

EJERCICIOS PRÁCTICOS: XML

=====

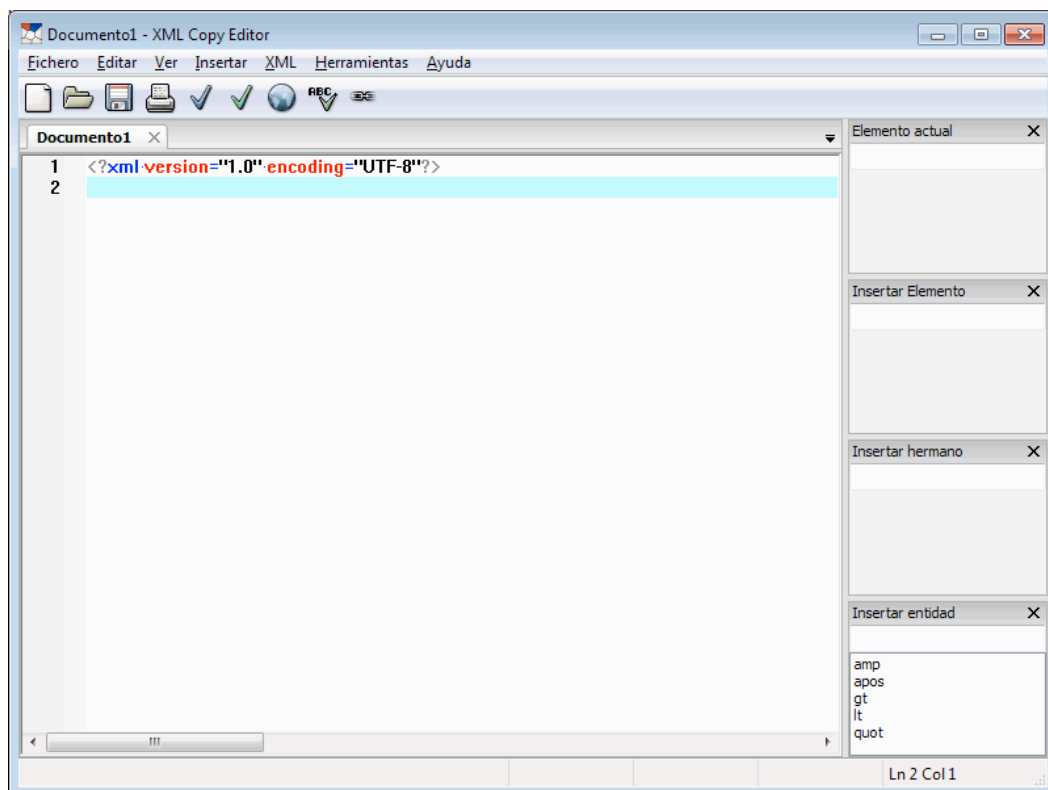
I. XML Copy Editor

XML Copy Editor es un editor de documentos XML libre (GPL 2.0) y multiplataforma cuya página web es <http://xml-copy-editor.sourceforge.net/>

