

# Proyecto Elba

## **Procedimientos y normas generales usados en la edición del libro electrónico y en su almacenamiento en la biblioteca digitales**

Carolina García Cataño

Dabne - Tecnologías de la información

<http://dabne.net>

# Proyecto Elba

Dabne – Tecnologías de la información

## Índice

- 1- Algunos procedimientos y normas para la creación de objetos digitales. (*digitalización*)
- 2- Algunas características y normas de las publicaciones digitales y de los repositorios digitales. (*edición y almacenamiento*)
- 3- Algunas ideas sobre la estructura y funciones de las Webs general y locales. (*difusión*)

# Proyecto Elba

Dabne – Tecnologías de la información

## 1.- Algunos procedimientos y normas para la creación de objetos digitales.

El objetivo de la digitalización es convertir la información (en forma de textos, imágenes fija, música y voz, video, etc,) contenida en diversos soportes, a un soporte electrónico en el que información se codifica en forma numérica

# Proyecto Elba

Dabne – Tecnologías de la información

1.- Algunos procedimientos y normas para la creación de objetos digitales

## El código ASCII

ASCII es un código de caracteres basado en el alfabeto latino (inicialmente solo mayúsculas), creado por ASA (American Standards Association) en 1963, como una refundición y evolución de los conjuntos de códigos utilizados entonces en telegrafía. Más tarde, en 1967, se incluyen las minúsculas y se redefinen algunos códigos de control para formar el código.

# Proyecto Elba

Dabne – Tecnologías de la información

1.- Algunos procedimientos y normas para la creación de objetos digitales

## Ejemplo de ASCII

Binary	Decimal	Hex	Graphic
0100 0000	64	40	@
0100 0001	65	41	A
0100 0010	66	42	B
0100 0011	67	43	C
0100 0100	68	44	D

# Proyecto Elba

Dabne – Tecnologías de la información

1.- Algunos procedimientos y normas para la creación de objetos digitales

## Objeto digital

*objeto digital*, ítem considerado como un todo coherente, que es el equivalente del objeto físico, libro, folleto, imagen o audio del que proceden. Por ejemplo, un artículo, un libro, etc.

*objeto digital secundario*: en ciertos casos, conviene considerar cada subdivisión coherente de un objeto digital, por ejemplo en un concierto cada pieza de un autor

# Proyecto Elba

Dabne – Tecnologías de la información

1.- Algunos procedimientos y normas para la creación de objetos digitales

## Proceso general de creación de objetos digitales

- 1.- Selección del material a digitalizar
- 2.- Digitalización
- 3.- Adecuación- OCR
- 4.- Corrección
- 5.- Almacenamiento en el depósito

# Proyecto Elba

Dabne – Tecnologías de la información

1.- Algunos procedimientos y normas para la creación de objetos digitales

## Digitalización de textos

a) *Escaneo*: Consiste en obtener una imagen (facsimil) del texto en formato digital, es decir como una secuencia binaria.

