

Tema 1: Documentos electrónicos y la Edición Digital



INDICE

La Biblioteca Digital

Documentos electrónicos:

-Digitalización de Documentos

-Imágenes digitales: principios y características

a) mapas pixeles

b) vectoriales

-Formatos de escritura

-Formatos de imagenes

-e-book readers: principios y formatos

-Aspectos legales

-Preservación de documentos digitales

•

Bibliotecas Digitales. Hispanoamericanas

- **Biblioteca Virtual Miguel de Cervantes**

www.cervantesvirtual.com

Objetivo: digitalizar del patrimonio bibliográfico, documental y crítico de la cultural española e hispanoamericana.

- **Biblioteca Somni (Universitat de València) <http://somni.uv.es>**

Objetivo: consulta, preservación y difusión pública través de Internet de valioso patrimonio bibliográfico conservado en la Universidad de Valencia

Bibliotecas Digitales. Iniciativas Privadas

- **Biblioteca Digital de Google <http://books.google.es/books>**
 - Digitalizar 15 millones de libros (asociaron bibliotecas norteamericanas y europeas)
 - Críticas por derechos de autor
- **Biblioteca Digital de Yahoo (www.opencontentalliance.org)**
 - Pretende digitalizar 150.000 libros en una fase
 - Se ha comprometido a pedir el permiso correspondiente para la publicación en Internet de cada imagen digitalizada
 - Intervienen Microsoft, HP y Adobe y la biblioteca Británica(acuerdo digitalizar 100.000 libros)

Digitalización.

- **Documento Digital (documento electrónico)**

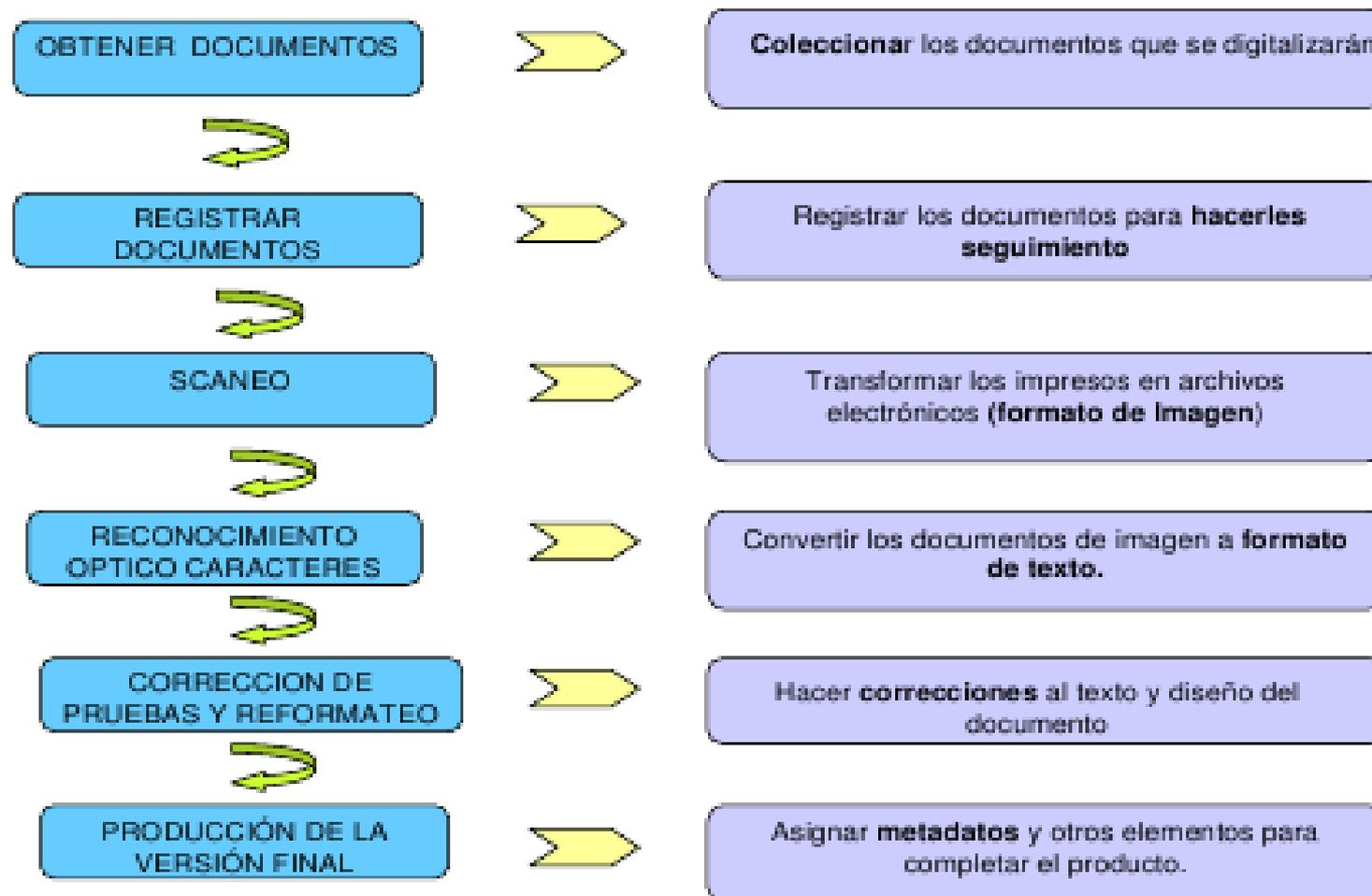
- Aquel documento que necesita de una máquina para ser leído
- Si está codificado en binario-->precisa de un ordenador para ser leído.
- Contiene texto, imágenes y otra información que puede ser leída por el ordenador.

- **Formato de Archivo**

-Manera particular de codificar la información para almacenarla en un archivo informático

- Dependiendo del tipo de datos a almacenar:
 - formatos de imagen (almacenar imágenes)
 - formatos de escritura (almacena texto)
 - formatos multimedia (audio y video)

2. Generación de documentos digitales.

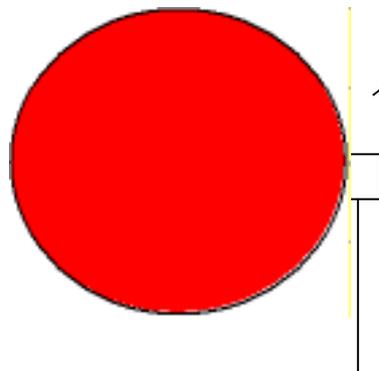


Imágenes.

