



UD1 – Introducció als llenguatges de marques

2n CFGS
Desenrotllament d'Aplicacions Web
2024-25

1.- Introducció

Quin dels següents és un ordinador?



1



2



3



4



5



6



7



8



9



10



11



12



13

1.- Introducció

Què és un ordinador?

1.- Introducció

Què és un ordinador?

Màquina electrònica que permet **emmagatzemar i tractar informació** amb la finalitat de **resoldre problemes**.

Antigament els problemes a resoldre eren habitualment matemàtics, hui dia:

- **Oci.**
- **Xarxes socials.**
- Realitzar càlculs.
- Escriure textos.
- Presentacions.
- Aplicacions interactives.
- Control de màquines.
- ...

1.- Introducció

Els ordinadors, en ser electrònics **treballen amb voltatge**, concretament amb dos nivells de voltatge, la qual cosa dona el **codi binari** (0 i 1).

El codi binari és l'única cosa que entenen els ordinadors, així, el text, la música, les imatges, els vídeos... s'emmagatzemen en binari en l'ordinador.

Com **les persones no entenem el binari** ha d'haver-hi una transformació entre la informació que usem les persones i la que usen els ordinadors, a esta transformació se'l coneix com a **codificació**.

1.- Introducció

Exemples de codificació:

Tipo de dato	Dato	Codificación	Representación	Binario
Texto	J	ASCII	74	1001010
	J	UNICODE	U+004A	0000000001001010
	23	UNICODE	U+0032 U+0033	0000000000110010 0000000000110011
Color	Naranja	RGB decimal	255, 128, 0	11111111 10000000 00000000
	Naranja	RGB hexadecimal	#FF8000	11111111 10000000 00000000

Evidentment un arxiu només pot estar codificat d'una manera.

1.- Introducció

Com es codificaria el següent text en ASCII?

Jo **soc** programador.



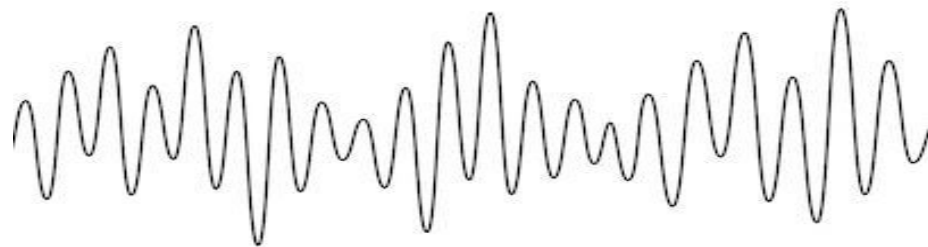
Y	o		s	o	y		p	r	o	g	r	a	m	a	d	o	r	.
89	111	32	115	111	121	32	112	114	111	103	114	97	109	97	100	111	114	46
1011001	1101111	0100000	1110011	1101111	1111001	0100000	1110000	1110010	1101111	1100111	1110010	1100001	1101101	1100001	1100100	1101111	1110010	0101110
101100111011110100000111001111011111110010100000111000011101011011111100111111001011000011101 1011100001110010011011111100100101110																		

Seria una codificació correcta?

1.- Introducció

Alguns tipus de dades es poden emmagatzemar directament amb la seua codificació binària com per exemple el so ja que tota la informació és d'un mateix tipus.

El so és una ona que en codificar-la es convertix en una seqüència en binari que permet reconstruir esta ona.



010110101100111100101100010111010111001110101000011101011100
11011101011100110101100111100101100010111010111001110101001...

1.- Introducció

El text pot a més pot contindre **informació sobre el format**.

Jo **soc** programador.

Per a emmagatzemar la informació de manera correcta l'ordinador ha d'incorporar informació addicional per a saber com representar posteriorment la informació.

S'haurà de **marcar el text** per a plasmar la informació addicional.

1.- Introducció

La idea de **marcar el text** prové del terme anglés "**marking up**" usat en la impremta tradicional.

Esta tècnica consistia a realitzar anotacions a mà en el manuscrit que servien com a instruccions per a l'impressor.

Els processadors de text van utilitzar de manera interna eixa mateixa tècnica però sense mostrar a l'usuari eixa informació.