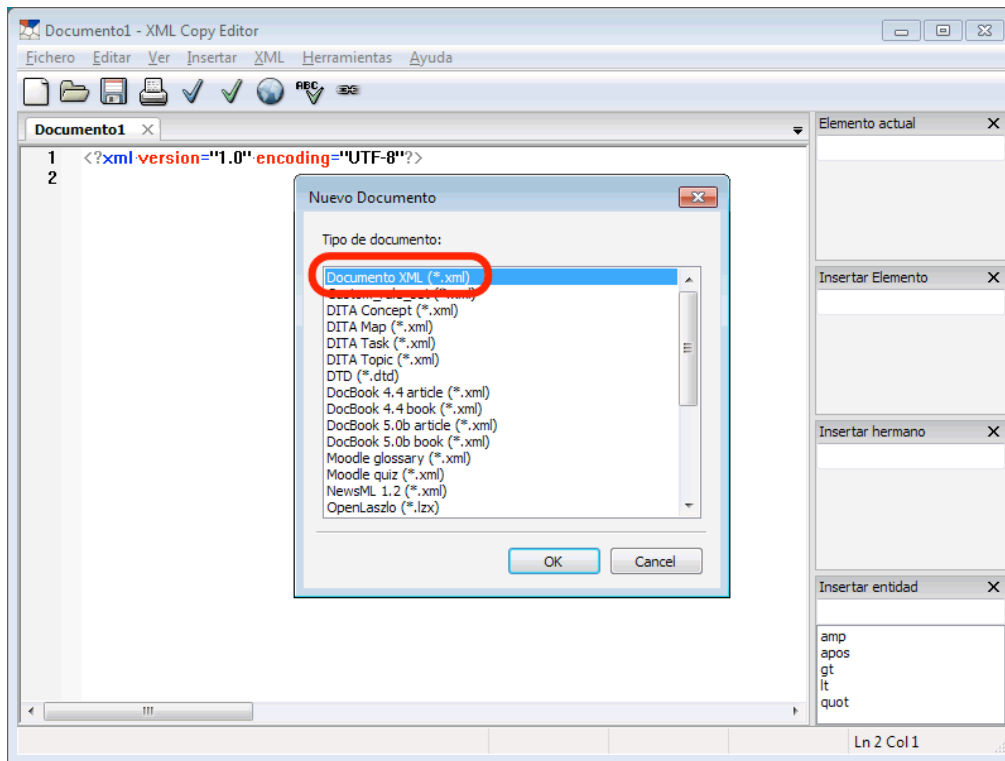
Accede a la página de la herramienta e instálatala en tu equipo.

Interfaz y configuración

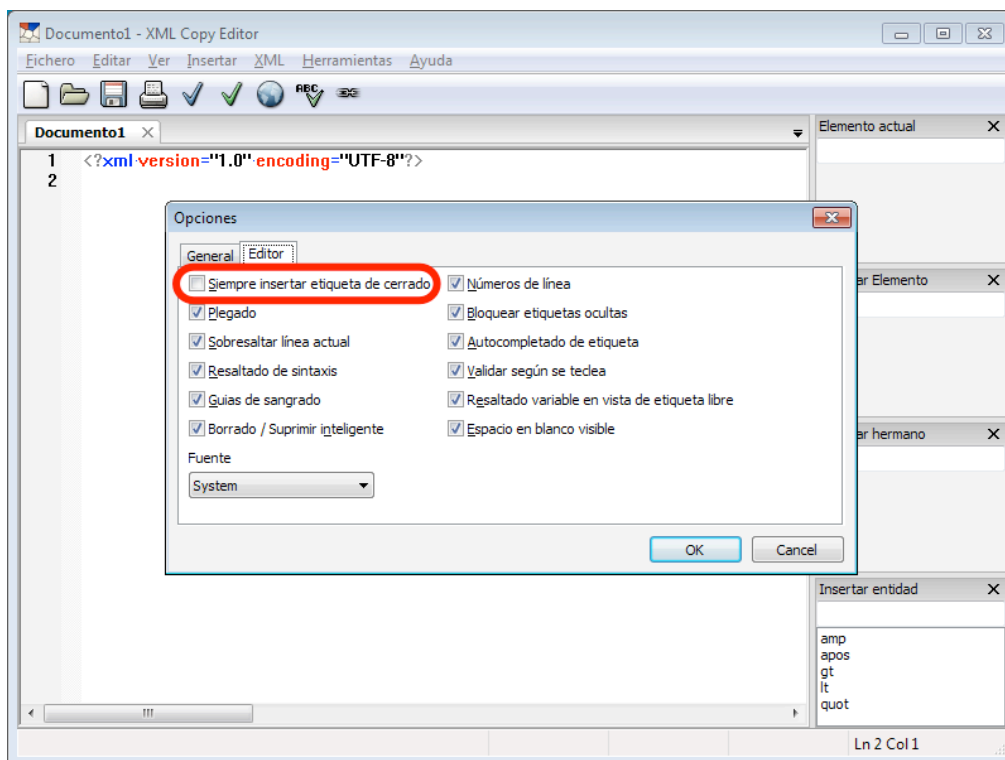
El interfaz de XML Copy Editor es similar al de cualquier editor de texto sin formato:



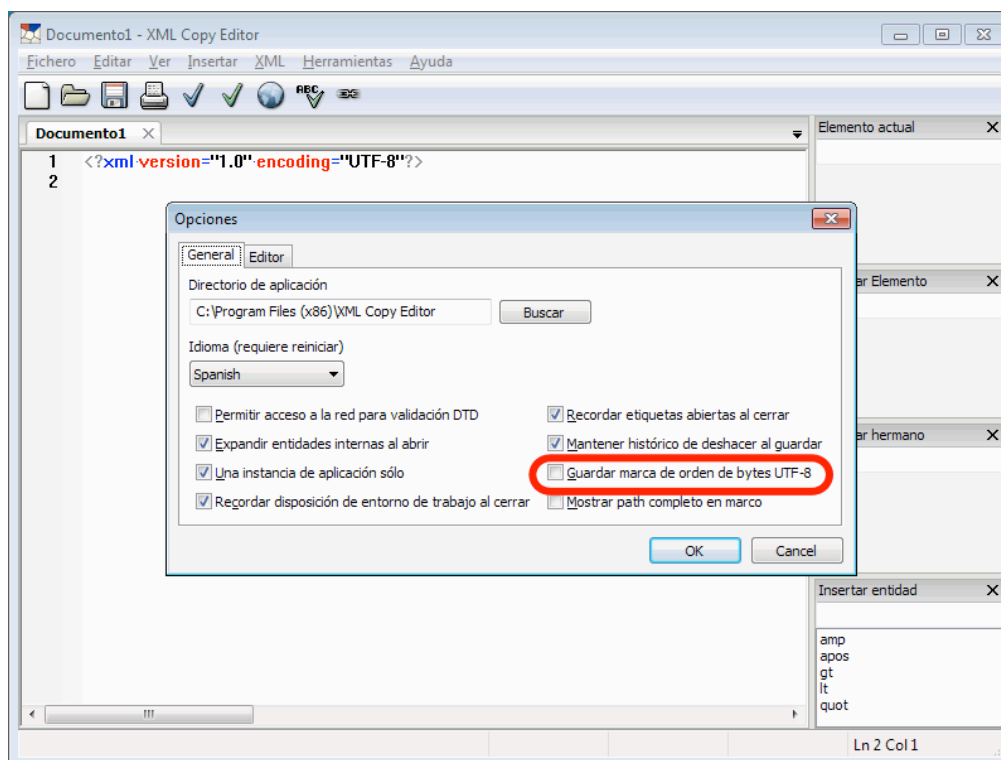
Al abrir XML Copy Editor el editor muestra la plantilla de un documento XML con la declaración XML. Si se quiere crear un documento XML desde cero, se puede elegir el menú *Fichero*→*Nuevo* y elegir la opción Documento XML (*.xml):



El menú *Herramientas*→*Opciones*→*Editor* permite configurar el comportamiento de XML Copy Editor al editar archivos XML. Consejo: desmarcar la casilla "Siempre insertar etiqueta de cerrado".



XML Copy Editor utiliza el navegador predeterminado del sistema. Se aconseja desactivar la opción "Guardar marca de orden de bytes UTF-8", ya que la marca de orden de bytes (BOM) no tiene mucho sentido en UTF-8, puesto que en UTF-8 no existe el problema del orden de los bytes en los caracteres que se codifican con varios caracteres.



Juego de caracteres

La declaración XML indica el juego de caracteres del documento. Dos de los juegos de caracteres más habituales son iso-8859-1 o utf-8, que se indicarían con las siguientes declaraciones xml:

```
<?xml version="1.0" encoding="iso-8859-1"?>
```

```
<?xml version="1.0" encoding="utf-8"?>
```