- **Representación de Imágenes digitales:**

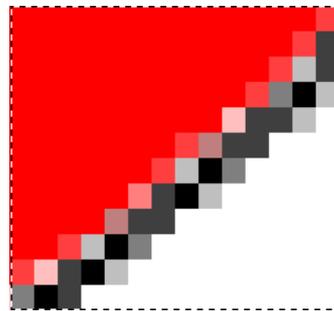
1) **Mapas de píxeles (o raster):** representación bidimensional de elementos de color que reproducen un objeto real o sintético

- 2) **Gráficos vectoriales:** formados por objetos geométricos independientes (círculos, arcos, polígonos..etc) definidos con una serie de atributos (forma, posición, color, profundidad..etc). Se describen las imágenes por los atributos que necesita el objeto geométrico en lugar de por sus píxeles.



Radio = 10
Centro = (0,0)
Color interior = rojo
Color línea = negro

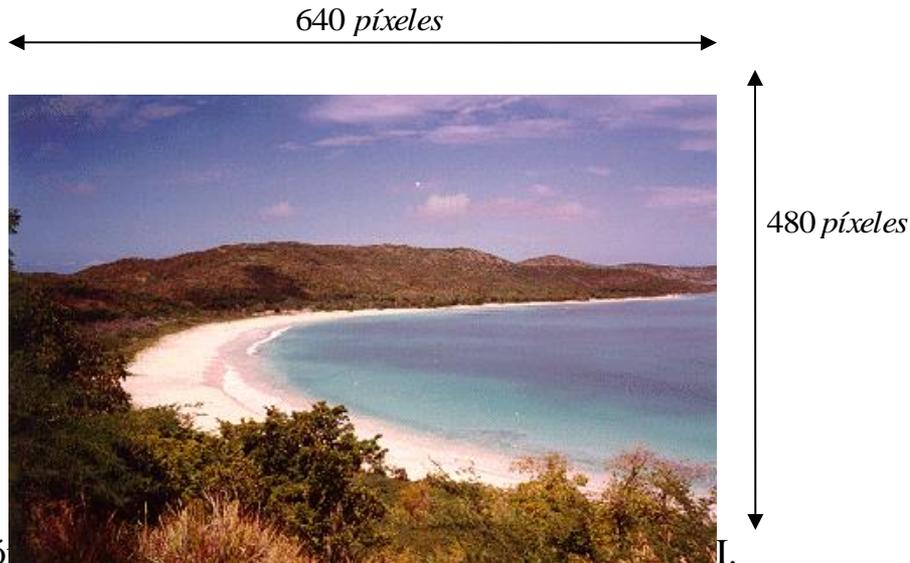
Círculo en un gráfico vectorial.



Círculo en un mapa de píxeles.

Mapa de Píxeles

- **Píxel**=picture element
 - Unidad mínima de una imagen
 - Almacena la información de color de una imagen
- **Mapa de píxeles**
 - Matrix de puntos llamados píxeles
 - Para cada punto(píxel) se almacena un color
- **Tamaño de imagen:** M x N píxeles --->ejemplo: imagen formada por 640 filas y 460 columnas tiene un tamaño 640x460 píxeles
- **Tamaño de archivo:** es la cantidad de información que contiene medida en bits o en algunos de sus múltiplos: bytes, megabytes, etc



Mapa de Píxeles

•Cámaras digitales o móviles: Megapíxeles

Un megapíxel, literalmente, equivale a un millón de píxeles. Por lo general, usamos esta medida para referirnos a la superficie de la imagen digital. De este modo, una fotografía digital que mida 3000 x 2000 píxeles, tiene 6.000.000 píxeles o, lo que es lo mismo, 6 megapíxeles.

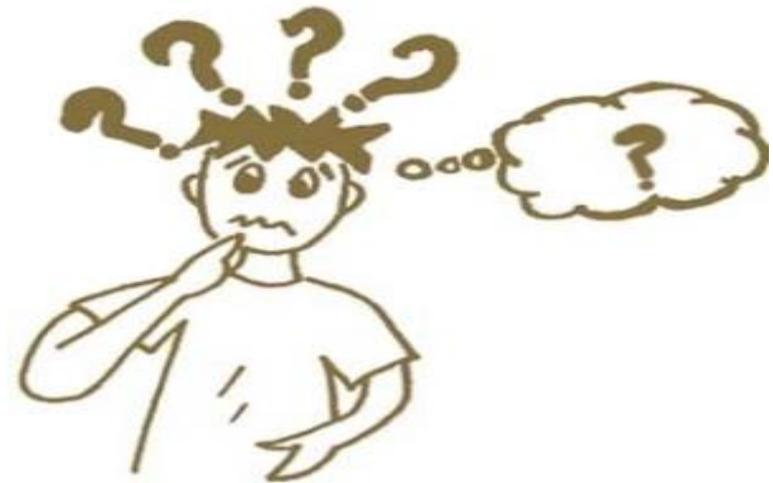
Problema: no da información acerca de las dimensiones de la imagen, sólo de la cantidad de píxeles que contiene

Ejemplo: una imagen de 800x500, otra 1000x400 ambas ocupan 0,4 megapíxeles

Mapa de Píxeles

Cuando digitalizamos una imagen hay dos parámetros importantes que deben ser considerados:

- Resolución espacial
- Profundidad de color



Mapas de píxeles. Resolución

Resolución Espacial hace referencia a:

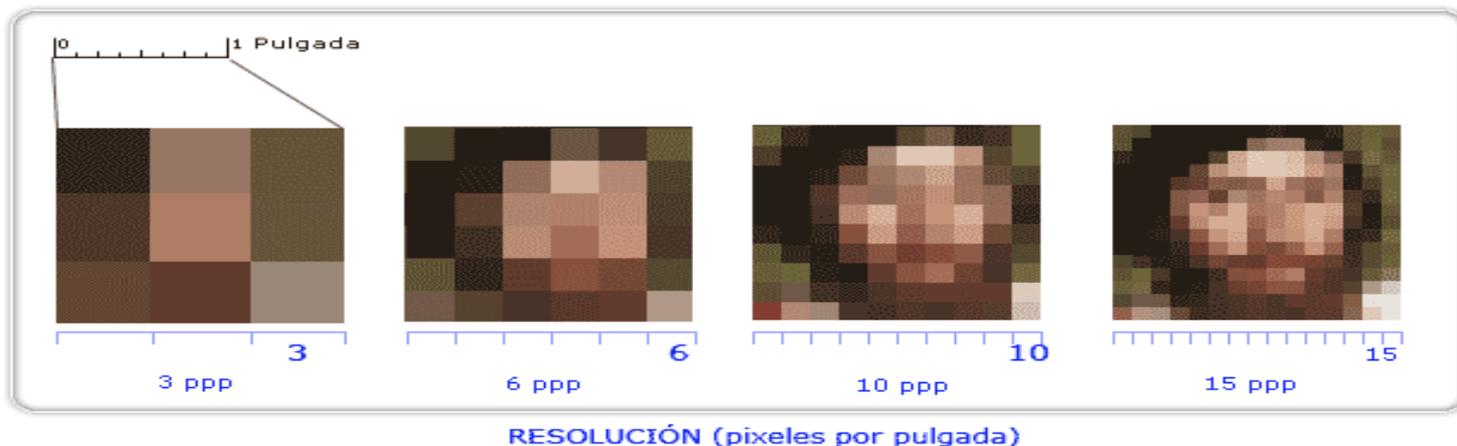
-Cantidad de puntos por unidad de longitud

-Número de puntos o pixels por pulgada (**dpi**, dot per inch, **ppp** puntos/pixels por pulgada, **Ppi** pixels per inch) -> impresora, escáner

-Pulgada=2,54 cm

-Indica la relación entre las medidas digitales y las físicas (mundo real).

Medidas físicas= Dimensiones digitales/resolución



Mapas de píxeles. Resolución

Resolución de entrada: es el número de píxeles que el dispositivo de entrada (escáner) **genera** por cada pulgada, sea papel o negativo.



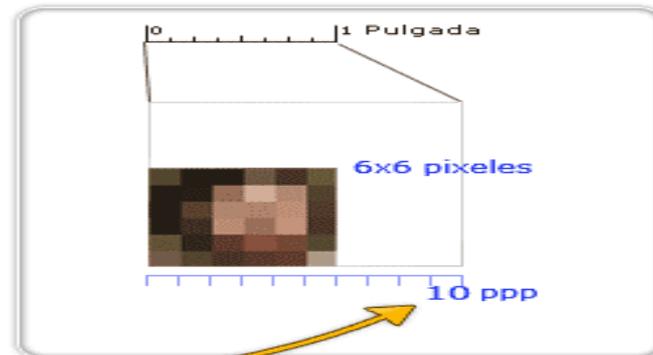
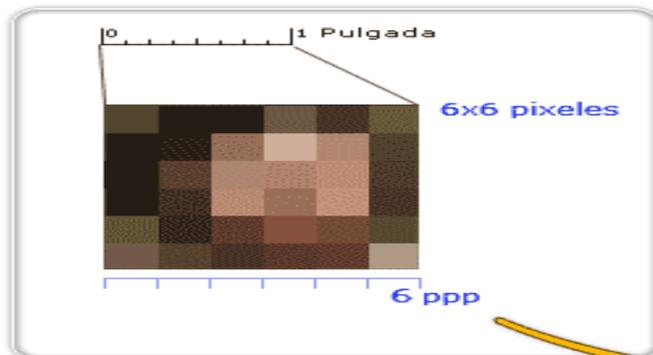
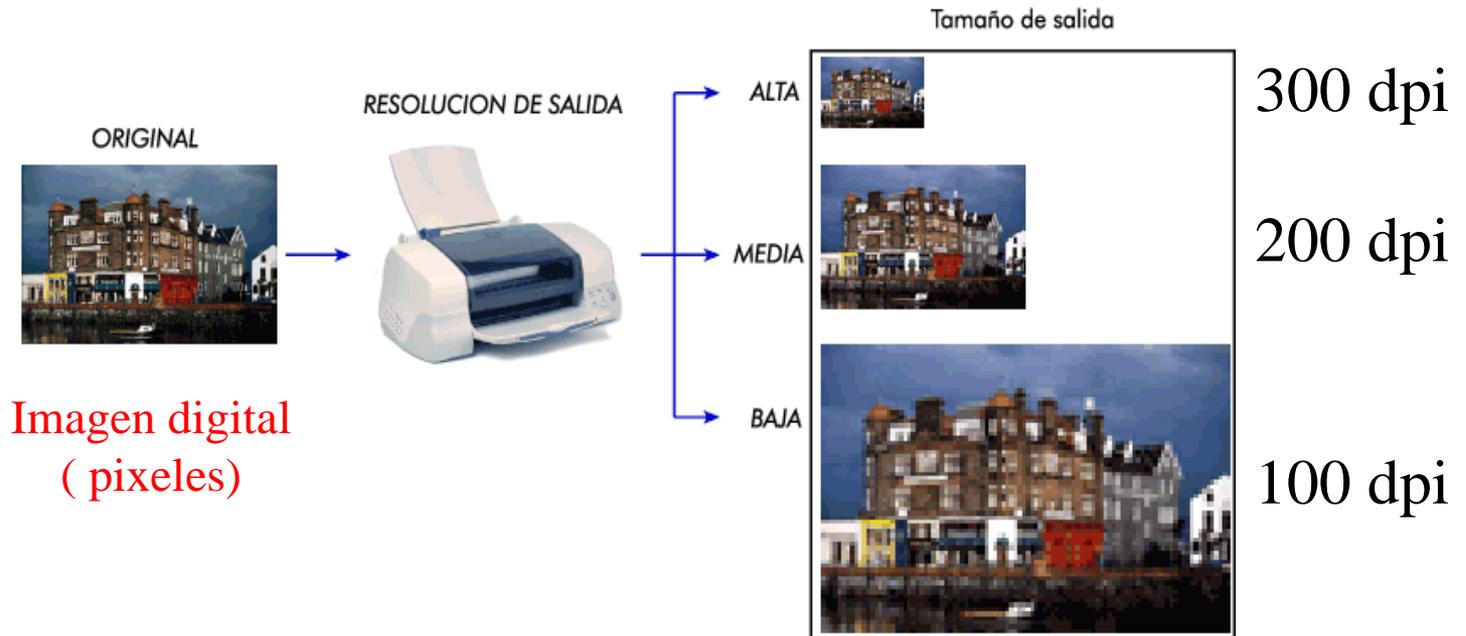
Papel (medidas físicas o reales: pulgadas o cm)

Imagen digital (píxeles)¹²

Mapas de píxeles. Resolución

Resolución de salida: es el número de píxeles por cada pulgada que el ordenador, envía a un dispositivo (como la impresora) para imprimir.

