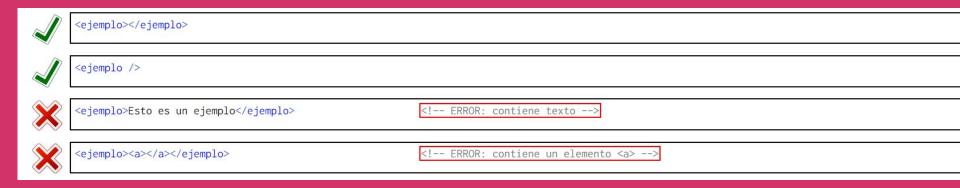
- **EMPTY:** significa que el elemento es vacío, es decir, que no puede tener contenido.
- **(#PCDATA):** significa que el elemento puede contener texto. #PCDATA debe escribirse entre paréntesis.
- **ANY:** significa que el elemento puede contener cualquier cosa (texto y otros elementos). ANY debe escribirse sin paréntesis.

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo EMPTY>
]>
```

¿Cuáles de las siguientes etiquetas son correctas y cuáles no?

```
<ejemplo></ejemplo>
<ejemplo />
<ejemplo>Esto es un ejemplo</ejemplo>
<ejemplo><a></a></ejemplo>
```

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo EMPTY>
]>
```



```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo (#PCDATA)>
]>
```

```
<ejemplo />
<ejemplo>Esto es un ejemplo</ejemplo>
<ejemplo><a></a></ejemplo>
```

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo (#PCDATA)>
]>
```

¿Cuál de los siguientes elementos son correctos y cuáles no?



<ejemplo />



<ejemplo>Esto es un ejemplo</ejemplo>



<ejemplo><a></ejemplo>

<!-- ERROR: contiene un elemento <a> -->

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo ANY>
    <!ELEMENT a ANY>
]>
```

```
<ejemplo />
<ejemplo>Esto es un ejemplo</ejemplo>
<ejemplo>Esto es <a>un ejemplo</a></ejemplo>
```

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo ANY |
    <!ELEMENT a ANY |
]>
```

¿Cuál de los siguientes elementos son correctos y cuáles no?



<ejemplo />



<ejemplo>Esto es un ejemplo</ejemplo>



<ejemplo>Esto es <a>un ejemplo</ejemplo>

Declaración de elementos

- Para indicar que un elemento puede o debe contener otros elementos se deben indicar los elementos, utilizando los conectores y modificadores siguientes:
 - , (coma): significa que el elemento contiene los elementos en el orden indicado.

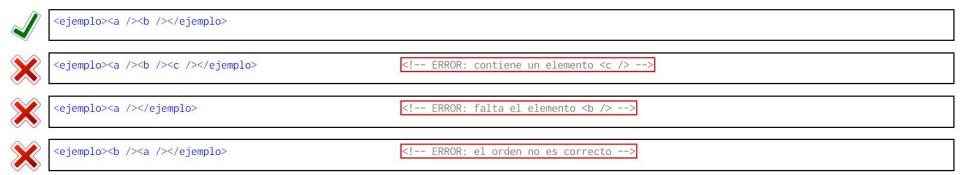
- | (o lógico): significa que el elemento contiene uno de los dos elementos.

- ? (opcional): significa que el elemento puede aparecer o no, pero sólo una vez.
- *: significa que el elemento puede aparecer o no una o más veces.
- **+ (obligatorio):** significa que el elemento tiene que aparecer una o más veces (no puede no aparecer).
- (): permite agrupar expresiones.

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo (a, b) >
    <!ELEMENT a EMPTY>
    <!ELEMENT b EMPTY>
]>
```

```
<ejemplo><a /><b /><ejemplo><a /><b /><c /></ejemplo>
<ejemplo><a /><ejemplo><b /><a />
```