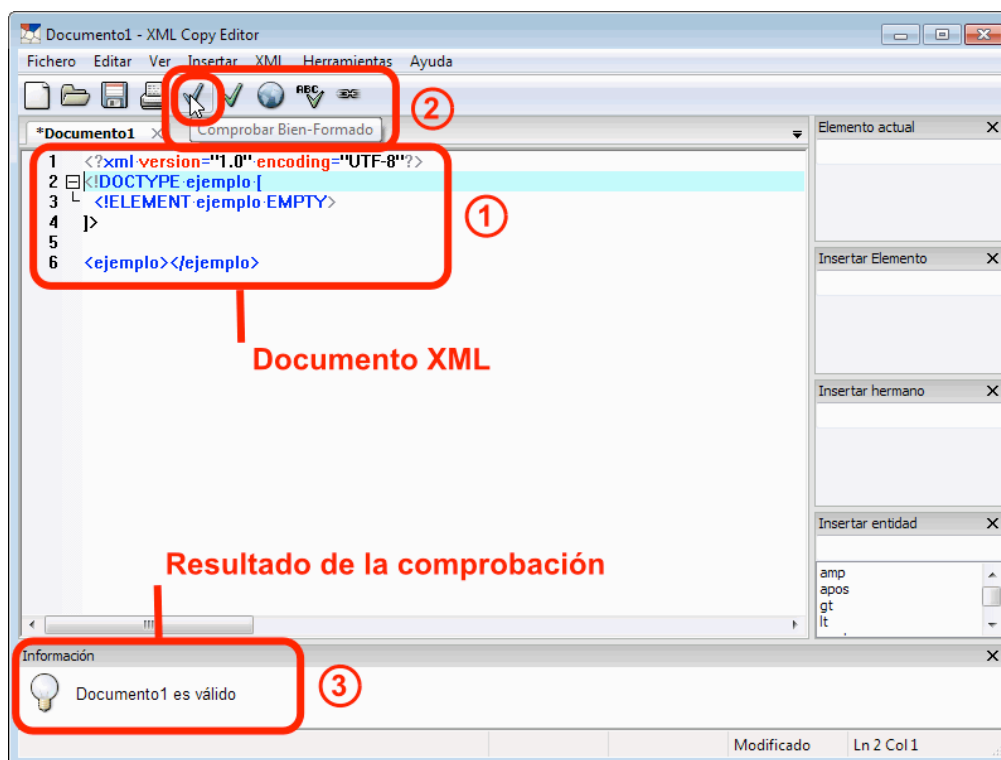
Es importante que el juego de caracteres que aparece en la declaración sea el juego de caracteres en que realmente está guardado el documento, porque si no el procesador XML puede tener problemas leyendo el documento.

XML Copy Editor tiene en cuenta el juego de caracteres indicado en la declaración. Si se modifica la declaración, al guardar el documento se guarda en el juego correspondiente. Pero hay que tener en cuenta que otros editores, como el bloc de notas de Windows, no lo hace.

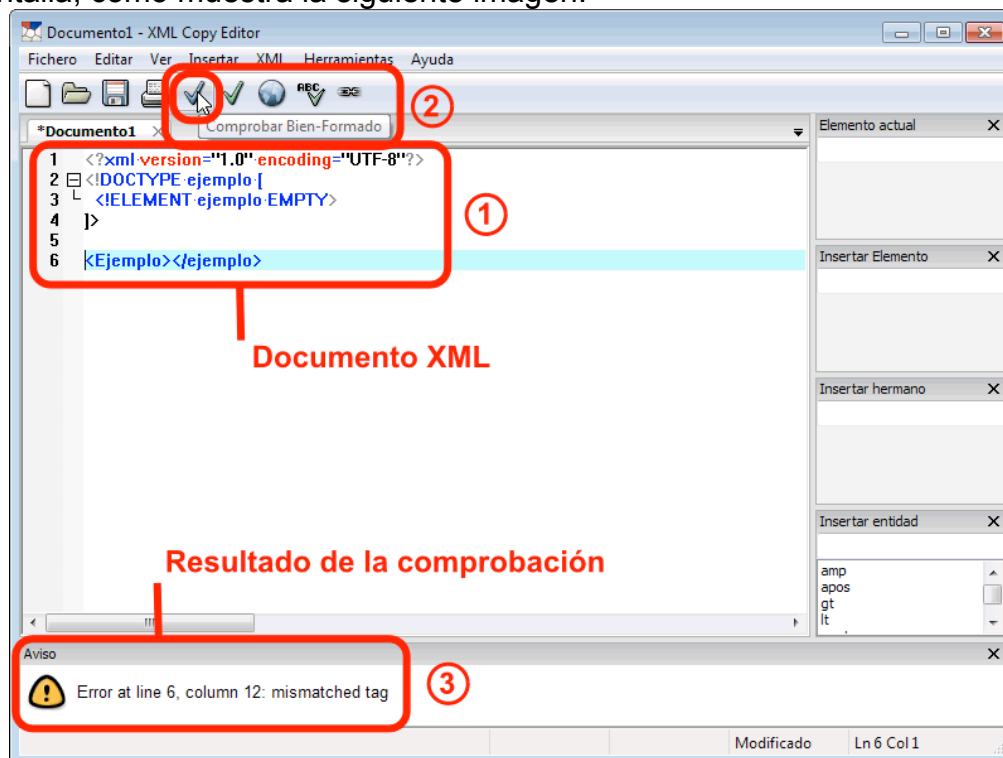
Comprobar que un documento está bien formado

Para comprobar si un documento está bien formado, se puede elegir el menú *XML → Comprobar Bien-Formado*, hacer clic en el botón correspondiente, o pulsar la tecla F2.