Papel (medidas físicas o reales: pulgadas o cm)



Mapas de píxeles. Resolución

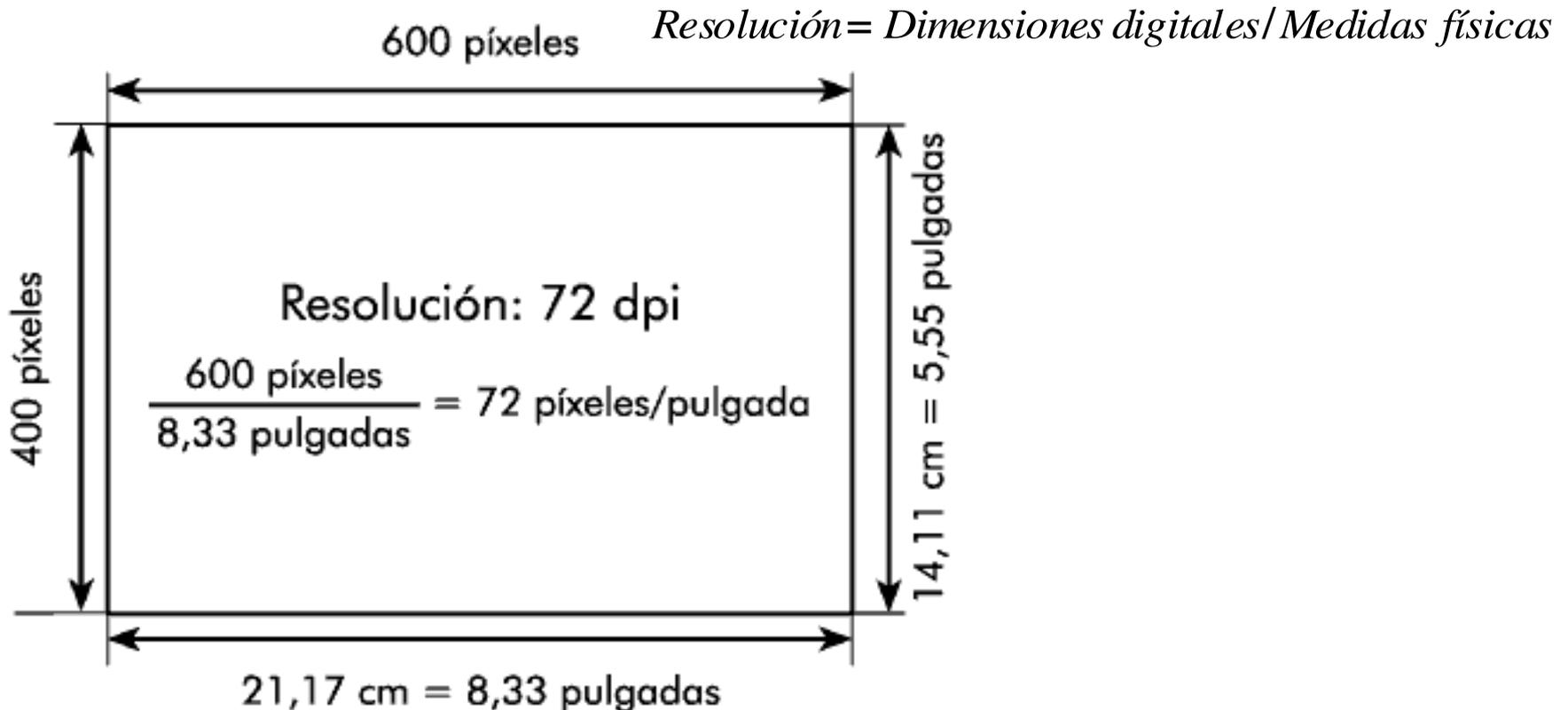
PROBLEMA

Si una imagen tiene unas dimensiones en píxeles de 600x400 y se imprime o está mostrando en pantalla con unas dimensiones de 8,33x5,55 pulgadas, ¿Cual es su resolución?

Medidas físicas= Dimensiones digitales / resolución

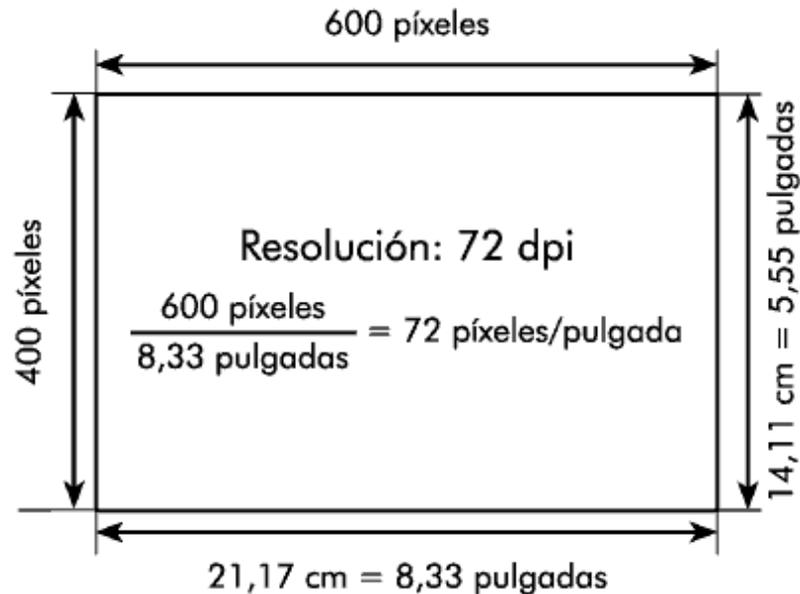
Mapas de píxeles. Resolución

Si una imagen tiene unas dimensiones en píxeles de 600 x400 y se imprime con unas dimensiones de 8,33x5,55 pulgadas, ¿Cual es su resolución? **Sol: 600/8,33pulgadas= 72 dpi.**



Mapas de píxeles. Resolución

PREGUNTA: Imagina ahora que se imprime esta imagen de 600x400 pixels) con una resolución superior, por ejemplo 200 ppp. ¿Cual será el tamaño físico(pulgadas) de la nueva imagen? ¿Igual, Menor, mayor?