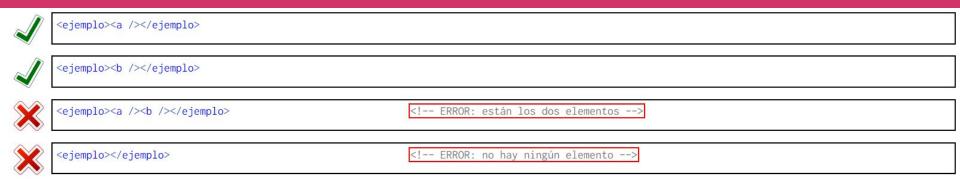
```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo (a, b)>
    <!ELEMENT a EMPTY>
    <!ELEMENT b EMPTY>
]>
```



```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo (a | b)>
    <!ELEMENT a EMPTY>
    <!ELEMENT b EMPTY>
]>
```

```
<ejemplo><a /></ejemplo>
<ejemplo><b /></ejemplo>
<ejemplo><a /><b /></ejemplo>
<ejemplo></ejemplo>
```

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo (a | b)>
    <!ELEMENT a EMPTY>
    <!ELEMENT b EMPTY>
]>
```



```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo (a, b?)>
    <!ELEMENT a EMPTY>
    <!ELEMENT b EMPTY>
]>
```

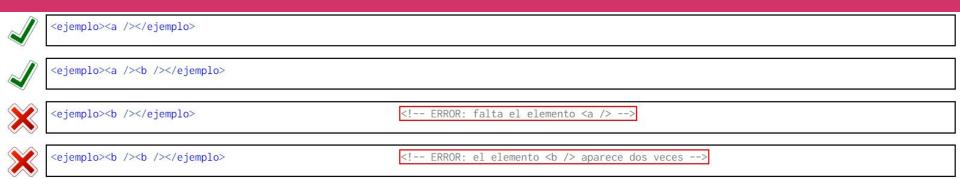
```
<ejemplo><a /></ejemplo>

<ejemplo><a /><b /></ejemplo>

<ejemplo><b /></ejemplo>

<ejemplo><b /></ejemplo>
```

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo (a, b?)>
    <!ELEMENT a EMPTY>
    <!ELEMENT b EMPTY>
]>
```



```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo (a*, b)
    <!ELEMENT a EMPTY>
    <!ELEMENT b EMPTY>
]>
```

```
<ejemplo><b /><ejemplo><b /><ejemplo><b /><ejemplo><b /><ejemplo><a /><b /></ejemplo><b /><ejemplo><b /><ejemplo><b /></ejemplo>
```

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo (a*, b)
    <!ELEMENT a EMPTY>
    <!ELEMENT b EMPTY>
]>
```

¿Cuál de los siguientes elementos son correctos y cuáles no?



<ejemplo></ejemplo>



<ejemplo><a /></ejemplo>



<ejemplo><a /><a /></ejemplo>



<ejemplo><a /></ejemplo>

<!-- ERROR: el elemento <a /> aparece después de -->

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo (a+, b)
    <!ELEMENT a EMPTY>
    <!ELEMENT b EMPTY>
]>
```

```
<ejemplo><a /><b /><ejemplo><a /><ejemplo><a /><a /><b /></ejemplo><a <ejemplo><b /><ejemplo><b /></ejemplo>
```

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo (a+, b)>
    <!ELEMENT a EMPTY>
    <!ELEMENT b EMPTY>
]>
```

¿Cuál de los siguientes elementos son correctos y cuáles no?



<ejemplo><a /></ejemplo>



<ejemplo><a /><a /></ejemplo>



<ejemplo></ejemplo>

<!-- ERROR: falta el elemento <a /> -->

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo (a, (a|b))>
    <!ELEMENT a EMPTY>
    <!ELEMENT b EMPTY>
]>
```

```
<ejemplo><a /><a /></ejemplo>
<ejemplo><a /><b /></ejemplo>
<ejemplo><a /></ejemplo>
```

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo (a, (a|b))
    <!ELEMENT a EMPTY>
    <!ELEMENT b EMPTY>
]>
```

¿Cuál de los siguientes elementos son correctos y cuáles no?



<ejemplo><a /><a /></ejemplo>



<ejemplo><a /></ejemplo>



<ejemplo><a /></ejemplo>

<!-- ERROR: falta el elemento <a /> o -->

<u>Ejemplo</u>

Unos programadores necesitan un formato de fichero para que sus distintos programas intercambien información sobre ventas. El acuerdo al que han llegado es que su XML debería tener esta estructura:

- El elemento raíz será < listaventas>
- Toda < listaventas > tiene una o más ventas.
- Toda <venta> tiene los siguientes datos:
- Importe.
- Comprador.
- Vendedor.
- Fecha (optativa).
- Un código de factura.

Posible solución:

```
<!DOCTYPE listaventas[</pre>
  <!ELEMENT listaventas (venta+)>
  <!ELEMENT venta (importe, comprador,
vendedor, fecha?, codigofactura) >
  <!ELEMENT importe (#PCDATA)>
  <!ELEMENT comprador (#PCDATA)>
  <!ELEMENT vendedor (#PCDATA)>
  <!ELEMENT fecha (#PCDATA)>
  <!ELEMENT codigofactura (#PCDATA)>
```

Classroom

Declaración de atributos

Los atributos permitidos para un elemento se especifican con ATTLIST y sigue la siguiente sintaxis:

<!ATTLIST nombreElemento nombreAtributo tipoAtributo valorInicialAtributo >

Donde:

- **nombreElemento:** es el nombre del elemento para el que se define un atributo.
- nombreAtributo: es el nombre del atributo.
- **tipoAtributo:** es el tipo de datos.
- valorInicialAtributo: es el valor predeterminado del atributo (aunque también puede indicar otras cosas).

Declaración de atributos

ilmportante! Para definir varios atributos de un mismo elemento, se puede utilizar una o varias declaraciones de atributos.

Declaración de atributos

Los tipos de definiciones de atributos que podemos encontrarnos son:

- CDATA: el atributo contiene caracteres (sin restricciones).
- NMTOKEN: el atributo sólo contiene letras, dígitos, y los caracteres ".", "-", "_" y ":".
- **NMTOKENS:** el atributo sólo contiene letras, dígitos, y los caracteres ".", "-", "_", ":" (como el tipo NMTOKEN) y también espacios en blanco.
- valores concretos: el atributo sólo puede contener uno de los términos de una lista. La lista se escribe entre paréntesis, con los términos separados por una barra vertical "|".
- **ID:** el valor del atributo (no el nombre) debe ser único y no se puede repetir en otros elementos o atributos.

Declaración de atributos

Los tipos de definiciones de atributos que podemos encontrarnos son:

- **IDREF:** el valor del atributo debe coincidir con el valor del atributo ID de otro elemento.
- **IDREFS:** el valor del atributo es una serie de valores separados por espacios que coinciden con el valor del atributo ID de otros elementos.
- **ENTITY:** el valor del atributo es alguna entidad definida en la DTD.
- **ENTITIES:** el valor del atributo es alguna de las entidades de una lista de entidades definida en la DTD.
- NOTATION: el valor del atributo es alguna notación definida en la DTD.

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo EMPTY>
    <!ATTLIST ejemplo color CDATA #REQUIRED>
]>
```

```
<ejemplo color="" />
<ejemplo color="amarillo" />
<ejemplo color="azul marino #000080" /
<ejemplo />
<ejemplo />
<ejemplo sabor="dulce" />
```

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo EMPTY>
    <!ATTLIST ejemplo color CDATA #REQUIRED>
]>
```

¿Cuál de los siguientes elementos son correctos y cuáles no?



<ejemplo color="" />



<ejemplo color="amarillo" />



<ejemplo color="azul marino #000080" />



<ejemplo />
<!-- ERROR: falta el atributo "color", obligatorio debido al #REQUIRED -->



```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo EMPTY>
    <!ATTLIST ejemplo color NMTOKEN #REQUIRED>
]>
```

```
<ejemplo color="" />
<ejemplo color="azul-marino" />
<ejemplo color="1" />
<ejemplo color="azul marino" />
<ejemplo color="#F0F0F0" />
```

```
<!DOCTYPE ejemplo [
 <!ELEMENT ejemplo EMPTY>
 <!ATTLIST ejemplo color NMTOKEN
```

¿Cuál de los siguientes elementos son correctos y cuáles no?



<eiemplo color="" />



<ejemplo color="azul-marino" />



<eiemplo color="1" />



<ejemplo color="azul marino" /> <!-- ERROR: hay un espacio en blanco



<eiemplo color="#F0F0F0" /> <!-- ERROR: contiene el carácter # -->



```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo EMPTY>
    <!ATTLIST ejemplo color NMTOKENS #REQUIRED>
]>
```

```
<ejemplo color="" />
<ejemplo color="1" />
<ejemplo color="azul marino" />
<ejemplo color="2*2" />
```

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo EMPTY>
    <!ATTLIST ejemplo color NMTOKENS #REQUIRED>
]>
```

¿Cuál de los siguientes elementos son correctos y cuáles no?



