

Creació de documents XML

XML (eXtensible Markup Language) és un estàndard del W3C per documents.

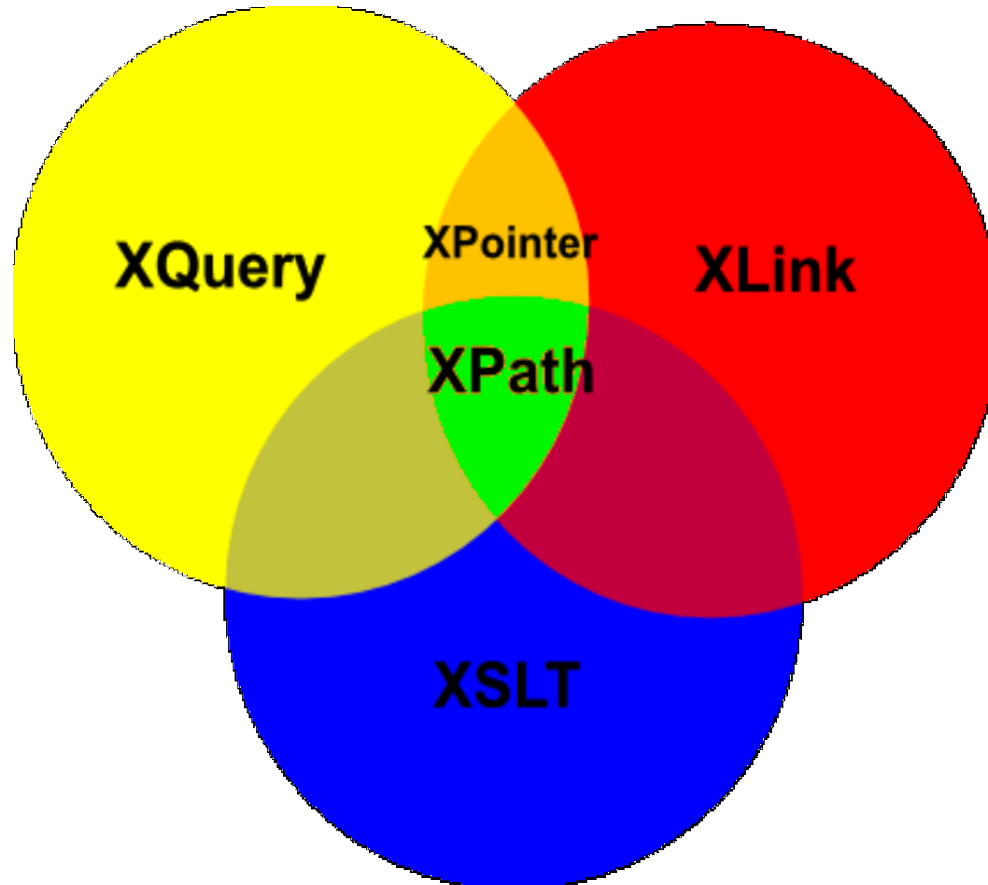
Defineix la sintaxi genèrica per marcar les dades textuais amb valors intel·ligibles per humans.

```
<persona>
  <nompersona>
    <nom>Xavier</nom>
    <cognom>Sala</cognom>
  </nompersona>
  <professio>
    Enginyer Tècnic
  </professio>
  <professio>
    Professor
  </professio>
</persona>
```

XML és un ll. De marques prou **flexible** per definir una **estructura** genèrica d'un document per posteriorment ser personalitzada a diferents dominis : BBDD, WWW, PDF, SMARTPHONE,...

