

## *Tipos de Información | LMWeb*

*Víctor Peinado*

*1 de marzo de 2016*

### *Referencias*

- Introducción de Salton <sup>1</sup>
- Formatos de ficheros <sup>2</sup>
- Texto plano <sup>3</sup>
- Texto con formato <sup>4</sup>

<sup>1</sup> Salton, G. "The Information Environment".

<sup>2</sup> Wikipedia: Formatos de ficheros  
[http://en.wikipedia.org/wiki/List\\_of\\_file\\_formats#Document](http://en.wikipedia.org/wiki/List_of_file_formats#Document)

<sup>3</sup> Wikipedia: Texto plano [http://en.wikipedia.org/wiki/Plain\\_text](http://en.wikipedia.org/wiki/Plain_text)

<sup>4</sup> Wikipedia: Texto con formato [http://en.wikipedia.org/wiki/Formatted\\_text](http://en.wikipedia.org/wiki/Formatted_text)

### *Tipos de Información*

La información puede presentarse de tres formas distintas:

#### *Texto escrito*

El texto escrito es el medio básico para la comunicación formal entre seres humanos. Puede tomar la forma de notas, cartas, mensajes escritos, libros, periódicos, etc. Para manejarse con él, es indispensable saber leer y escribir. Esto que parece una obviedad en nuestro entorno, no lo ha sido hasta hace relativamente poco tiempo.

#### *Habla*

El habla es el medio informal y natural de comunicación entre seres humanos. Es accesible para gente que no sabe leer ni escribir. El reto actual (donde los lingüistas tenemos mucho que aportar) es que el habla sea también el vehículo de comunicación entre seres humanos y máquinas.

#### *Imágenes y vídeo*

Las imágenes y el vídeo se suelen utilizar para complementar información que se proporciona en forma de texto: como cuando utilizamos gráficos o diagramas. En cualquier caso, puede utilizar sin texto, en forma de ilustraciones, fotografía, pinturas, películas.

Representar, manejar y procesar estos tres tipos de información presenta distintos problemas y retos. Veamos cómo almacenamos estos tipos de información en formato digital.

## Documentos de texto

Distinguimos dos tipos de documentos textuales: texto plano y texto con formato.

### Texto plano

- secuencia ordinaria de caracteres, legible como texto.
- contenido textual intercambiable, sin añadidos.
- no contiene información de apariencia o estilo.
- puede contener información de estructura: textos con **marcas** o meta-datos.
- visualizable y editable con cualquier **editor de texto**

### Texto con formato

- secuencia de caracteres + información de estilo (tipos de letra, tamaños, colores, estructura, hiperenlaces)
- no necesariamente intercambiable, sino transformable.
- visualizable y editable con **procesador de texto**, habitualmente de tipo WYSIWYG.

### Formatos de documentos de texto

De todos los formatos de documentos de texto,<sup>5</sup> los más habituales son:

Entre los formatos de **texto** más habituales, tenemos:

- texto plano: \*.txt, \*.text
- lenguajes de marcado: \*.html, \*.xml, \*.md, \*.tex
- código de lenguajes de programación: \*.java, \*.py, \*.pl, \*.c, \*.php, \*.go

Tenemos también formatos de **ficheros binarios**.

- formato MS-Word: \*.doc, \*.docx
- formato de intercambio de MS: \*.rtf
- formato OpenDocument: \*.odt
- formato PDF: \*.pdf

### Formatos de ficheros de audio

De todos los formatos de documentos de audio,<sup>6</sup> los más habituales son:

#### 1. audio sin pérdida

Hay centenares de editores de textos, algunos más sencillos, otros más potentes, algunos orientados a escribir código de programación. En negrita, algunos de los que utilizaremos más adelante en clase: Notepad, **Notepad++**, **atom**, SublimeText, UltraEdit, Emacs, Vim, nano, Gedit, Textmate, SimpleText, TextEdit, etc.

Conoces y usas habitualmente algún procesador de texto incluido alguna *suite* ofimática: MS-Word, LibreOffice/OpenOffice, Google Drive, KOffice, Abiword, MacWrite, WordStar, WordPerfect, etc.

<sup>5</sup> Wikipedia: Ficheros de texto [http://en.wikipedia.org/wiki/List\\_of\\_file\\_formats#Document](http://en.wikipedia.org/wiki/List_of_file_formats#Document)

<sup>6</sup> Wikipedia: Ficheros de audio [http://en.wikipedia.org/wiki/List\\_of\\_file\\_formats#Sound\\_and\\_music](http://en.wikipedia.org/wiki/List_of_file_formats#Sound_and_music)

- audio no comprimido: \*.wav, \*.aiff
- audio comprimido: \*.wma, \*.flac, \*.m4a

## 2. **audio con pérdida** (y comprimido)

- \*.mp3, \*.aac, \*.ra, \*.speex

## *Formatos de ficheros de imagen*

De todos los formatos de fichero de imagen,<sup>7</sup> los más habituales son:

<sup>7</sup> Wikipedia: Ficheros de imágenes  
[http://en.wikipedia.org/wiki/Image\\_file\\_formats](http://en.wikipedia.org/wiki/Image_file_formats)

### 1. **comprimido sin pérdida**

- \*.raw, \*.tiff

### 2. **comprimido con pérdida**

- \*.jpg, \*.jpeg, \*.png, \*.gif

### 3. **no comprimido**

- \*.bmp