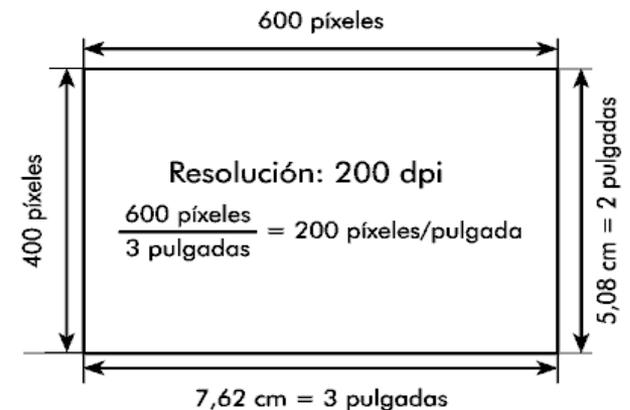
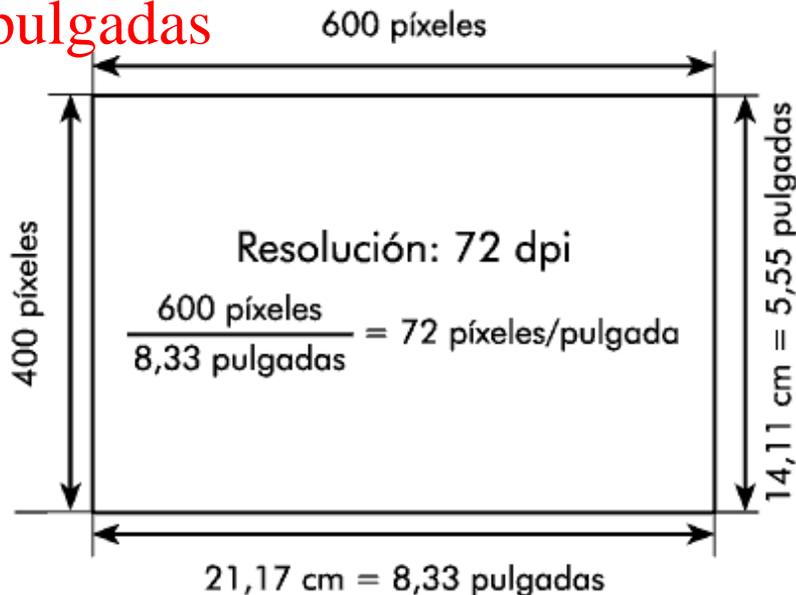


Mapas de píxeles. Resolución

SOLUCIÓN: Eso quiere decir que el ordenador le mandará imprimir la imagen a la impresora con un tamaño menor para así conseguir que haya 200 ppp y no 72 ppp de papel. Por tanto, la imagen impresa será más pequeña:

Solucion= $600\text{píxeles}/200\text{ dpi}=3\text{ pulgadas}$
 $400\text{Píxeles}/200\text{ dpi}=2\text{ pulgadas}$

Es decir en la imagen será más pequeña, se imprimirá en 3x2 pulgadas



Mapas de píxeles. Resolución

Recomendaciones

| Resolución | Destino |
|-------------------------|--|
| 72 ppp | <u>Monitor</u> Páginas web, multimedia, presentaciones power point <u>Impresión de Logos, lineas</u> |
| 150/200 ppp | <u>Impresión en general</u> Trabajaremos esta resolución para flyers, boletines, imágenes impresas en impresoras personales u otros trabajos en los que la calidad de imagen no es vital. |
| 300 ppp o superiores | <u>Impresión profesional de calidad óptima</u> |

Mapas de píxeles. Profundidad de color

Profundidad de Color o bits por píxel:

-Bits de codificación del color (o *profundidad de color*): número de bits que se utiliza para almacenar el color de un píxel → más bits por pixel, mayor calidad de la imagen.

-**Imagen B&N**-> 1 bit ($2^1=2$ tonos)

-**Imagen en escala de grises**-->1 byte=256 tonalidades de gris ($2^8=256$ tonalidades)

-**Color Verdadero (true Color)**: 24 bits (3 bytes) de Color (8 R 8G 8B)--->16 millones de colores distintos ($2^{24}=16$ millones)

Mapas de Píxeles. Profundidad de color



Distintas profundidades de color.

1 bit = Dos tonos, blanco y negro.

8 bits (1 byte) = 256 niveles de gris, imágenes de escala de grises.

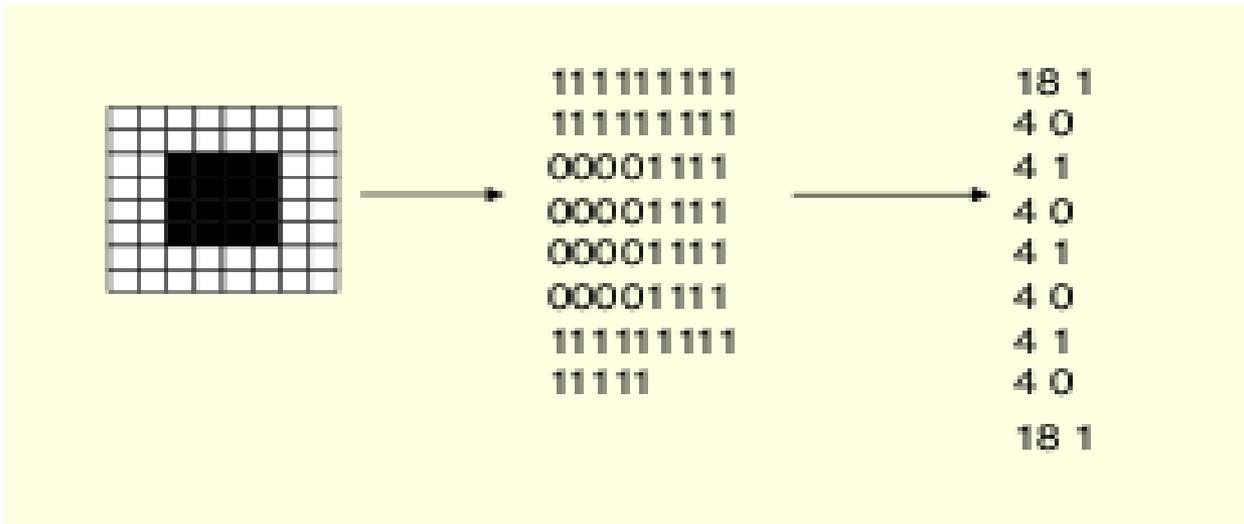
24 bits = 16 millones de colores, imágenes en color

Compresión

Compresión: Reducción de los tamaños de los ficheros (archivos) de imagen para ser procesados más fácilmente, almacenados o transmitidos con mayor velocidad a través de la red.