XML no és un ll. De programació ni una BBDD ni un protocol de xarxa

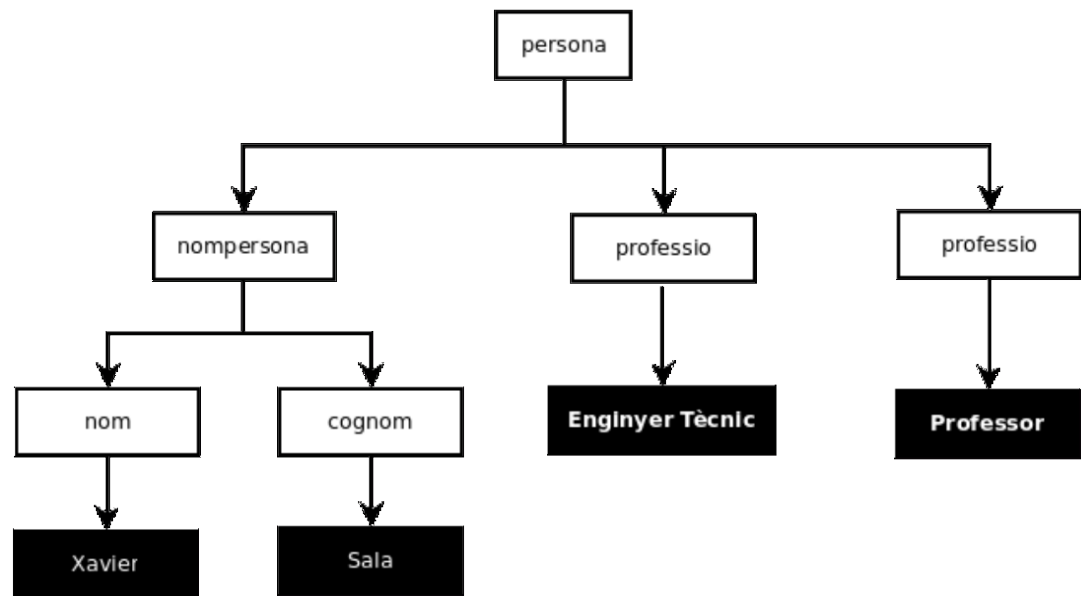
XML Integra diferents technologies



Representació XML en forma d'arbre

- Tots els documents XML **han de tenir** una etiqueta arrel que no és fill de ningú
- Les fulles de l'arbre són les dades contingudes en el document

```
<persona>  
  <nompsona>  
    <nom>Xavier</nom>  
    <cognom>Sala</cognom>  
  </nompsona>  
  <professio>  
    Enginyer Tècnic  
  </professio>  
  <professio>  
    Professor  
  </professio>  
</persona>
```