1.- Introducció

Un arxiu .docx de Word realment és un arxiu comprimit amb diversos documents que emmagatzemen la informació i com mostrar la mateixa.

The diagram illustrates the internal structure of a Word document. On the left, a screenshot of the Microsoft Word interface shows the 'Inicio' (Home) tab. The ribbon includes options for 'Pegar' (Paste), 'Fuente' (Font), and 'Portapapeles' (Clipboard). The text 'Yo soy programador.' is displayed in the document area, with the word 'programador' highlighted in blue. A blue arrow points from this text to the XML code on the right. The XML code is a snippet from a Word document, showing the structure of a paragraph. It includes tags for paragraph, run, and run properties. The text 'Yo soy programador.' is represented by the following XML structure:

```
<w:p w:rsidR="00AC663E" w:rsidRDefault="00ED2180">
  <w:r>
    <w:t xml:space="preserve">Yo </w:t>
  </w:r>
  <w:r w:rsidRPr="00ED2180">
    <w:rPr>
      <w:b/>
    </w:rPr>
    <w:t>soy</w:t>
  </w:r>
  <w:r>
    <w:t xml:space="preserve"> </w:t>
  </w:r>
  <w:r w:rsidRPr="00ED2180">
    <w:rPr>
      <w:color w:val="4472C4" w:themeColor="accent1"/>
    </w:rPr>
    <w:t>programador</w:t>
  </w:r>
  <w:r>
    <w:t>.</w:t>
  </w:r>
  <w:bookmarkStart w:id="0" w:name="_GoBack"/>
  <w:bookmarkEnd w:id="0"/>
</w:p>
```

The XML code shows that the word 'programador' is styled with the color '4472C4' and the theme color 'accent1'. A blue arrow points from the '4472C4' value in the XML to a callout box on the right. The callout box is a blue square with the text '4472C4' and three smaller boxes below it containing the numbers '68', '114', and '196'.

2.- Llenguatges de marques

Els llenguatges de marques permeten **representar la informació** dotant-la de **representació i/o significat** fent ús de **marques en el propi text**.

Tipus:

- Orientats a la **presentació**.
- Orientats a la **descripció**.
- Orientats a **procediments**.

2.- Llenguatges de marques

Orientats a la **presentació**:

Les marques indiquen **com s'ha de mostrar la informació**.

- Grandària de la pàgina
- Tipus de lletra
- Grandària
- Color
- Marges
- ...

Yo **soy** programador.

```
[Arial, 16, negro, normal]  
Yo  
[Arial, 16, negro, negrita]  
soy  
[Arial, 16, azul, normal]  
programador  
[Arial, 16, negro, normal]  
.
```

2.- Llenguatges de marques

Orientats a la **descripció**:

Les marques indiquen **el tipus d'informació** que s'està emmagatzemant.

```
<?xml version="1.0" encoding="UTF-8"?>
<serie tipo="animación" anyo="2013" terminada="no">
  <nombre>Rick y Morty</nombre>
  <personajes>
    <personaje especie="humano">
      <nombre>Rick</nombre>
      <apellido>Sanchez</apellido>
      <edad>70</edad>
      <estado>Vivo</estado>
      <origen>Tierra C-137</origen>
    </personaje>
    <personaje especie="humano">
      <nombre>Morty</nombre>
      <apellido>Smith</apellido>
      <edad>14</edad>
      <estado>Vivo</estado>
      <origen>Desconocido</origen>
    </personaje>
  </personajes>
  <temporadas>
    7
  </temporadas>
</serie>
```

2.- Llenguatges de marques

Orientats a procediments:

Les marques indiquen **instruccions que s'han de realitzar amb el text.**

- Girar-ho.
- Convertir-ho en una fórmula.
- Realitzar una suma.
- ...

3.- Evolució històrica

En la **dècada de 1960** les empreses que publicaven, distribuïen i manejaven documents electrònics **no seguien cap estàndard** per a descriure els diferents elements.

Al no existir un estàndard a seguir, cada aplicació seguia les seues pròpies regles, això **no permetia que els documents es pogueren usar en altres plataformes**.

A la fi de la dècada dels 1960, **Charles Goldfarb i el seu equip en IBM**, van proposar idees perquè els documents de text inclogueren la possibilitat de marcar el format d'este.

