XML

eXtensible Markup Language

Necessitat del XML

- Les aplicacions informàtiques fan servir una gran quantitat de dades, que han de ser consultades, intercanviades, manipulades etc.
- Cada aplicació desa les dades en un format propietari, de tal manera que per poder fer-los servir has de tenir el mateix programa.
- XML: **és una forma estándar d'intercanviar dades**. Per exemple Word (doc a docx), Excel (xls a xlsx) ja fan servir XML per desar les dades.

Característiques

- Document de text pla (editors bàsics)
- Llenguatge de marques (estructura dades)
- Dissenyat pel transport de dades, no per mostrar-los (format amb CSS).
- Les etiquetes XML no definides, es defineixen per a cada aplicació.
- XML és autodescriptiu.
- Desenvolupat per la W₃C.
- Eina independent del software i del hardware (per desar les dades en text pla).
- Tecnologies adicionals que complementen XML: XLST, Xpath ,Xquery ,Xforms...

Metallenguatge

- XML:= és una forma de crear llenguatges.
- Alguns llenguatges basats en XML:
 - GML (Geography Markup Language, Lenguaje de Marcado Geográfico).
 - MathML (Mathematical Markup Language, Lenguaje de Marcado Matemático).
 - RSS (Really Simple Syndication, Sindicación Realmente Simple).
 - SVG (Scalable Vector Graphics, Gráficos Vectoriales Escalables).
 - XHTML (eXtensible HyperText Markup Language, Lenguaje de Marcado de Hipertexto eXtensible).

És només informació entorn etiquetes

XML vs HMTL

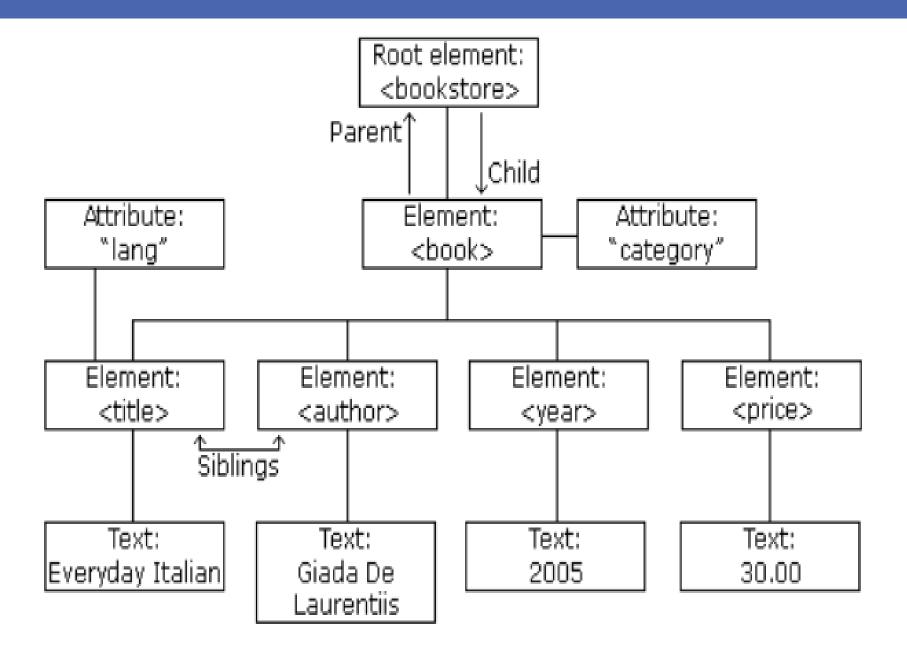
- XML va ser dissenyat per transportar dades, lo important són les dades .
- HTML va ser dissenyat per mostrar dades. Lo important és com es mostra les dades(estructura).
- XML no té les etiquetes predefinides, la persona autora és qui les defineix.
- HMTL sí té les etiquetes predefinides.
- XML és extensible.