Exemple: Representar l'arbre

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<montilivi>
  <alumne>
    <nom>Federicu Pi</nom>
    <any>2010</any>
    <matricula>
      <modul>XML</modul>
      <modul>Xarxes Locals</modul>
    </matricula>
  </alumne>
  <nom>Filomeno Garcia</nom>
  <any>2010</any>
  <matricula>
    <modul>XML</modul>
  </matricula>
</alumne>
</montilivi>
```

Document XML ben format

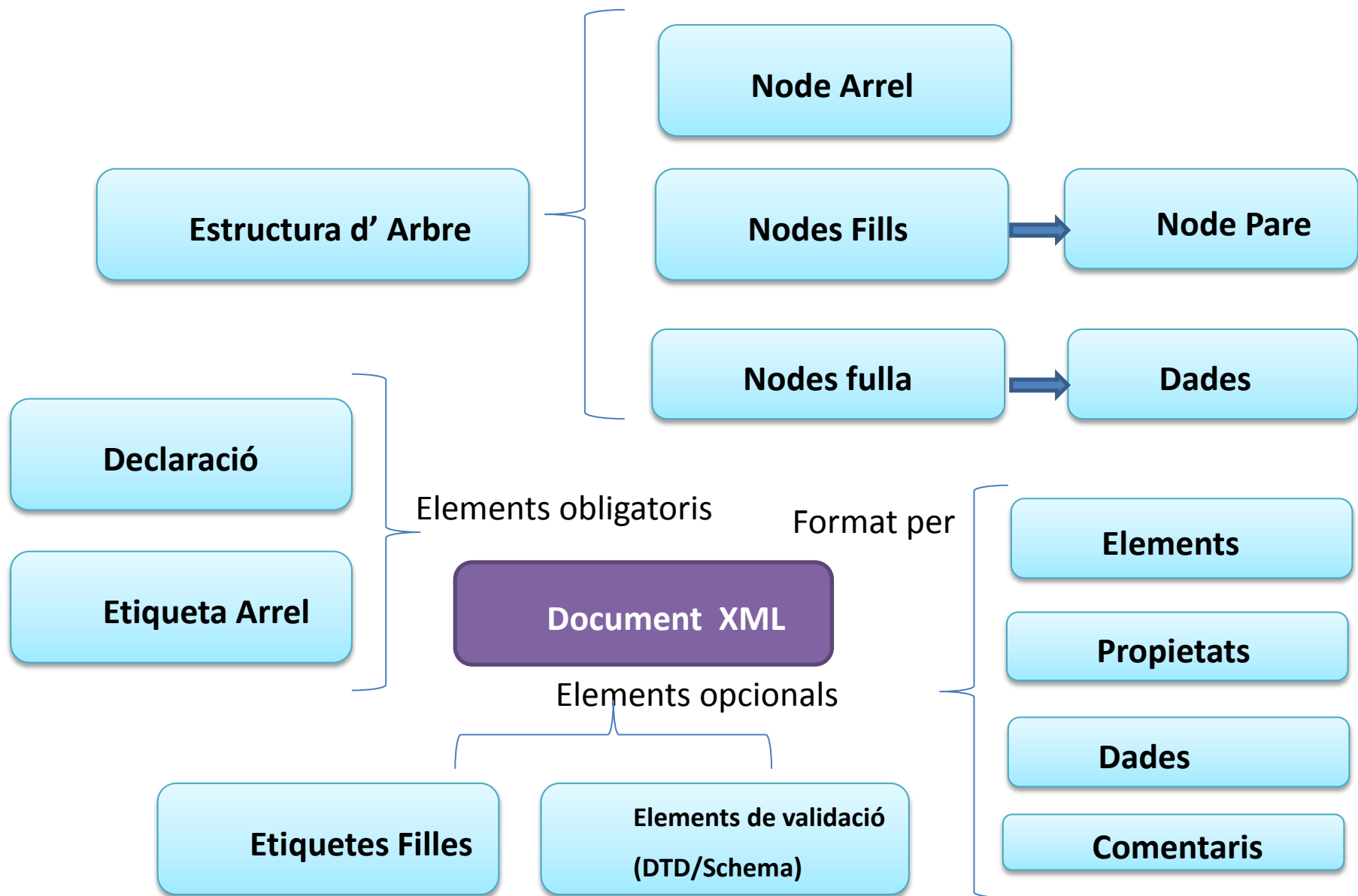
És aquell que segueix les regles de sintaxi marcades per l'standard (i que estudiarem en aquest tema)

Qualsevol navegador actual ens comprovarà si un document xml està o no ben format

Exemple: Documents ben formats / documents mal formats

- **Usar mozilla per deduir quin dels dos documents (Clients.xml i Clients2.xml) està mal format.**
- **Corregiu l'error que detecteu en el mal format i torneu a obrir el document amb mozilla**

Què cal saber identificar amb XML



Regles Sintàctiques d'un Document

Unes poques regles ... XML ben format

Els documents XML han de començar amb la **declaració** que indiqui quina versió estem fent servir d'XML . La declaració no és totalment obligatoria però és **altament recomanable**

```
<?xml version="1.0" ?>
```

Atribut "version": Generalment es fa servir la versió "1.0" . La 1.1 aporta poques novetats.

Atribut Encoding: Defineix el codi de caràcters que fem servir en ell document

Com que XML està definit per un organisme internacional fa servir Unicode per defecte (UTF-8 i UTF-16) però en pot fer servir d'altres

US-ASCII Anglès	Windows-1250 Europa central	ISO-8859-2 Est d'Europa
UTF-8 Unicode comprimit	Windows-1251 Ciríl·lic	ISO-8859-3 Sud d'Europa
UTF-16 UCS Comprimit	Windows-1253 Grec	ISO-8859-4 Nord d'Europa
Windows-1252 Oest d'Europa	ISO-8859-1 Oest d'Europa	ISO-2022-JP Japonès

Atribut standalone: Defineix si el document XML necessita o no un altre fitxer .

- **standalone="yes"**: Significa que el document és complet i per tant no li calen d' altres
- **standalone="no"**: Vol dir que el document no es pot entendre per si sol. Generalment degut a que li falta algun fitxer per validar-lo.

- Per tant la declaració d'un fitxer podria ser una cosa com això:

```
<?xml version="1.0" encoding="UTF-8" standalone="no" ?>
```

Segueix XML versió 1.0

Caràcters Unicode

Necessita un DTD per ser validat

Regles Sintàctiques d'un Document XML ben format

Més regles ...

<Import Brut>
<ImportBrut>
<Import_Brut>
<import_brut>
<impb>
<_ImportBrut>

Donant nom als elements (etiquetes)

- El nom ha de ser molt clar i identificar el significat del què contindrà
`<ciutat>Girona</ciutat>`
En lloc de
`<c>Girona</c>`
- Es permeten combinacions de lletres i numeros sense espais en blanc.
- Les etiquetes nom poden començar per xml
- Etiquetes són case sensitive
- Es poden incloure subguió , guió i punt però no es pot començar per cap d'aquests
- Cap altre símbol és acceptat com a component d'una etiqueta

Regles Sintàctiques d'un Document XML ben format

Cal que dedueu les regles sintàctiques d'un document XML ben format a partir de la web

http://www.w3schools.com/xml/xml_tree.asp

http://www.w3schools.com/xml/xml_syntax.asp

http://www.w3schools.com/xml/xml_elements.asp

http://www.w3schools.com/xml/xml_attributes.asp

Presenteu per parelles aquesta tasca en línia en l'espai moodle preparat a tal efecte.

Regles Sintàctiques d'un Document XML ben format

Més regles ...

Atributs predefinits ...

xml:lang: Permet especificar en quin idioma està el contingut. El valor ha de ser un codi d'idioma ISO

```
<missatge xml:lang="ca">Hola!</missatge>  
<missatge xml:lang="fr">Salut!</missatge>
```

xml:space: Per forçar a que es conservin els espais en blanc: 'default' o 'preserve'

```
<missatge xml:space="preserve">  
  Hola  
    Manolito!  
</missatge>
```

Regles Sintàctiques d'un Document XML ben format

Més regles ...

Espais de noms (namespaces)

Si ens cal distingir etiquetes iguals (perquè procedeixen de documents diferents o per qualsevol altre motiu) ens cal definir espais de noms

```
<alumnes:nom>Manolito Porrás</alumnes:nom>  
<professors:nom>Filomena Pi</professors:nom>
```

Regles Sintàctiques d'un Document XML ben format (namespaces)

- Els espais de noms s'han de carregar abans de l'inici de les dades del document XML

```
<?xml version="1.0">  
<classe xmlns:A="http://server/alumnes"  
        xmlns:P="http://server/professors">  
  <A:nom>Manolito Porras</A:nom>  
  <P:nom>Filomena Pi</P:nom>  
</classe>
```

- En la definició tenim

```
<classe xmlns:a="http://server/alumnes">
```

Paraula clau

Prefixe o alias de
l'espai de noms

Nom de l'espai

Regles Sintàctiques d'un Document XML ben format (namespaces)

- A pesar de la semblança dels noms amb les adreces web això només són noms
 - Si hi anem amb el navegador no hi trobarem cap document
 - Es van fer servir per aconseguir tenir noms d'espais de noms únics
- Podem definir quin és l'espai de noms per defecte sense posar-li prefix

```
<classe xmlns="http://server/alumnes">
```

Regles Sintàctiques d'un Document XML ben format (namespaces)

- No cal definir tots els espais de noms en l'arrel
- També es poden definir directament en els elements
 - També és vàlid pels seus descendents

```
<classe xmlns="http://server/alumnes">  
  <nom>Mariano Pardines</nom>  
  <nom xmlns="http://server/professors">  
    Federicu Pi  
  </nom>  
</classe>
```

Regles Sintàctiques d'un Document XML ben format (namespaces)

- Per tant és el mateix:

```
<?xml version="1.0">  
<classe xmlns:A="http://server/alumnes"  
        xmlns:P="http://server/professors">  
  <A:nom>Manolito Porras</A:nom>  
  <P:nom>Filomena Pi</P:nom>  
</classe>
```

```
<?xml version="1.0">  
<classe xmlns="http://server/alumnes"  
        xmlns:P="http://server/professors">  
  <nom>Manolito Porras</nom>  
  <P:nom>Filomena Pi</P:nom>  
</classe>
```

```
<?xml version="1.0">  
<classe xmlns="http://server/alumnes">  
  <nom>Manolito Porras</nom>  
  <nom xmlns="http://server/professors">Filomena Pi</nom>  
</classe>
```