4.- GML

Xarres Goldfarb, treballador i investigador d'IBM, va liderar l'equip que va crear el llenguatge **GML** (**G**eneralized **M**arkup **L**anguage).

Goldfarb va crear l'acrònim GML a partir de les inicials dels 3 membres de l'equip que van crear el llenguatge de marques GML: Xarres **G**oldfarb, Edward **M**osher i Raymond **L**orie.

L'ISO (**I**nternational **O**rganization for **S**tandardization) va crear un estàndard a partir de GML → SGML (**S**tandard **G**ML) que encara que no va acabar d'assentar-se, va crear els fonaments i les bases dels llenguatges de marques usats hui dia.

4.- GML

Codi GML:

:h0.Rick y Morty

:p.Ánimación - :hp1.2013:ehp1. - Sin terminar

:p.7 temporadas

:h1.Personajes

:hp1.Rick Sanchez:ehp1.

:p.70 años (Vivo) origen: Tierra C-137

:hp1.Morty Smith:ehp1.

:p.14 años (Vivo) origen: :hp2.Desconocido:ehp2.

Renderització del codi GML:

Rick y Morty

Ánimación - **2013** - Sin terminar

7 temporadas

Personajes

Rick Sanchez

70 años (Vivo) origen: Tierra C-137

Morty Smith

14 años (Vivo) origen: *Desconocido*

5.- SGML

Charles Goldfarb va crear un estàndard a partir de GML per a l'ISO (International Organization for Standardization).

SGML (**S**tandard **G**ML) que encara que no va acabar d'assentar-se, va crear els fonaments i les bases dels llenguatges de marques usats hui dia.

En SGML el marcat es realitza amb etiquetes entre els símbols < i >, a més les etiquetes tenen obertura i tancament:

<etiqueta>contingut</etiqueta>

Esta manera de marcar és la que s'utilitza en els llenguatges **XML** i **HTML** que són dos dels llenguatges de marques més populars de la història i que encara s'utilitzen.

Per esta raó es podria dir que SGML és **el pare dels llenguatges XML i HTML**,

6.- TeX i LaTeX

En la dècada de 1970 Donald Knuth, un dels enginyers informàtics més importants de la història, va crear el llenguatge **TeX**.

El llenguatge **TeX** permetia produir documents científics amb una tipografia i unes capacitats que permetia representar la informació de la mateixa manera siga el que siga la màquina on s'usava.

Posseïx **300 comandos** per a crear els documents.

Es necessita d'un programari capaç de convertir l'arxiu Tex a un format d'impressió.

De tots els derivats que es van produir de TeX, el més popular és **LaTeX**, que des que es va crear en 1984 ha patit moltes revisions i es continua utilitzant en l'actualitat.

6.- TeX i LaTeX

Codi LaTeX:

```
\documentclass[article]{article}

% Packages/Macros %
\usepackage{amsmath} % Standard packages

% Margins %
\addtolength{\textwidth}{1.0in}
\addtolength{\textheight}{1.00in}
\addtolength{\evensidemargin}{-0.75in}
\addtolength{\oddsidemargin}{-0.75in}
\addtolength{\topmargin}{-.50in}

% Document %

\begin{document}
\title{Rick y Morty}
\author{(Animación - \textbf{2013} - Sin terminar)}
\date{}
\maketitle

\begin{center}
7 temporadas
\end{center}

\section{Personajes}
\begin{Personajes}
\item {\bf Rich Sanchez} \\
\begin{theorem}
70 anyos (Vivo) origen: Tierra C-137
\end{theorem}
\item {\bf Morty Smith} \\
\begin{theorem}
14 anyos (Vivo) origen: \emph{Desconocido}
\end{theorem}
\end{Personajes}

\begin{center}
 $E=mc^2$ 
\end{center}

\end{document}
```

Renderització del codi LaTeX:

Rick y Morty
(Animacin - **2013** - Sin terminar)

7 temporadas

1 Personajes

Rich Sanchez
70 anyos (Vivo) origen: Tierra C-137

Morty Smith
14 anyos (Vivo) origen: *Desconocido*

$$E = mc^2$$

7.- RTF

El llenguatge RTF (**Rich Text Format**) va ser ideat en 1987 per Richard Brodie, Charles Simonyi i David Luebbert que pertanyien a l'equip **de desenrotllament de Microsoft Word**.