En caso de que el documento esté bien formado, el programa lo indica en la parte inferior de la pantalla, como muestra la imagen siguiente:



En caso de que el documento no esté bien formado, el programa lo indica en la parte inferior de la pantalla, como muestra la siguiente imagen:

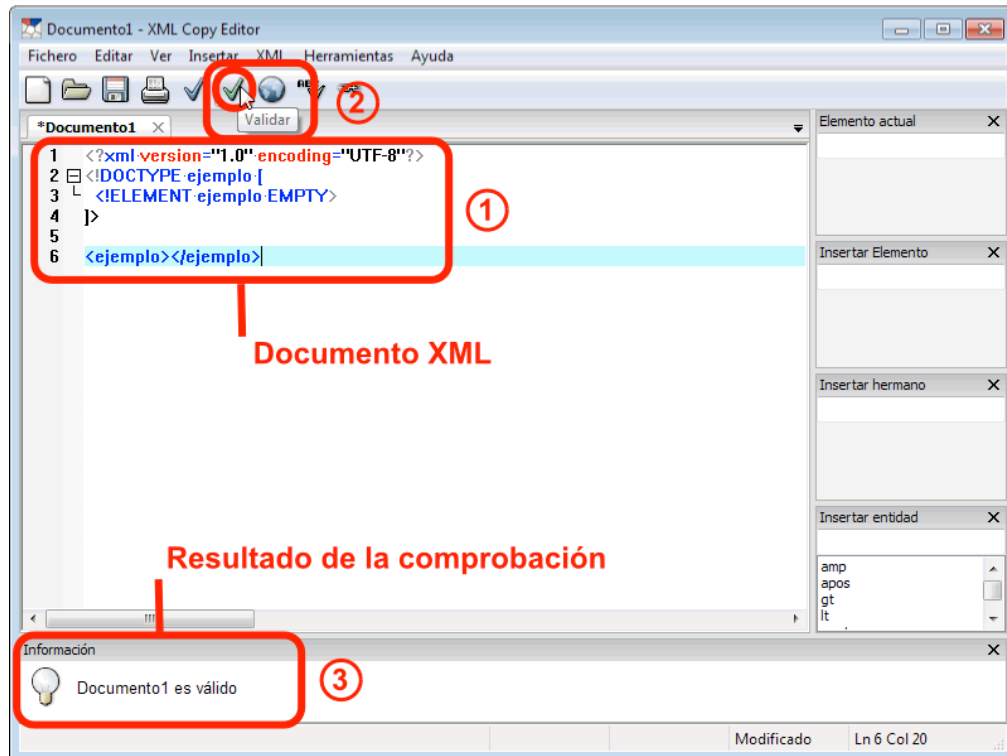


El mensaje inferior explica el tipo de error detectado y la línea en el que se ha detectado (en el ejemplo, que las etiquetas no coinciden). Si el documento contiene varios errores, sólo se muestra uno de ellos, por lo que una vez corregido un error es necesario repetir la comprobación hasta que el documento esté bien formado.