Arbre XML

```
<?xml version="1.0" encoding="UTF-8"?>
<bookstore>
  <book category="cooking">
    <title lang="en">Everyday Italian</title>
    <author>Giada De Laurentiis</author>
    <year>2005</year>
    <price>30.00</price>
  </book>
  <book category="children">
    <title lang="en">Harry Potter</title>
    <author>J K. Rowling</author>
    <year>2005</year>
    <price>29.99</price>
  </book>
</bookstore>
```

Pares, fills i germans

arbreXml.xml



Objectiu: arxius XML ben estructurats

• Per desar un valor:

```
<etiqueta>valor</etiqueta>
```

• Pot tenir un element buit:

```
<year>
<element></element>
<element />
```

- Els docs XML han de tenir un element arrel (<bookstore>).
- Declaració del XML:

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

- ✓ És opcional:
- ✓ Si es posa va a l'inici del document.
- ✓ S'especifica la versió XML emprada (abans de l'encoding).
- ✓ S'ha d'especificar la codificació emprada, la codificació de caracters predetreminada per

documents XML és la UTF-8 (igual que a HTML5, CSS, JS, PHP i SQL).

Relació pare-fills

- L'element persona té quatre fills.
- I l'element fecha-de-nacimiento tres.

• Correctament anidats : <i>This text is bold and italic</i>

En XML els elements també poden tenir atributs amb valors com l'HTML
 Aquests han d'estar citats.

- Tots els elements XML han d'estar tancats (els tags).
- Elements amb continguts mixte:

```
<persona>
  <nombre>Elsa</nombre> vive en <ciudad>Pamplona</ciudad>.
</persona>
```

• XML és case sensitive distingeix entre majúscules i minúscules.

<carta> no és lo mateix que <Carta>.

<Price>29.99</price> incorrecte
<Price>29,99</Price> correcte

• Espais en blanc:

</peliculas>

• Espais en blanc:

datos.xml

```
▼<series>
     <serie numeros="2 4 6 8"/>
     <serie numeros="3 6 9 12 15"/>
   </series>
<?xml version="1.0" encoding="UTF-8"?>
<datos>
  <dato>1</dato>
  <dato>2</dato>
  <dato>3</dato>
</datos>
<?xml version="1.0" encoding="UTF-8"?>
<datos><dato>1</dato><dato>2</dato><dato>3</dato></dato>
<?xml version="1.0" encoding="UTF-8"?>
<dato>><dato>2</dato>
<dato>3</dato></datos>
```

Atribut predefinit xml:space amb el valor "preserve"

```
1 Fernando Alonso 1:55.341
2 Lewis Hamilton 1:55.729
3 Sebastian Vettel 1:56.122
</clasificacion>
```

preserve.xml

Per indicar que els espais que apareixen en el contingut de l'element han de preservar-se.

No tots els programes el reconeixen.

Valors: preserve o default (el que es té per defecte).

• Referències a entitats, hi han cinc per XML:

<	<	less than	
>	>	greater than	
&	&	ampersand	
'	1	apostrophe	_
"	ıı	quotation mark	<

```
<year> any <2005</year>
<year> any &lt;2005</year>
```

entidades.xml

Incorrecto

Correcto

<condicion>a<b< condicion=""></b<></condicion>	<condicion>a<b</condicion>
<condicion>a==1 && b==2</condicion>	<pre><condicion>a==1 && b==2</condicion></pre>
<dato)"="" caracter="comilla doble("></dato>	<pre><dato caracter='comilla doble(")'></dato></pre>
<dato)'="" caracter="comilla simple("></dato>	<pre><dato caracter="comilla simple(')"></dato></pre>
	<pre><dato caracter="comilla simple(')"></dato> <dato caracter='comilla doble(")'></dato></pre>

Caràcters Unicode (ex:€)

```
<?xml version="1.0" encoding="UTF-8"?>
oductos>
   <nombre precio="12.56&#8364;">Gorro de lana</no</pre>
                <!-- valor decimal-->
   <nombre precio="16.99&#x20AC;">Gorro polar</nom</pre>
                <!-- valor hexadecimal-->
</productos>
                           <?xml version="1.0" encoding="UTF-8"?>
```

cproductos>

</productos>

<nombre precio="12.56€">Gorro de lana</nombre>
<nombre precio="16.99€">Gorro polar</nombre>

productos.xml

Elements-Nomenclatura

Els noms:

- Han de començar amb una lletra o un guió baix
- No poden començar amb les lletres XML. Es poden fer servir totes les paraules (excepte xml)
- Sí poden contenir lletres, dígits, guions, guions baixos i punts
- No poden tenir espais.