És un format que permet incloure anotacions en el text que indiquen com s'ha de formatar el text.

Codifica el text usant 8 bits, per la qual cosa per a incloure caràcters que no es troben en ASCII s'han d'usar seqüències de fuga, això fa que siga inviable com a format llegible sense renderitzar el text (en les últimes versions té suport per a UNICODE).

7.- RTF

El seu èxit radica en el fet que les indicacions de format són molt potents i són més llegibles per les persones que el format natiu dels processadors de text.

Com a llenguatge de marcat és un dels més críptics.

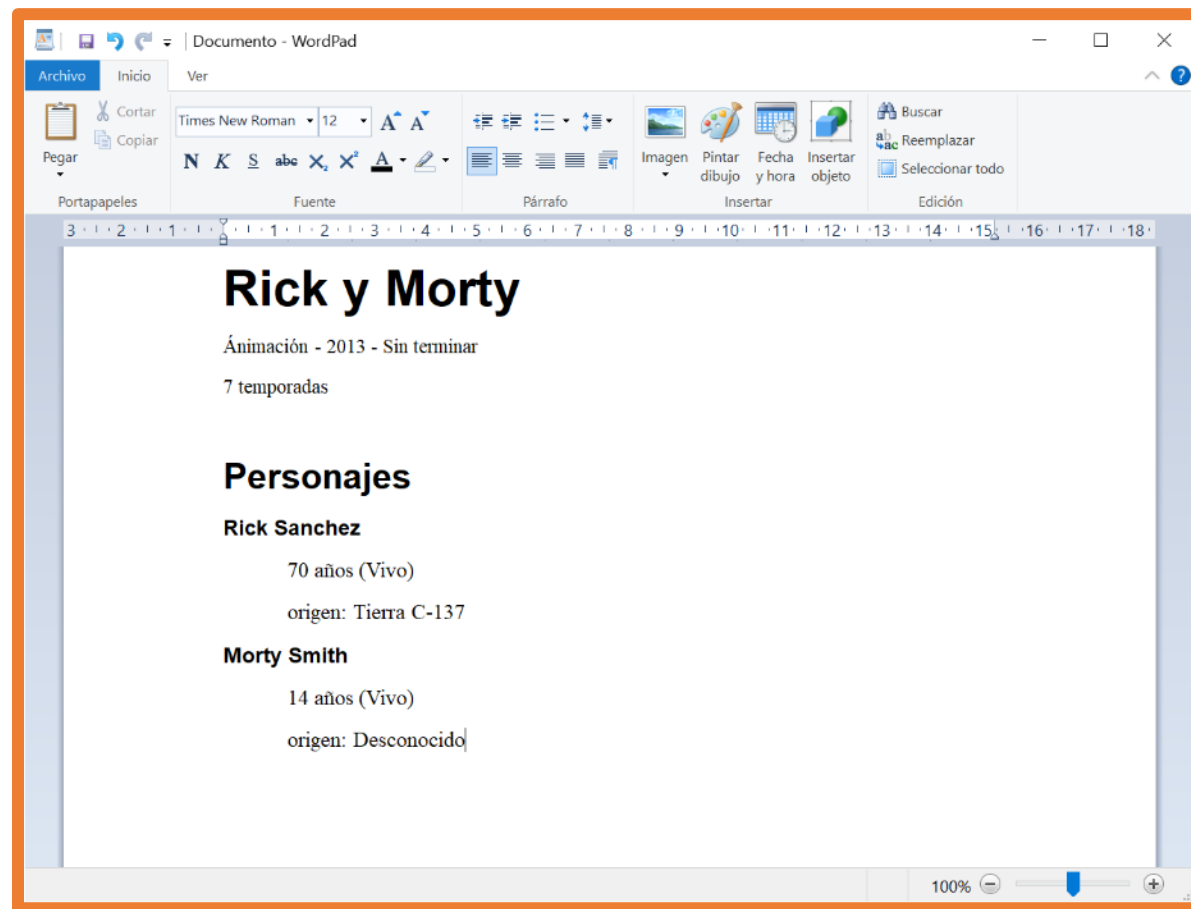
Encara que és propietat de Microsoft, és un format reconegut per la majoria de processadors de text.