<ejemplo color="" />



<ejemplo color="1" />



<ejemplo color="azul marino" />



<ejemplo color="2*2" />

<!-- ERROR: hay un asterisco -->

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo EMPTY>
    <!ATTLIST ejemplo color (azul|blanco|rojo) #REQUIRED>
]>
```

```
<ejemplo color="" />
<ejemplo color="azul" />
<ejemplo color="verde" />
```

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo EMPTY>
    <!ATTLIST ejemplo color (azul|blanco|rojo) #REQUIRED>
]>
```

¿Cuál de los siguientes elementos son correctos y cuáles no?



<ejemplo color="" />



<ejemplo color="azul" />



<ejemplo color="verde" />
<!-- ERROR: "verde" no está en la lista de valores -->

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo (libro*)>
    <!ELEMENT libro (#PCDATA) >
    <!ATTLIST libro codigo ID #REQUIRED>
]>
```

```
<ejemplo>
    libro codigo="L1">Poema de Gilgamesh</libro>
    libro codigo="L2">Los preceptos de Ptah-Hotep</libro>
</ejemplo>
    libro codigo="1">Poema de Gilgamesh</libro>
    libro codigo="L2">Los preceptos de Ptah-Hotep</libro>
</ejemplo>
    libro codigo="L2">Los preceptos de Ptah-Hotep</libro>
</ejemplo>
    libro codigo="L1">Poema de Gilgamesh
</ejemplo>
    libro codigo="L1">Los preceptos de Ptah-Hotep</libro>
</ejemplo></ejemplo>
```

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo (libro*)>
    <!ELEMENT libro (#PCDATA) >
    <!ATTLIST libro codigo ID #REQUIRED>
]>
```



```
<ejemplo>
  libro codigo="L1">Poema de Gilgamesh</libro>
  libro codigo="L2">Los preceptos de Ptah-Hotep</libro>
  </ejemplo>
```



```
<ejemplo>
    clibro codigo="1">Poema de Gilgamesh</libro>
    clibro codigo="L2">Los preceptos de Ptah-Hotep</libro>
</ejemplo>
```



```
<ejemplo>
   clibro codigo="L1">Poema de Gilgamesh</libro>
   libro codigo="L1">Los preceptos de Ptah-Hotep</libro>

</pr
```

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo ((libro|prestamo)*)>
    <!ELEMENT libro (#PCDATA) >
    <!ATTLIST libro codigo ID #REQUIRED>
    <!ELEMENT prestamo (#PCDATA) >
    <!ATTLIST prestamo libro IDREF #REQUIRED>
]>
```

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo ((libro|prestamo)*)>
    <!ELEMENT libro (#PCDATA) >
    <!ATTLIST libro codigo ID #REQUIRED>
    <!ELEMENT prestamo (#PCDATA) >
    <!ATTLIST prestamo libro IDREF #REQUIRED>
]>
```

```
<ejemplo>
  libro codigo="L1">Poema de Gilgamesh</libro>
  <prestamo libro="L1">Numa Nigerio</prestamo>
  </ejemplo>
```



```
<ejemplo>
   clibro codigo="L1">Poema de Gilgamesh</libro>
   <prestamo libro="L2">Numa Nigerio</prestamo>

</pre
```

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo ((libro|prestamo)*)>
    <!ELEMENT libro (#PCDATA) >
    <!ATTLIST libro codigo ID #REQUIRED>
    <!ELEMENT prestamo (#PCDATA) >
    <!ATTLIST prestamo libro IDREFS #REQUIRED>
]>
```