Elements-Nomenclatura-Recomanacions

- Noms descriptius: <persona>, <Nom>, <cognom>.
- Noms curts i simples, <titulo_libro> no com aquest: <mi_titulo_de_un_libro>.
- Evitar "-" (alguns programes poden pensar que és una resta)
- Evitar "." (propietat a OOP)
- Evitar ": " (reservat pels espais de noms)
- Evitar les lletres no angleses(á,Á,ñ,Ñ..) encara que sí están permeses, però poden tenir incompatibilitats amb alguns programes que no les reconeixen.
- Es permet un espai en blanc o un salt de línea darrera del nom de l'etiqueta (davant no).

<ciudad >Pamplona</ciudad

```
<Ciudad>Pamplona</Ciudad>
<Ciudad>Pamplona</ciudad>
                                                  <día>18</día>
<día>18</dia>
                                                  <mes>6</mes>
<mes>6<mes/>
                                                  <ciudad>Pamplona</ciudad>
<ciudad>Pamplona</finciudad>
                                                  <_rojo/>
<_rojo>
                                                  <colores2>Rojo y Naranja</colores2>
<2colores>Rojo y Naranja</2colores>
                                                  <Aficiones >Cine, Bailar, Nadar</Aficiones >
< Aficiones >Cine, Bailar, Nadar</ Aficiones >
                                                  <persona><nombre>Elsa</nombre></persona>
<persona><nombre>Elsa</persona></nombre>
                                                  <color_favorito>azul</color_favorito>
<color favorito>azul</color favorito>
```

Incorrecte

Recomanació

Sonia Sánchez Sierra

Estils que es fan servir

Style	Example
Lower case	<firstname></firstname>
Upper case	<firstname></firstname>
Underscore	<first_name></first_name>
Pascal case	<firstname></firstname>
Camel case	<firstname></firstname>

XML extensible

Partim d'una aplicació que extrau les dades d'un document XML com el següent:

Si es fan canvis en el XML l'aplicació pot continuar funcionant igual, ja que continúa trobant els elements pels quals estaba definit.

Atributs

```
<year>2005
```

Pot tenir text i atributs:

```
<book category="children">
  <title lang="en">Harry Potter</title>
  <author>J K. Rowling</author>
```

L'atribut serveix per donar informació extra sobre l'element que el conté.

Per exemple, per desar les dades: codi:47P, nom: falda, color:vermell,preu:14.50

Representació XML:

Atributs Sintaxis

```
<person gender="female">Ok <person gender='female'> Ok

<gangster name='George "Shotgun" Ziegler'>

<gangster name="George &quot;Shotgun&quot; Ziegler"> Ok
```

Tots els atributs d'un element han de ser únics.

Atributs <u>Elements o atributs?</u>

```
<person gender="female">
   <firstname>Anna</firstname>
   <lastname>Smith</lastname>
</person>
```

```
<person>
    <gender>female</gender>
    <firstname>Anna</firstname>
    <lastname>Smith</lastname>
</person>
```

Informació

Atributs

Els atributs no poden contenir múltiples valors, els elements sí.

Els atributs no poden tenir estrutura d'arbre, els elements sí.

Els atributs són difícils d'expandir després si es volen fer canvis.

```
<note day="10" month="01" year="2008"
to="Tove" from="Jani" heading="Reminder"
body="Don't forget me this weekend!">
</note>
```



Atributs

• Els atributs (per ex id) poden servir per identificar elements XML com els id de

l'HTML.

Ens serveix per identificar les diferents notes, és a dir és una dada de les dades (metadades), no són dades per si mateixes.

Recomanació: Les metadades s'han de desar com atributs i les dades en si com a elements.

Sonia Sánchez Sierra

Espai de noms

• Si volem fer servir diferents XML pot passar:

```
Conflicte de noms
```

```
<name>African Coffee Table</name>
<width>80</width>
<length>120</length>
```

Espai de noms

- És una recomanació W3C per proporcionar elements i atributs amb nom únic en un arxiu XML.
- Un arxiu XML pot contenir noms d'elements o atributs amb origen de més d'un vocabulari XML.
- Si a cada vocabulari se li dóna un espai de noms, un espai propi es resól el conflicte de noms que pot provocar els elements o atributs amb el mateix nom.
- És defineix fent referència a una URI.
- No es tracta de fer servir la URI com un enllaç, ni ha de tenir contingut, només es fa servir perquè el nom sigui únic.