7.- RTF

Codi RTF:

```
{\rtf1\ansi\ansicpg1252\deff0\nouicompat\deflang3082
{\fonttbl{\f0\fnl\fcharset0 Arial;}
{\f1\fnl\fcharset0 Times New Roman;}
}
{* \generator Riched20 10.0.19041} \viewkind4 \uc1
\pard \sa200 \sl240 \slmult1 \b \fs56 \lang10 Rick y Morty \par
\b0 \f1 \fs22 'c1nimaci' \f3n - 2013 - Sin terminar \par
7 temporadas \par
\f0 \par
\b \fs40 Personajes \b0 \par
\b \fs24 Rick Sanchez \par
\b0 \f1 \tab 70 a \f1os (Vivo) \par
\tab origen: Tierra C-137 \par
\b \f0 Morty Smith \par
\b0 \f1 \tab 14 a \f1os (Vivo) \par
\tab origen: Desconocido \par
}
```

Renderització RTF:



8.- PostScript

Va començar el seu desenvolupament en 1976 per part de John Warnock, va continuar el seu desenvolupament en l'empresa Xerox i finalment quan John Warnock va crear Adobe Systems 1985 es continua el seu desenvolupament des d'allí.

Realment és un llenguatge de programació que indica la forma en la qual s'ha de mostrar la informació i que pot incloure text i el tipus de lletra d'este, a més de píxels individuals i formes vectorials.

```
%colocar el cursor  
100 100 moveto  
%dibujar cuadrado  
100 200 lineto  
200 200 lineto  
200 100 lineto  
100 100 lineto  
%relleno  
stroke
```

9.- HTML

HTML – HyperText Markup Language

Creat per **Tim Berners Lee** en la dècada de 1990 en base de les directrius de SGML.

Permetia crear documents transportables a través d'Internet en els quals era possible l'ús d'hipertext (marcar paraules que servien d'enllaç a altres documents).

Encara que va tardar a ser acceptat, el seu èxit va ser rotund i hui dia és el llenguatge usat per a crear les pàgines web.

9.- HTML

En els seus inicis els documents HTML s'havien de visualitzar utilitzant **intèrprets de text**, com **Lynx** d'UNIX.

Estos intèrprets de text **acolorien el text i marcaven els enllaços** que apareixien en ell (hipertext).

Més endavant **el programari va millorar** i van aparéixer els **navegadors** que augmentaven la capacitat incloent formats més avançats i visuals.

```
<!DOCTYPE html>
<html lang="es">
<head>
  <meta charset="UTF-8">
  <meta name="viewport" content="width=device-width, initial-scale=1.0">
  <title>Rick y Morty</title>
</head>
<body>
  <h1>Rick y Morty</h1>
  Animación - 2013 - Sin terminar<br>
  7 temporadas

  <h2>Personajes</h2>
  <h3>Rick Sanchez</h3>
  <p>
    70 años (Vivo)<br>
    origen: Tierra C-137
  </p>

  <h3>Morty Smith</h3>
  <p>
    14 años (Vivo)<br>
    origen: Desconocido
  </p>
</body>
</html>
```

10.- XML

XML – e**X**tensible **M**arkup **L**anguage

És un subconjunt de SGML que **permet definir llenguatges de marcat** usant una sintaxi més estricta però més comprensible.

Des de finals de la dècada de 1990 es popularitza incorporant nombrosos llenguatges al seu voltant per aconseguir documents molt dinàmics i amb gran capacitat de format.

Fins i tot es va crear una versió d'HTML sobre les bases de XML per a estandarditzar la manera en la qual els navegadors havien de mostrar els documents HTML: **XHTML**.

10.- XML

En l'actualitat, un dels usos més estesos de XML és el de **serialització de dades**, que consistix a **exportar la informació** dotant-la de sentit semàntic.