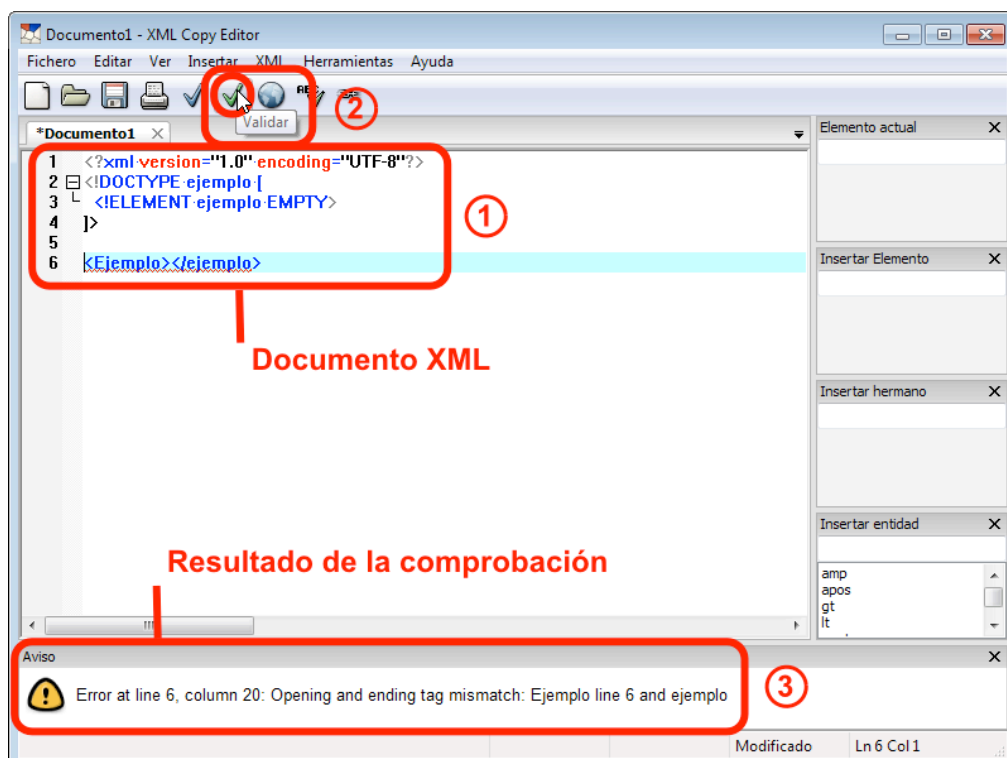
Comprobar que un documento es válido

Para comprobar si un documento es válido, se puede elegir el menú *XML→Validar→DTD/XML Schema*, hacer clic en el botón correspondiente, o pulsar la tecla F5.

En caso de que el documento sea válido, el programa lo indica en la parte inferior de la pantalla:



En caso de que el documento no sea válido, el programa lo indica en la parte inferior de la pantalla:



El mensaje inferior explica el tipo de error detectado y la línea en el que se ha detectado (en el ejemplo, que la etiqueta contiene texto). Si el documento contiene varios errores, sólo se muestra uno de ellos, por lo que una vez corregido un error es necesario repetir la validación hasta que el documento sea válido.

II. Ejercicios

Documentos XML bien formados

Los siguientes documentos no están bien formados. Cada uno contiene dos errores (dos errores del mismo tipo cuentan como uno sólo). Corrige los errores y comprueba con XML Copy Editor que ya son documentos bien formados (si para corregir algún error hay que inventarse una etiqueta o atributo, utiliza un nombre que tenga relación con la información contenida en el documento).

XML deportistas

```
<?xml version="1.0" encoding="UTF-8"?>
<deportistas>
  <deportista>
    <deporte Atletismo />
    <nombre>Jesse Owens</nombre>
  <deportista>
    <deporte Natación />
    <nombre>Mark Spitz</nombre>
  </deportista>
</deportistas>
```

XML películas

```
<?xml version="1.0" encoding="UTF-8"?>
<pelicula>
  <titulo>Con faldas y a lo loco</titulo>
  <director>Billy Wilder</director>
</pelicula>
<pelicula>
  <director>Leo McCarey</director>
  <titulo>Sopa de ganso</titulo>
</pelicula>
<autor />barto</autor>
```

XML texto

```
<?xml version="1.0" encoding="UTF-8"?>
<texto>
  <Titulo>XML explicado a los niños</titulo>
  <párrafo>El <abreviatura>XML</abreviatura>define cómo crear
  lenguajes de marcas.</párrafo>
  <párrafo>Las marcas se añaden a un documento de texto
  para añadir información.</párrafo>
  <http://>www.example.org</http://>
</texto>
```

XML información geográfica

```
<?xml version="1.0" encoding="UTF-8"?>
<geografia mundial>
  <pais>
    <pais>España</pais>
    <continente>Europa</continente>
    <capital></capital nombre="Madrid">
  </pais>
</geografia mundial>
```

Elementos sin atributos

Los siguientes documentos no son válidos porque contienen uno o dos errores (los errores **no** están en el DTD interno). Corrige los errores y compruebe con XML Copy Editor que ya son documentos válidos.