Espai de noms - sintaxi

```
xmlns:prefijo="URI"
```

Per exemple: aquí es defineixen dos espais de noms amb els prefixes e1 i e2.

```
<e1:ejemplo xmlns:e1="http://www.abrirllave.com/ejemplo1"
xmlns:e2="http://www.abrirllave.com/ejemplo2">
```

Les URI no tenen perquè tenir res, només serveixen per ser únics.

Espai de noms - Exemples

• Si volem posar en un XML aquests dos

```
<carta>
     <palo>Corazones</palo>
     <numero>7</numero>
</carta>
```

```
<?xml version="1.0" encoding="UTF-8"?>
<e1:ejemplo xmlns:e1="http://www.abrirllave.com/ejemplo1"
  xmlns:e2="http://www.abrirllave.com/ejemplo2">
  <e1:carta>
      <e1:palo>Corazones</e1:palo>
      <e1:numero>7</e1:numero>
   </el:carta>
   <e2:carta>
      <e2:carnes>
         <e2:filete_de_ternera precio="12.95"/>
         <e2:solomillo_a_la_pimienta precio="13.60"/>
      </e2:carnes>
      <e2:pescados>
         <e2:lenguado al horno precio="16.20"/>
         <e2:merluza_en_salsa_verde precio="15.85"/>
      </e2:pescados>
   </e2:carta>
</e1:ejemplo>
```

Espai de noms - Exemples

Els espais de noms es poden definir en cada element o des de l'arrel com abans.

```
<?xml version="1.0" encoding="UTF-8"?>
<e1:ejemplo xmlns:e1="http://www.abrirllave.com/ejemplo1">
   <e1:carta>
      <e1:palo>Corazones</e1:palo>
      <e1:numero>7</e1:numero>
   </el:carta>
   <e2:carta xmlns:e2="http://www.abrirllave.com/ejemplo2">
      <e2:carnes>
         <e2:filete_de_ternera precio="12.95"/>
         <e2:solomillo a la pimienta precio="13.60"/>
      </e2:carnes>
      <e2:pescados>
         <e2:lenguado_al_horno precio="16.20"/>
         <e2:merluza en salsa verde precio="15.85"/>
      </e2:pescados>
   </e2:carta>
</el:ejemplo>
```

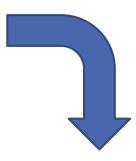
- Definir diferents espais de noms pels elements.
- I cap per algún element (="").

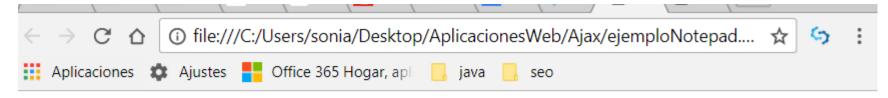
```
<?xml version="1.0" encoding="UTF-8"?>
<ejemplo xmlns="http://www.abrirllave.com/ejemplo1">
   <carta>
      <palo>Corazones</palo>
      <numero>7</numero>
   </carta>
   <carta xmlns="http://www.abrirllave.com/ejemplo2">
      <carnes>
         <filete_de_ternera precio="12.95"/>
         <solomillo_a_la_pimienta precio="13.60"/>
     </carnes>
      <pescados xmlns="">
         <lenguado_al_horno precio="16.20"/>
         <merluza_en_salsa_verde precio="15.85"/>
      </pescados>
   </carta>
</ejemplo>
```

Visualització al Navegador

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

<root>
    <element nombre="primero" />
        <element nombre="segundo" />
-</root>
```



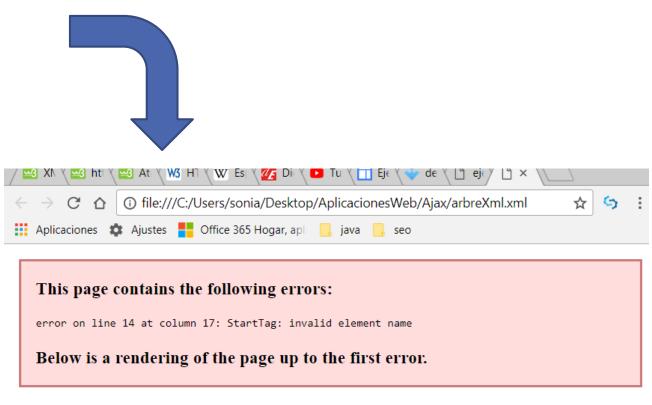


This XML file does not appear to have any style information associated with it. The document tree is shown below.