Encara que hui dia continua tenint ús, ha sigut sobrepassat per **JSON** a causa de les característiques d'este últim.

```
<?xml version="1.0" encoding="UTF-8"?>
<serie tipo="animación" anyo="2013" terminada="no">
  <nombre>Rick y Morty</nombre>
  <personajes>
    <personaje especie="humano">
      <nombre>Rick</nombre>
      <apellido>Sanchez</apellido>
      <edad>70</edad>
      <estado>Vivo</estado>
      <origen>Tierra C-137</origen>
    </personaje>
    <personaje especie="humano">
      <nombre>Morty</nombre>
      <apellido>Smith</apellido>
      <edad>14</edad>
      <estado>Vivo</estado>
      <origen>Desconocido</origen>
    </personaje>
  </personajes>
  <temporadas>
    7
  </temporadas>
</serie>
```

11.- JSON

JSON – JavaScript Object Notation

És una manera de representar (serialitzar) les dades contingudes en els objectes en el llenguatge JavaScript (concretament en la versió ECMAScript de 1999).

Per la seua versatilitat s'ha convertit en una notació independent de JavaScript que competix amb els XML usats per a exportar informació.

Realment JSON **no és un llenguatge de marcat** però realitza les mateixes funcions que XML d'una manera molt més eficient.

11.- JSON

En JSON el text es dividix en **dades** i **metadades**.

Una dada indica el tipus d'informació i una metadada indica la informació en si.

Dades i metadades s'associen mitjançant el símbol : (dos punts), a més es poden usar les claus { } i els claudàtors [] per a realitzar agrupacions.

```
{  
  "serie": "Rick y Morty",  
  "anyo": "2013",  
  "terminada": "no",  
  "personajes": [  
    {  
      "tipo": "humano",  
      "nombre": "Rick",  
      "apellido": "Sanchez",  
      "edad": 70,  
      "estado": "Vivo",  
      "origen": "Tierra C-137"  
    },  
    {  
      "tipo": "humano",  
      "nombre": "Morty",  
      "apellido": "Smith",  
      "edad": 14,  
      "estado": "Vivo",  
      "origen": "Desconocido"  
    }  
  ],  
  "temporadas": 7  
}
```

12.- YAML

YAML, igual que JSON, **no és un llenguatge de marcat** però permet les mateixes funcionalitats que JSON.

YALM Ain't Markup Language.

Encara que al principi del seu desenrotllament a principis dels 2000, el seu significat original era:

Yet Another Markup Language
(Un altre Llenguatge de Marcat Més)

En YAML s'utilitza la sagnia amb espais per a indicar que un element es troba dins d'un altre.

```
services:
  pihole:
    container_name: pihole
    image: pihole/pihole:latest
    environment:
      TZ: 'Europe/Madrid'
      # WEBPASSWORD: 'set a secure password here or it will be random'
      WEBPASSWORD: cacafuti
      FTLCONF_LOCAL_IPV4: 192.168.1.250
    volumes:
      - './etc-pihole:/etc/pihole'
      - './etc-dnsmasq.d:/etc/dnsmasq.d'
    restart: unless-stopped
```


13.- Markdown

Markdown és un llenguatge de marcat lleuger disenrotllat en 2004 per John Gruber amb l'ajuda d'Aaron Swartz.

L'objectiu de Markdown és crear documents senzills que es puguin convertir en documents HTML en ser renderitzats.

El seu èxit ha sigut tan notable que GitHub, Reddit o StackExange entre altres ho han adoptat perquè els usuaris publiquen contingut amb format.

13.- Markdown

Codi Markdown:

```
# Rick y Morty
Animación - **2013** - Sin terminar
7 temporadas

## Personajes
### Rick Sanchez
| 70 años (Vivo) origen: Tierra C-137
### Morty Smith
| 14 años (Vivo) origen: Desconocido
```

Renderització Markdown:

Rick y Morty

Animación - **2013** - Sin terminar 7 temporadas

Personajes

Rick Sanchez

70 años (Vivo) origen: Tierra C-137

Morty Smith

14 años (Vivo) origen: Desconocido

14.- Extensions d'arxius de llenguatges de marques

Lenguaje	Extensión
GML	.gml
SGML	.sgml
Tex	.tex
LaTeX	.tex
RTF	.rtf
PostScript	.ps .prn
HTML	.htm .xhtml .html (según la versión usada)
XML	.xml
Markdown	.md
JSON	.json
YAML	.yaml .yml

15.- Característiques i components

Etiquetes

Els llenguatges de marques utilitzen **etiquetes** per a **marcar el text**.