```
<eiemplo>
 libro codigo="L1">Poema de Gilgamesh</libro>
 libro codigo="L2">Los preceptos de Ptah-Hotep</libro>
 color="L1 L2">Numa Nigerio/prestamo>
</eiemplo>
<eiemplo>
 libro codigo="L1">Poema de Gilgamesh</libro>
 libro codigo="L2">Los preceptos de Ptah-Hotep</libro>
 continuo libro="L1">Numa Nigerio/prestamo>
</eiemplo>
<ejemplo>
 libro codigo="L1">Poema de Gilgamesh</libro>
 libro codigo="L2">Los preceptos de Ptah-Hotep</libro>
 </ejemplo>
```

```
<!DOCTYPE ejemplo [
  <!ELEMENT ejemplo ((libro|prestamo)*)>
  <!ELEMENT libro (#PCDATA) >
  <!ATTLIST libro codigo ID #REQUIRED>
  <!ELEMENT prestamo (#PCDATA) >
 <!ATTLIST prestamo libro IDREFS #REQUIRED>
]>
```



```
<ejemplo>
 libro codigo="L1">Poema de Gilgamesh</libro>
 libro codigo="L2">Los preceptos de Ptah-Hotep</libro>
 cprestamo libro="L1 L2">Numa Nigerio</prestamo>
</eiemplo>
```



```
<eiemplo>
 libro codigo="L1">Poema de Gilgamesh</libro>
 libro codigo="L2">Los preceptos de Ptah-Hotep</libro>
 cprestamo libro="L1">Numa Nigerio</prestamo>
</eiemplo>
```



```
<ejemplo>
 libro codigo="L1">Poema de Gilgamesh</libro>
 libro codigo="L2">Los preceptos de Ptah-Hotep</libro>
 ERROR: el valor "L3" no es ID de ningún elemento
</eiemplo>
```

10. Definición de elementos en un DTD

Declaración de atributos

Por otra parte, los atributos pueden tomar valores iniciales, que son los siguientes:

- #REQUIRED: el atributo es obligatorio, aunque no se especifica ningún valor predeterminado.
- **#IMPLIED:** el atributo no es obligatorio y no se especifica ningún valor predeterminado.
- **#FIXED valor:** el atributo tiene un valor fijo.
- valor concreto: el atributo tiene un valor predeterminado.

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo EMPTY>
    <!ATTLIST ejemplo color CDATA #REQUIRED>
]>
```

```
<ejemplo color="" />
<ejemplo color="amarillo" />
<ejemplo color="azul marino #000080" />
<ejemplo />
```

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo EMPTY>
    <!ATTLIST ejemplo color CDATA #REQUIRED>
]>
```

¿Cuál de los siguientes elementos son correctos y cuáles no?



<ejemplo color="" />



<ejemplo color="amarillo" />



<ejemplo color="azul marino #000080" />



<ejemplo />

<!-- ERROR: falta el atributo "color" -->

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo EMPTY>
    <!ATTLIST ejemplo color CDATA #IMPLIED>
]>
```

```
<ejemplo />
<ejemplo color="" />
<ejemplo color="amarillo" />
<ejemplo color="azul marino #000080" />
```

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo EMPTY>
    <!ATTLIST ejemplo color CDATA #IMPLIED>
]>
```

```
<
```

```
<ejemplo color="" />
```

```
<ejemplo color="amarillo" />
```

```
<ejemplo color="azul marino #000080" />
```

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo EMPTY>
    <!ATTLIST ejemplo color CDATA #FIXED "verde">
]>
```

```
<ejemplo />
<ejemplo color="verde" />
<ejemplo color="" />
<ejemplo color="amarillo" />
```

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo EMPTY>
    <!ATTLIST ejemplo color CDATA #FIXED "verde">
]>
```

¿Cuál de los siguientes elementos son correctos y cuáles no?



<ejemplo />



<ejemplo color="verde" />

<eiemplo color="amarillo" />



<ejemplo color="" />
<!-- ERROR: el atributo "color" no tiene el valor "verde" --:</pre>



<!-- ERROR: el atributo "color" no tiene el valor "verde" -->

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo EMPTY>
    <!ATTLIST ejemplo color CDATA "verde">
]>
```

```
<ejemplo />
<ejemplo color="" />
<ejemplo color="amarillo" />
<ejemplo color="verde" />
```

```
<!DOCTYPE ejemplo [
    <!ELEMENT ejemplo EMPTY>
    <!ATTLIST ejemplo color CDATA "verde">
]>
```



Classroom