```
▼<root>
<element nombre="primero"/>
<element nombre="segundo"/>
</root>
```

Visualització al Navegador(errors)

```
<?xml version="1.0" encoding="UTF-8"?>
<bookstore> <!-- Element arrel-->
  <book category="cooking">
         elements secundaris
         title, author, yera, price-->
    <title lang="en">Everyday Italian</title>
    <author>Giada De Laurentiis
    <year>2005</year>
    <price>30.00</price>
  </book>
  <book category="children">
    <title lang="en">Harry Potter</title>
   <author>J K. Rowling</author>
    <year> any <2005</pre>
    <year> any &lt;2005</year>
    <Price>29.99</price> incorrecte
    <Price>29,99</Price> correcte
  </book>
</bookstore>
```



Everyday Italian Giada De Laurentiis 2005 30.00 Harry Potter J K. Rowling any

Exercicis XML ESTRUCTURA

Associar un arxiu CSS a un document XML

```
<?xml-stylesheet type="text/css" href="miestilo.css"?>
```

XML amb CSS (catalog.css i catalog.xml)

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/css" href="catalog.css"?>
<CATALOG>
  <CD>
    <TITLE>Empire Burlesque</TITLE>
    <ARTIST>Bob Dylan
    <COUNTRY>USA</COUNTRY>
    <COMPANY>Columbia</COMPANY>
    <PRICE>10.90</PRICE>
    <YEAR>1985</YEAR>
  </CD>
  <CD>
    <TITLE>Hide your heart</TITLE>
    <ARTIST>Bonnie Tyler
    <COUNTRY>UK</COUNTRY>
    <COMPANY>CBS Records</COMPANY>
    <PRICE>9.90</PRICE>
    <YEAR>1988</YEAR>
  </CD>
</CATALOG>
```

```
CATALOG {
    background-color: #ffffff;
    width: 100%;
CD {
    display: block;
    margin-bottom: 30pt;
    margin-left: 0;
TITLE {
    display: block;
    color: #ff0000;
    font-size: 20pt;
ARTIST {
    display: block;
    color: #0000ff;
    font-size: 20pt;
COUNTRY, PRICE, YEAR, COMPANY {
    display: block;
    color: #000000;
    margin-left: 20pt;
```

Empire Burlesque Bob Dylan USA

Columbia 10.90 1985

Hide your heart Bonnie Tyler

UK CBS Records 9.90 1988

No recomanable XML amb CSS. Millor JS o XSLT.

Sonia Sánchez Sierra

Exercici

- En un XML es desa la informació del nom d'un animal i el número de potes que té.
- Volem una sortida al navegador del tipus:

perro4pato2ballena0

Solució

(animales.css, animales.xml)

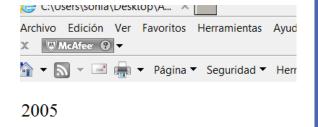
```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/css" href="animales.css"?>
<animales>
   <animal>
      <nombre>perro</nombre>
      <patas>4</patas>
   </animal>
                                           nombre{color:blue;font-size:40px}
   <animal>
                                           patas{color:red;font-size:22px}
      <nombre>pato</nombre>
      <patas>2</patas>
   </animal>
   <animal>
      <nombre>ballena</nombre>
      <patas>0</patas>
    </animal>
</animales>
```

Exercicis XML CSS Errores

Parser (analitzador)

- Les etiquetes del XML fan possible poder tenir una representació interna com el DOM d'HTML.
- Parser:= script,mòdul,biblioteca o programa que s'encarrega de transformar un arxiu de text en una representació interna.
- EL XML DOM (Document Object Model) defineix les propietats i mètodes per accedir i editar XML.
- Per poder accedir al XML s'ha de carregar el DOM XML.
- Els navegadors tenen un parser integrat que pot convertir el text en un objecte DOM XML.

```
<!DOCTYPE html>
<html>
<body>
<script>
var parser, xmlDoc;
var text = "<bookstore><book>" +
"<title>Everyday Italian</title>" +
"<author>Giada De Laurentiis</author>" +
"<year>2005</year>" +
"</book></bookstore>";
parser = new DOMParser();
xmlDoc = parser.parseFromString(text, "text/xml");
document.getElementById("demo").innerHTML =
xmlDoc.getElementsByTagName("year")[0].childNodes[0].nodeValue;
</script>
</body>
</html>
```