Quan s'interpreta o renderitza el text, **estes etiquetes ajuden el programari a saber com representar la informació (format)**.

Prenent com a exemple el llenguatge de marques XML, les etiquetes s'escriuen en minúscules entre els símbols **<** i **>**.

A més, la manera de marcar el text és tancant-lo entre una etiqueta d'obertura i una de **tancament**, l'etiqueta de tancament sempre comença amb el símbol **/**.

15.- Característiques i components

Jerarquia

La posició de les marques és molt important ja que determina una jerarquia dins de la informació.

Si una marca es troba dins d'una altra, significa que eixa informació pertany a la seua jerarquia → pare – fill.

Si una etiqueta conté a una altra, no es pot tancar una etiqueta pare si abans no s'ha tancat l'etiqueta fill.

La jerarquia és molt important tant en els llenguatges de marques orientats a la presentació com en els orientats a la descripció.

15.- Característiques i components

Atributs

Mitjançant els atributs es pot afegir informació o propietats a les etiquetes.

Els atributs no permeten jerarquia.

Una etiqueta pot no tindre atributs o pot tindre un o més atributs.

Els atributs s'indiquen dins de l'etiqueta d'obertura, se'ls assigna valor mitjançant el símbol = i el valor de l'atribut s'ha d'indicar entre cometes ".

Un atribut pot tindre diversos valors, per a això dins de les cometes se separen els valors mitjançant l'espai.

