

# **Aplicații Multimedia**

## **Laborator 3**

### **Documente Multimedia**

#### Cuprins

<b>Fișiere XML.....</b>	<b>1</b>
<b>Fișiere DOCX .....</b>	<b>2</b>
<b>Fișiere PDF .....</b>	<b>3</b>
<b>Fișiere CSV .....</b>	<b>3</b>

## Fișiere XML

Fișierele XML (eXtensible Markup Language) sunt folosite pentru stocarea și transportul datelor, mai ales atunci când o bază de date nu s-ar preta aplicației.

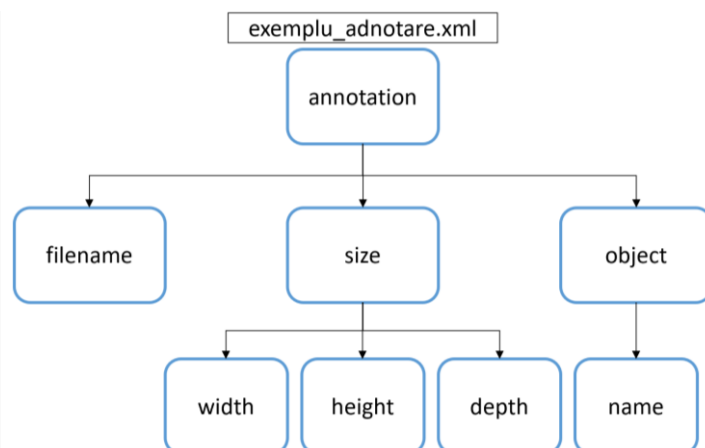
XML este un limbaj ce folosește etichete (<exemplu\_eticheta> </exemplu\_eticheta>), asemenea HTML, cu distincția că etichetele nu sunt predefinite, ca în cazul HTML, ci sunt alese de cel care scrie sau structurează fișierul.

XML pune la dispoziție o modalitate de a structura datele, nu de a le prezenta, fiind necesar un program care să folosească respectivele fișiere (să le creeze, scrie, citească sau afișeze).

Dintre numeroasele sale aplicații, vom folosi în acest laborator un exemplu de fișier XML utilizat la stocarea de adnotări pentru identificarea de obiecte în imagini, pentru aplicații de procesare de imagine sau de inteligență artificială. Acestea conțin aspecte precum denumirea imaginii la care se face referire, dimensiunea acesteia și numele obiectelor ce se găsesc în aceasta și poziția lor, în pixeli sau în procente.

Exemplul de mai jos este derivat din formatul de adnotări Pascal VOC, despre care puteți citi mai multe [aici](#). Fișierele XML au o structură de arbore, pornind de la elementul rădăcină, care se ramifică până la frunze, care poartă denumirea de copii. În exemplul de mai jos, <annotation> este elementul rădăcină, fiind părinte pentru subelementul <size>, copilul său.

```
<!--exemplu_adnotare.xml-->
<annotation>
  <filename>exemplu.jpeg
</filename>
  <size>
    <width>1248</width>
    <height>384</height>
    <depth>3</depth>
  </size>
  <object>
    <name>Door</name>
  </object>
</annotation>
```

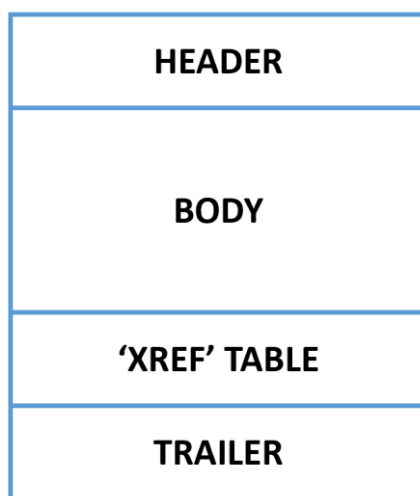




## Fișiere PDF

PDF (Portable Document Format) este un format creat inițial de Adobe, ca o modalitate universală de a prezenta documente cu diverse tipuri de conținut, de la link-uri, butoane, câmpuri de introdus text, la audio, video și chiar modele 3D. PDF este un format open, menținut de ISO (International Organization for Standardization).

Structura de bază a unui document PDF este următoarea:



Puteți citi mai multe despre structura unui PDF, organizarea acestuia și tipurile de date ce pot fi folosite accesând [acest articol](#).

Documente PDF pot fi generate din numeroase aplicații, însă pot fi și construite programatic sau prin alte limbaje de markup. Poate cea mai cunoscută modalitate de a genera documente PDF este prin intermediul limbajului LaTeX, folosit în special în mediul academic și științific.

## Fișiere CSV

CSV (Comma-Separated Value) este un format simplu de fișier care utilizează virgula ca separator între valorile pe care le conține. În acest mod sunt stocate date care, de exemplu, ar putea fi conținute într-un tabel, sau într-un program precum Microsoft Excel sau Google Spreadsheet.

Primul rând al unui fișier CSV trebuie să conțină capul de tabel, adică denumirile coloanelor, pentru ca următoarele rânduri să fie reprezentate de datele efective.

Ca regulă de folosire, dacă o celulă este goală, fișierul CSV nu va conține niciun caracter între cele 2 virgule care delimitează câmpul respectiv.

```
"LatD", "LatM", "LatS", "NS", "LonD", "LonM", "LonS", "EW", "City", "State"
41, 5, 59, "N", 80, 39, 0, "W", "Youngstown", OH
42, 52, 48, "N", 97, 23, 23, "W", "Yankton", SD
46, 35, 59, "N", 120, 30, 36, "W", "Yakima", WA
42, 16, 12, "N", 71, 48, 0, "W", "Worcester", MA
43, 37, 48, "N", 89, 46, 11, "W", "Wisconsin Dells", WI
36, 5, 59, "N", 80, 15, 0, "W", "Winston-Salem", NC
49, 52, 48, "N", 97, 9, 0, "W", "Winnipeg", MB
39, 11, 23, "N", 78, 9, 36, "W", "Winchester", VA
34, 14, 24, "N", 77, 55, 11, "W", "Wilmington", NC
39, 45, 0, "N", 75, 33, 0, "W", "Wilmington", DE
48, 9, 0, "N", 103, 37, 12, "W", "Williston", ND
41, 15, 0, "N", 77, 0, 0, "W", "Williamsport", PA
37, 40, 48, "N", 82, 16, 47, "W", "Williamson", WV
```

Exemplu CSV text

LatD	"LatM"	"LatS"	"NS"	"LonD"	"LonM"	"LonS"	"EW"	"City"	"State"
41	5	59	"N"	80	39	0	"W"	"Youngstown"	OH
42	52	48	"N"	97	23	23	"W"	"Yankton"	SD
46	35	59	"N"	120	30	36	"W"	"Yakima"	WA
42	16	12	"N"	71	48	0	"W"	"Worcester"	MA
43	37	48	"N"	89	46	11	"W"	"Wisconsin Dells"	WI
36	5	59	"N"	80	15	0	"W"	"Winston-Salem"	NC
49	52	48	"N"	97	9	0	"W"	"Winnipeg"	MB
39	11	23	"N"	78	9	36	"W"	"Winchester"	VA
34	14	24	"N"	77	55	11	"W"	"Wilmington"	NC
39	45	0	"N"	75	33	0	"W"	"Wilmington"	DE
48	9	0	"N"	103	37	12	"W"	"Williston"	ND
41	15	0	"N"	77	0	0	"W"	"Williamsport"	PA
37	40	48	"N"	82	16	47	"W"	"Williamson"	WV

Exemplu CSV deschis in Excel