Ejemplo de parser 'interno'

analitzador.html

Sonia Sánchez Sierra

- Creem un DOMParser, un analitzador XML DOM : parser= new DomParser()
- Parsear XML des d'una cadena de text, mètode: parseFromString() per obtenir un objecteDOM XML

```
var parser = new DOMParser();
var |xmlDoc = parser.parseFromString(text,"text/xml");
```

```
var text = "<bookstore><book>" +
"<title>Everyday Italian</title>" +
"<author>Giada De Laurentiis</author>" +
"<year>2005</year>" +
"</book></bookstore>";
```

- Hi ha tres resultats diferents possibles, dependent del tipus MIME donat:
 - ✓ application/xml, ojecte resultant serà un Document
 - ✓ image/svg+xml ojecte resultant serà un SVGDocument
 - ✓ text/html ojecte resultant serà un HTMLDocument.

```
var parser = new DOMParser();
var doc = parser.parseFromString(stringContainingXMLSource, "application/xml");
// returns a Document, but not a SVGDocument nor a HTMLDocument

parser = new DOMParser();
doc = parser.parseFromString(stringContainingXMLSource, "image/svg+xml");
// returns a SVGDocument, which also is a Document.

parser = new DOMParser();
doc = parser.parseFromString(stringContainingHTMLSource, "text/html");
// returns a HTMLDocument, which also is a Document.

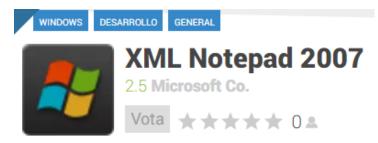
ierra
```

DOM

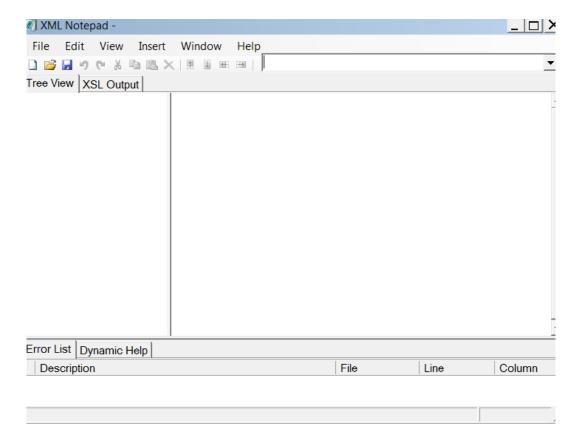
• HTML DOM defineix una forma estàndard per accedir i manipular documents HTML. Presenta un document HTML como una estructura d'arbre.

• XML DOM defineix una forma estàndard per accedir i manipular documents XML. Presenta un document XML como una estructura d'arbre.

XML Notepad

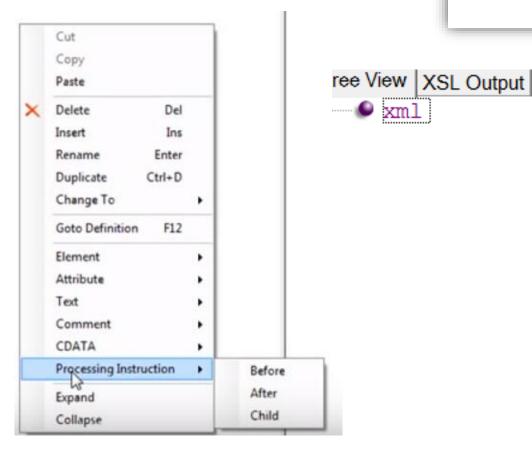


• https://xml-notepad-2007.uptodown.com/windows



Anem a fer el següent arxiu xml:

Definim la versió:



```
<?xml version="1.0" encoding="utf-8"?>
<root>
    <element nombre="primero" />
        <element nombre="segundo" />
</root>
```

version="1.0" encoding="uti-8"