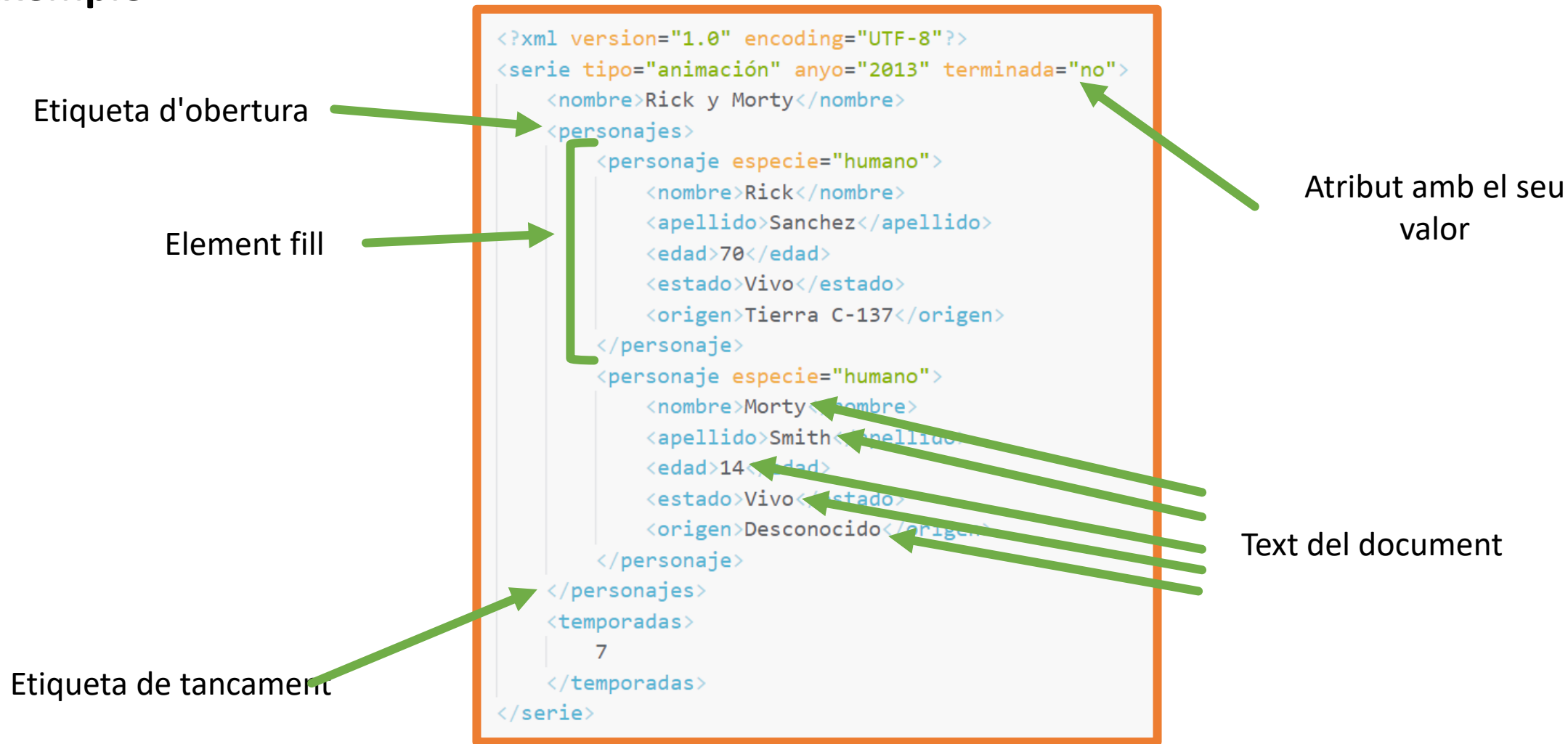
15.- Característiques i components

Exemple XML

```
<?xml version="1.0" encoding="UTF-8"?>
<serie tipo="animación" anyo="2013" terminada="no">
  <nombre>Rick y Morty</nombre>
  <personajes>
    <personaje especie="humano">
      <nombre>Rick</nombre>
      <apellido>Sanchez</apellido>
      <edad>70</edad>
      <estado>Vivo</estado>
      <origen>Tierra C-137</origen>
    </personaje>
    <personaje especie="humano">
      <nombre>Morty</nombre>
      <apellido>Smith</apellido>
      <edad>14</edad>
      <estado>Vivo</estado>
      <origen>Desconocido</origen>
    </personaje>
  </personajes>
  <temporadas>
    7
  </temporadas>
</serie>
```

15.- Característiques i components

Exemple



15.- Característiques i components

Estàndard

Per a poder transmetre la informació es necessita parlar el mateix idioma.

En els llenguatges de marques ocorre el mateix, cada llenguatge de marques ha de tindre un estàndard que han de seguir aquelles aplicacions que el vulguen usar, d'eixa manera, les aplicacions representaran la informació de la mateixa manera.

Per als llenguatges de marques moderns que partixen de SGML com són XML i HTML els estàndards s'establixen per la [W3C](#) – **World Wide Web Consortium**.

Un document "ben format" és aquell en el qual la seua estructura sintàctica és vàlida, això vol dir que es compon els elements, atributs i comentaris tal com indica l'estàndard.

15.- Característiques i components

Validació

Que un document estiga "ben format" no significa que supere la validació.

La validació indica com han d'usar-se les etiquetes del llenguatge de marcat, per exemple, si una etiqueta pot tindre o no un atribut o si una etiqueta pot ser filla d'una altra.

Per això és molt important que els documents realitzats també superen la validació perquè els programes usats per a la seua visualització els sàpien interpretar correctament.

Alguns llenguatges de marques necessiten d'un document extern en el qual es definix com es poden utilitzar les etiquetes.

16.- Ferramentes

Les ferramentes utilitzades en el treball amb llenguatges de marques són:

- Editor de text.
- Validador.
- Programa de renderització.

Encara que es pot utilitzar qualsevol editor de text, l'habitual és utilitzar algun editor que continga característiques que faciliten el treball per a cada tipus de llenguatge de marques:

- Ressaltat d'etiquetes.
- Identificació d'atributs.
- autocompletat d'etiquetes i tancat automàtic d'estes..
- Anàlisi sintàctica i formal del codi escrit.
- Contracció jeràrquica d'etiquetes.
- ...

L'editor Visual Studio Code permet treballar amb molts tipus de llenguatges, inclosos els llenguatges de marques. A més, se li poden instal·lar extensions per a potenciar-ho. És l'editor que es recomana per a usar en el mòdul.

17.- Enllaços interessants

Editor i visor de LaTeX: <https://es.overleaf.com/>

Validador JSON: <https://jsonlint.com/>

Editor i visor de *Markdown: <https://markdownlivepreview.com/>

Pràctica

Activitat 1:

Visual Studio Code i llenguatges de marques.