Las técnicas de compresión utilizadas se basan generalmente en:

- Abreviar información repetida
- Eliminar información que el OjO humano tiene dificultades para apreciar



Compresión

Tipos de compresión:

1) **Sin pérdidas (lossless)**: la imagen descomprimida será idéntica a la original, es decir no se ha eliminado ninguna información

-Algoritmos: RLE, LZW, Huffman, Delta-Encoding Lempel-Ziv-Welch

2) **Con pérdidas(lossy)**: la imagen descomprimida no es idéntica a la original, se ha perdido información.

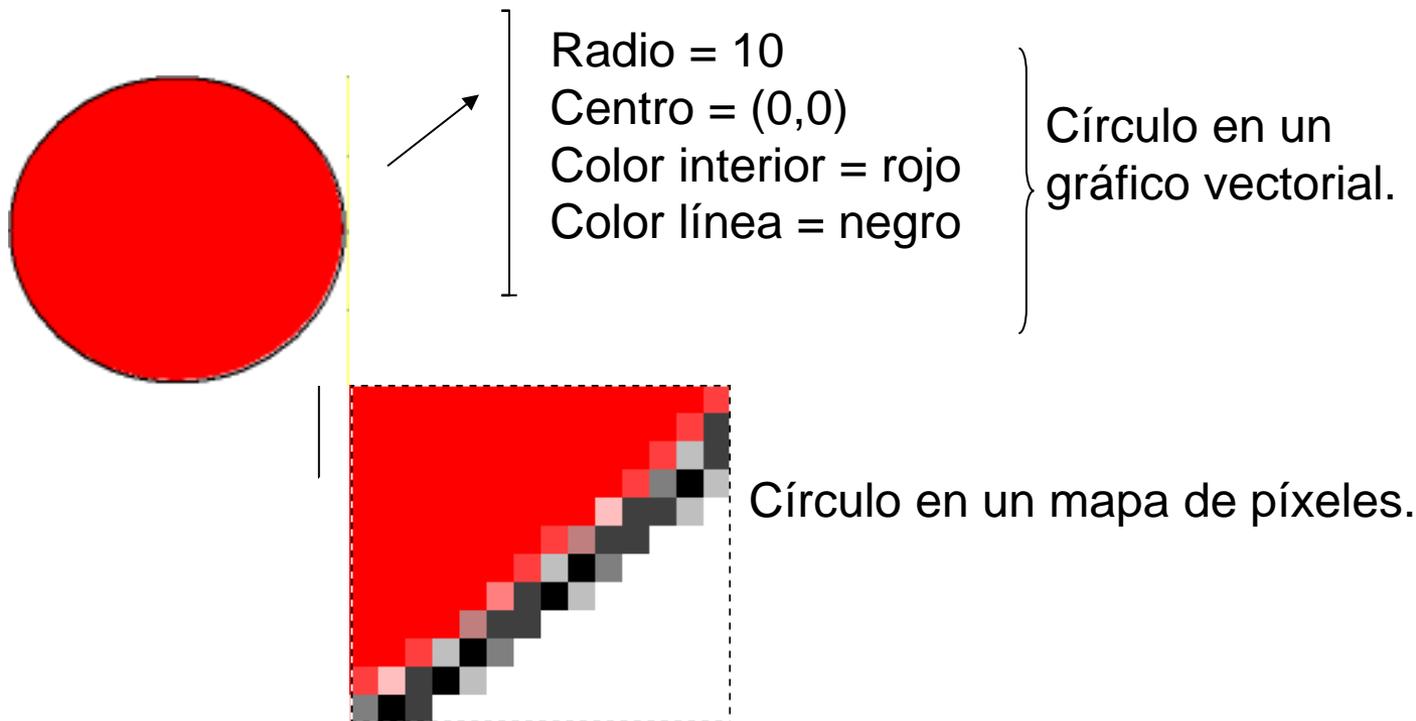
Algoritmos: Wavelets, fractales, JPEG

Ratio de Compresión: número de bits de la imagen original dividido por el número de bits de la imagen comprimida

Ejemplo: ratio 1:30 \rightarrow 30 de bits de la imagen original son representados por un bit en la imagen comprimida

Gráficos Vectoriales

- **Gráficos vectoriales:** Definen una imagen como un conjunto de objetos que son primitivas geométricas: rectas, circunferencias, elipses, curvas de Bezier..
- Es sencillo aplicarles transformaciones geométricas
- Mantienen su resolución frente al escalado.



Gráficos Vectoriales vs Mapa de bits

| Mapas de píxeles | Gráficos vectoriales |
|--|---|
| Almacenan información de Cada pixel | Descripción matemática |
| El tamaño del fichero es función del tamaño de la imagen | El tamaño es función del número de elementos de la imagen-->ocupan menos espacio en memoria |
| Se puede obtener a partir de la imagen original | Interpretados |
| Manipulación de la imagen basada en la manipulación de los píxeles | Transformaciones geométricas |
| Pérdida de calidad frente al escalado | Mantiene la calidad frente al escalado |

Vectorial o mapa de píxeles???

PREGUNTA??

- Dibujos arquitectónicos.?
- Gráficos de barras. ?
- Huellas dactilares.?
- Un mapa del mundo. ?
- Un scanner del cerebro.?
- Una reproducción de la mona lisa?.