Números

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE numeros [
  <!ELEMENT numeros (#PCDATA)>
]>

<numeros>
  <numero>25</numero>
</numeros>
```

Letras

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE letras [
  <!ELEMENT letras (letra)>
  <!ELEMENT letra (#PCDATA)>
]>

<letras>
  <letra>m</letra>
  <letra>uve doble</letra>
</letras>
```

Colores

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE colores [
  <!ELEMENT colores (color*)>
  <!ELEMENT color (#PCDATA)>
]>

<colores>
  <color>azul marino</color>
  negro
  <color>amarillo</color>
</colores>
```

Flores

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE flores [
  <!ELEMENT flores (flor+)>
  <!ELEMENT flor (#PCDATA)>
]>
<flores>
</flores>
```

Animales

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE animales [
  <!ELEMENT animales (animal*)>
  <!ELEMENT animal (#PCDATA)>
]>

<animales>
  <perro>Caniche</perro>
  <gato>Siamés</gato>
</animales>
```

Escritores

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE escritores [
  <!ELEMENT escritores (escritor*)>
  <!ELEMENT escritor (nombre, nacimiento)>
  <!ELEMENT nombre (#PCDATA)>
  <!ELEMENT nacimiento (#PCDATA)>
]>

<escritores>
  <escritor>
    <nombre>Mario Vargas Llosa</nombre>
    <nacimiento>28 de marzo de 1936</nacimiento>
  </escritor>
  <escritor>
    <nacimiento>1 de abril de 1929</nacimiento>
    <nombre>Milan Kundera</nombre>
  </escritor>
</escritores>
```


Los siguientes documentos no son válidos porque contienen errores (los errores **están** en el DTD interno). Corrige los errores y comprueba con XML Copy Editor que ya son documentos válidos.

NOTA: *Casi siempre habrá varias maneras de corregir la DTD, intenta elegir la más sencilla.*

Marcadores

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE favoritos [
  <!ELEMENT favoritos (marcador)>
  <!ELEMENT marcador (nombre, uri)>
  <!ELEMENT nombre (#PCDATA)>
  <!ELEMENT uri (#PCDATA)>
]>

<marcadores>
  <marcador>
    <nombre>W3C</nombre>
    <uri>http://www.w3.org/</uri>
  </marcador>
  <marcador>
    <nombre>Web Hypertext Application Technology Working Group
(WHATWG)</nombre>
    <uri>http://www.whatwg.org/</uri>
  </marcador>
</marcadores>
```

Efemerides

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE efemerides [
  <!ELEMENT efemerides (efemeride)>
  <!ELEMENT efemeride (fecha, hecho)>
]>

<efemerides>
  <efemeride>
    <fecha>20 de julio de 1969</fecha>
    <hecho>Llegada del hombre a la Luna</hecho>
  </efemeride>
  <efemeride>
    <fecha>12 de octubre de 1492</fecha>
    <hecho>Llegada de Colón a América</hecho>
  </efemeride>
  <efemeride>
    <fecha>6 de abril de 1909</fecha>
    <hecho>llegada de Robert Peary al Polo Norte</hecho>
  </efemeride>
</efemerides>
```

Aeropuertos

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE aeropuertos [
  <!ELEMENT aeropuertos (aeropuerto*)>
  <!ELEMENT aeropuerto (nombre, cerrado)>
  <!ELEMENT nombre (#PCDATA)>
  <!ELEMENT cerrado (#PCDATA)>
]>

<aeropuertos>
  <aeropuerto>
    <nombre>Berlín Schönefeld (SFX)</nombre>
  </aeropuerto>
  <aeropuerto>
    <nombre>Berlín Tempelhof (THF)</nombre>
    <cerrado />
  </aeropuerto>
</aeropuertos>
```

Vuelos

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE vuelos [
  <!ELEMENT vuelos (vuelo*)>
  <!ELEMENT vuelo (origen, destino)>
  <!ELEMENT origen (#PCDATA)>
  <!ELEMENT destino (#PCDATA)>
]>

<vuelos>
  <vuelo>
    <origen>Valencia (VLC)</origen>
    <destino>Londres Heathrow (LHR)</destino>
  </vuelo>
  <vuelo>
    <destino>Berlín Schönefeld (SFX)</destino>
    <origen>Paris Charles de Gaulle (CDG)</origen>
  </vuelo>
</vuelos>
```