MAPA DE PIXELS: TIFF, JPEG, BMP, PNG

| | | | | | | | |
|---------------------------------------|---------------|---|--|----|-----------------------|--|---|
| TIFF (Tagged Image File Format) | .tif .tiff | 1) sin pérdidas ó 2) con pérdidas | de grises; hasta 64 bits en imágenes de color | No | Propietario | Archivos originales para posteriormente se editados o impresos | disco comparado con JPEG. |
| GIF (Graphics Interchange Format) | .gif | Si, y esta es sin pérdidas | Escala de grises o color entre 1 a 8 bits | si | Propietario | Diagramas en internet. Rótulos, logotipos. Se usa mucha en Internet | Admite animaciones y transparencias |
| JPEG (Joint Photographs Expert Group) | .jpeg .jpg | Si y esta puede ser con: 1) pérdidas ó 2) con pérdidas | | no | Propietario | Ideal para guardar fotografías realistas con mucho detalle,. Se usa mucho en la Web | Permite utilizar grado de compresión variables. Pérdidas de resolución durante la compresión. |
| BMP (Windows Bitmap) | .bmp | Si admite compresión, pero el ratio de compresión es muy bajo->ocupa mucho espacio en memoria | 1, 4, 8 y 24 bits imágenes de color | no | Propietario (windows) | Se usa cuando ocupa poco espacio-->no se usa mucho en la Web | “Pesa mucho”. Formato simple que mantiene un gran información de la imagen. |
| PNG (Portable Network Graphics) | .png | Si admite y esta es sin pérdidas | 1, 2, 4, 16 bits escala de grises; 64 bits en imágenes de color | no | Libre | Ideal para imágenes formadas por grandes áreas de color plano o con pocas variaciones de color (gráficos,dibujos...etc) Se usa mucho en la Web | -Comprime bien los dibujos y gráficos -Permite transparencias |



GRAFICOS VECTORIALES: VML, SVF, EPS

| | | | /Propietario | | |
|-----------------------------------|------------------------------------|---|-----------------------------|--|--|
| VML (vector Markup lenguaje) | No admite | Permite mostrar animaciones vectoriales | Libre | -Se integra en el código html, javascript -->permite inclusión de imágenes en la Web -Ya no se utiliza mucho porque muchos navegadores no lo admiten -Microsoft lo sigue usando | -Lenguaje basada en xml. -Tiene ciertas características que permite dibujar como si estuviésemos trabajando con mapas de bits |
| SVG (scalable Vector graphics) | Con Gzip y pasa a ser llamado SVGZ | Permite mostrar animaciones vectoriales | Libre | Para la inclusión de imágenes en páginas Web, aplicaciones de teléfonos móviles y cartografía | -Requiere menor espacio en disco que un bitmap -Pueden ser buscadas e indexadas en bases de datos |
| EPS (encapsulated postscript) | Comprime imágenes a mapa de bits | No admite animación | Propietario: Compañía Adobe | -Muy utilizado en la impresión profesional -Intercambio de archivos entre programas de maquetación (autoedición) | -Puede contener una combinación de texto e imágenes -Puede guardar mapas de bits -Compatible con muchos programas de autoedición |

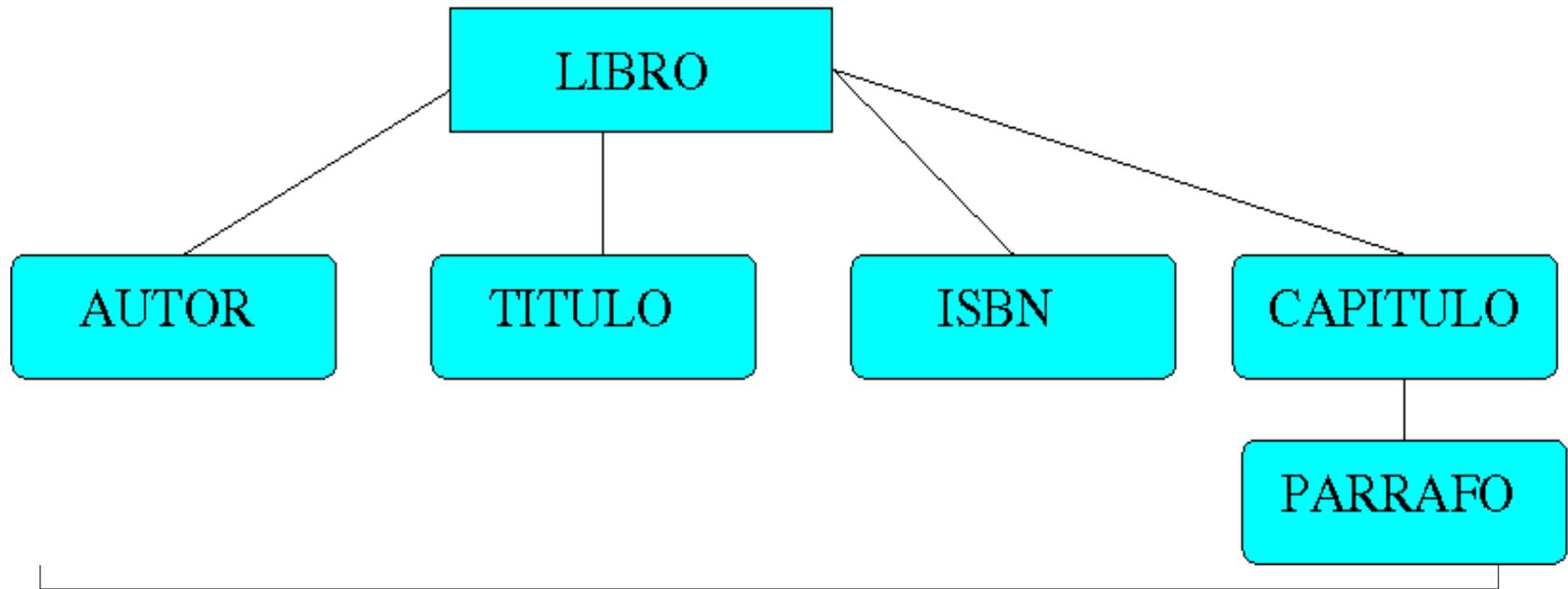
Formatos de Escritura

Los formatos de escritura pueden ser:

- **1) No estructurados:** orientados a la presentación final
- **2) Estructurados:**
 - dan énfasis a la estructura lógica del documento, en vez de la forma como es representado
 - se utilizan para definir tipos de estructuras de documentos
 - normalmente usan etiquetas para marcar las partes (estructuras) de un texto

Ventajas: -separa contenido de la presentación final
-reutilizables a largo plazo

Formatos de Escritura. Estructurados



Ejemplo: para definir un libro tendríamos uno o varios elementos autores, un elemento título, un elemento isbn, uno o varios elementos capítulos y uno o varios elementos párrafo

Formatos de Escritura. Estructurados

<libro>

<autor>

Arturo Pérez Reverte

<autor/>

<titulo>

La reina del Sur

<titulo/>

<isbn>

978-84-204-7198-3

<isbn/>

<capitulo>

<parrafo>

Estábamos en sentados en un bar cuando de repente....