Reyes españoles

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE reyesEspañoles [
  <!ELEMENT reyesEspañoles (rey*, reina*)>
  <!ELEMENT rey (nombre, padre, madre)>
  <!ELEMENT nombre (#PCDATA)>
  <!ELEMENT padre (#PCDATA)>
  <!ELEMENT madre (#PCDATA)>
]>

<reyesEspañoles>
  <rey>
    <nombre>Felipe III</nombre>
    <padre>Felipe II</padre>
    <madre>Ana de Austria</madre>
```

```

</rey>
<reina>
  <nombre>Juana la Loca</nombre>
  <padre>Fernando el Católico</padre>
  <madre>Isabel la Católica</madre>
</reina>
<rey>
  <nombre>Carlos I</nombre>
  <padre>Felipe el Hermoso</padre>
  <madre>Juana la Loca</madre>
</rey>
</reyesEspañoles>

```

Elementos con atributos

Los siguientes documentos no son válidos porque contienen uno o dos errores (los errores **no** están en el DTD interno). Corrige los errores y comprueba con XML Copy Editor que ya son documentos válidos.

Datos personales

```

<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE persona [
  <!ELEMENT persona EMPTY>
  <!ATTLIST persona nombre CDATA #IMPLIED>
]>

<persona dni="03141592E" />

```

Película

```

<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE pelicula [
  <!ELEMENT pelicula EMPTY>
  <!ATTLIST pelicula titulo CDATA #IMPLIED>
]>

<pelicula titulo="La diligencia" genero="oeste" />

```

Cuadros

```

<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE cuadros [
  <!ELEMENT cuadros (cuadro*)>
  <!ELEMENT cuadro EMPTY>
  <!ATTLIST cuadro titulo ID #REQUIRED>
  <!ATTLIST cuadro autor CDATA #REQUIRED>
]>

<cuadros>
  <cuadro titulo="Adán y Eva" autor="Alberto Durero" />
  <cuadro autor="Lucas Cranach, el viejo" titulo="Adán y Eva" />
</cuadros>

```

Lista de la compra

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE listaCompra [
  <!ELEMENT listaCompra (item*)>
  <!ELEMENT item EMPTY>
  <!ATTLIST item nombre CDATA #REQUIRED>
  <!ATTLIST item cantidad CDATA #REQUIRED>
]>

<listaCompra>
  <leche cantidad="12 litros" ></leche>
  <pan cantidad="3 barras de cuarto" />
</listaCompra>
```

Los siguientes documentos no son válidos porque contienen errores (los errores **están** en el DTD interno). Corrige los errores y comprueba con XML Copy Editor que ya son documentos válidos.

Libro

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE libro [
  <!ELEMENT libro EMPTY>
  <!ATTLIST libro autor NMTOKEN #REQUIRED>
]>

<libro autor="Mario Vargas Llosa" />
```

Inventores

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE inventores [
  <!ELEMENT inventores>
  <!ELEMENT inventor EMPTY>
  <!ATTLIST inventor invento CDATA #REQUIRED>
  <!ATTLIST inventor nombre ID #REQUIRED>
]>

<inventores>
  <inventor nombre="Robert Adler" invento="Mando a distancia" />
  <inventor nombre="Laszlo Josef Biro" invento="Bolígrafo" />
  <inventor nombre="Josephine Garis Cochran" invento="Lavaplatos" />
  <inventor invento="Fuego" />
</inventores>
```

Cosas por hacer

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE cosasPorHacer [
  <!ELEMENT cosasPorHacer (cosa)>
  <!ELEMENT cosa EMPTY>
  <!ATTLIST cosa fecha CDATA #REQUIRED>
  <!ATTLIST cosa asunto CDATA #REQUIRED>
  <!ATTLIST cosa fechaLimite CDATA #REQUIRED>
]>

<cosasPorHacer>
  <cosa fecha="20 de febrero de 2011" fechaLimite="1 de marzo de 2011">
    Preparar ejercicios de DTDs</cosa>
  <cosa fecha="21 de febrero de 2011" fechaLimite="5 de marzo de 2011">
    Preparar tema XSLT</cosa>
</cosasPorHacer>
```