<parrafo/>

<capitulo/>

<libro/>

Formatos de Escritura: No Estructurados

| | Libre | Propiet. | Mod. | No Mod. | Característica principal | Ventaja ppal. | Inconveniente |
|--------------|------------|----------|------|---------|--|----------------------------|-----------------------|
| ASCII | x | | x | | Código de texto plano representados en 7-8 bit. | Incluido en estandar UTF-8 | Límite de caracteres |
| DOC | (Software) | x | x | | Formato de documentos de carácter ofimático | Sencillo y extendido | No es multiplataforma |
| ODF | x | | x | | Formato de documentos de carácter ofimático | Multiplataforma | No estandar |
| PDF | | x | | x | Representa el doc. Original sin pérdida de información | Multiplataforma | No modificable |

Formatos de Escritura: Estructurados

| | Libre | Prop. | Mod. | No mod. | Característica Principal | Ventaja principal | Inconveniente |
|-------------|----------------------|-------|------|---------|---|---|--|
| SGML | x | | x | | Metalinguaje complejo. | Portabilidad | Muy complejo |
| XML | x | | x | | Metalinguaje derivado de SGML. | Separación de estruct / contenido. Etiquetas ilimitadas | Uso necesario de XSL o CSS |
| HTML | x | | x | | Hipertextualidad | Sencillo para el usuario | Etiquetas limitadas y no extensibles |
| PS | | x | x | | Lenguaje de descripción de páginas. | Alta calidad de impresión | Propietario. |
| TeX | x | | x | | Documentos para el ámbito científico | Exportar a otros formatos. Facilidad introducir ecuaciones, imágenes, índice | Complejidad aprendizaje / Paquetes auxiliares |
| RTF | (formato y software) | x | x | | Compatibilidad entre versiones de Microsoft | Codifica texto e imagen | No compatible con otros SSOO |

Preservación de Documentos Digitales:

- Entendemos por preservación digital todo el proceso y el conjunto de tareas que se llevan a cabo para proteger los documentos frente a posibles pérdidas, para conservarlos por su gran valor patrimonial y para poder ofrecerlos a los usuarios desde lugares remotos.
- La preservación es una cuestión importante en el ciclo de vida de la documentación, así que debemos tenerla en cuenta desde la creación de los propios documentos. Por eso todas las unidades de información documental están creando políticas y estrategias para cumplir los objetivos.

Preservación de Documentos Digitales:

Las principales estrategias para la preservación de documentos digitales son:

- **Conservación**
- **Emulación**
- **Encapsulado**
- **Migración**

Preservación de Documentos Digitales:

- **Conservación:** consiste en conservar adecuadamente el soporte en el que se encuentra la información es decir, conservar adecuadamente los sistemas informáticos. Este proceso implica la congelación del software y del hardware de los diferentes ordenadores en un momento determinado

- **Desventajas:**

- Elevado coste
- La conservación debe realizarse desde la creación del documento
- Escasa utilización
- Estrategia para uso a corto o medio plazo

Preservación de Documentos Digitales:

Emulación: consiste en recrear la apariencia y el funcionamiento original de un documento digital en el software original en el que fue creado. De esta manera, el ordenador más moderno podrá emular tanto el sistema operativo obsoleto como los programas que utilizaba.

•Desventajas:

-Puede ser un proceso complejo y no conseguir el efecto deseado

Preservación de Documentos Digitales:

Encapsulado: El encapsulado consiste en conservar dentro de un mismo paquete y, junto con el código fuente del documento, la cadena de bits del software y del sistema que permita reproducirlo.

Para ello hay que encapsular tres elementos de información:

1. El documento y su entorno de software. Esto consiste en el código fuente del documento que hay que conservar, más los datos del sistema operativo que permitía que se ejecutase dicho software.
2. El emulador de la plataforma original. Esto no quiere decir que este el ejecutable, solo la especificación de los atributos necesarios para que el emulador pueda recrear el comportamiento del documento y del software originales.
3. Una explicación, que sea comprensible para todos, en la que se explique el software y el hardware emulados, los metadatos, su ciclo de vida, y contexto de creación

Preservación de Documentos Digitales:

Migración : Método encargado de convertir los documentos almacenados a nuevos formatos existentes con el objetivo de no perder la información que contienen.

-Se utiliza para copiar o convertir datos desde una tecnología a otra, tanto si se trata de hardware como de software, conservando las características esenciales de los datos.

-El propósito de la migración es preservar la integridad de los objetos digitales y mantener la posibilidad por parte de los usuarios de recuperar, visualizar y utilizarlos en una perspectiva de constante cambio tecnológico.

Bibliografía:

- “La biblioteca digital”* E. García Camarero, L.A. García, Ed. Arco/Libros, 2001
- “Los documentos electrónicos, Qué son y cómo se tratan”*,
Jordi Serra Serra, Ediciones Trea, 2008
- “La fotografía digital en los archivos, Qué es y cómo se trata”*, David Iglesias Franch, Ediciones Trea, 2008.
- La edición digital*, Javier Barroso, apuntes Universidad Politécnica de València
- *Lectores electrónicos*, v. 17, n.4 ,E.Monteoliva-García, C. Pérez-Ortiz, Rafa Repiso-Caballero, 2008 revista “el profesional de la información”,
www.elprofesionaldelainformacion.com.
- Acceso abierto a la información*, Remedios Melero, v. 14, n.5 , 2008 revista “el profesional de la información”, www.elprofesionaldelainformacion.com.
- Las licencias creative Commons una alternativa al copyright*, Raquel Xalabarder, apuntes de la Universitat Oberta de Catalunya.

Temario grupos del trabajo en casa:

- **1) Formatos de Escritura (estructurados vs no estructurados)**
- **2) Formatos de Imágenes (mapas de píxeles y vectoriales)**
- **3) Aspectos Legales de la edición electrónica**
- **4) Dispositivos de Lectura, e-book readers: principios y formatos**
- **5) Preservación de Documentos digitales**