El *hardware* consiste en un scanner,

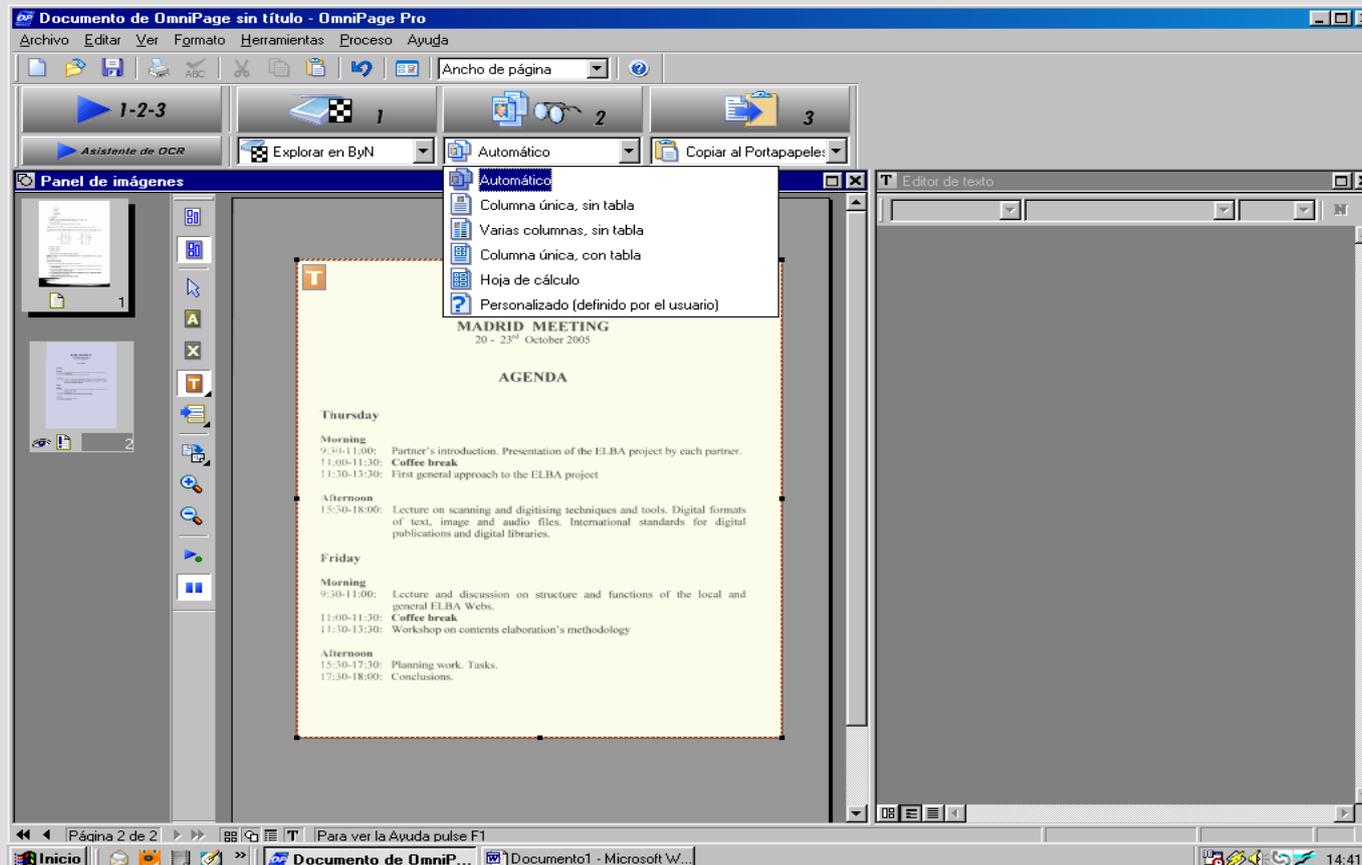
También existen diversos *software* necesarios para la digitalización, es de muy sencilla operación, aprende con uso a reconocer tipos especiales y el estilo de los textos que se están digitalizando.

# Proyecto Elba

## Dabne – Tecnologías de la información

1.- Algunos procedimientos y normas para la creación de objetos digitales

- Imagen del aspecto gráfico del software de digitalización



# Proyecto Elba

Dabne – Tecnologías de la información

1.- Algunos procedimientos y normas para la creación de objetos digitales

Formatos de imagen al realizar el escaneo

*TIFF Tagged Image File Format* (formato de archivo de imágenes con etiquetas) archiva información sobre las características de la imagen, que sirve para su tratamiento posterior.

*JPEG, (Joint Photographic Experts Group)* es un algoritmo diseñado para comprimir imágenes. JPEG es también el formato de fichero que utiliza este algoritmo para comprimir imágenes.

*PNG, (PNetwork Graphics)* es un formato gráfico basado en un algoritmo de compresión sin pérdida para bitmaps no sujeto a patentes.

# Proyecto Elba

Dabne – Tecnologías de la información

1.- Algunos procedimientos y normas para la creación de objetos digitales

## *Reconocimiento óptico de caracteres (OCR):*

Haciendo uso de un software se traducirán las imágenes de un texto escrito a un texto que pueda ser editado por el ordenador.

Para hacerlo de una manera adecuada, en los programas de OCR deberemos especificar en las imágenes que parte es texto, que parte es imagen y así desechar cualquier mancha

# Proyecto Elba

Dabne – Tecnologías de la información

1.- Algunos procedimientos y normas para la creación de objetos digitales

## *Corrección del OCR*

El software no reconoce los caracteres de forma óptima, por tanto es necesario revisar los textos y corregir o comprobar que es idéntico al original

# Proyecto Elba

Dabne – Tecnologías de la información

## 2.- Edición y almacenamiento

Otras de las tareas que han de realizarse con objetos digitales son las relativas a su edición en la forma en que queremos visualizar nuestros textos y al almacenamiento ordenado de los distintos materiales asociados con los objetos digitales.

Estos materiales cumplen diversas funciones y en cada caso se usaran los formatos mas adecuados a ellas, que también deberán tenerse en cuenta en la organización y forma de almacenarse los objetos digitales.

# Proyecto Elba

Dabne – Tecnologías de la información

## 2.- Edición y almacenamiento

En la edición electrónica queremos conservar los textos y mostrarlos con toda la riqueza tipográfica de los originales de que proceden debemos usar otros formatos, en general llamados de marcas, etiquetas o marcadores, con los que asignamos algunos calificativos para describir las diferentes peculiaridades, tipográficas u otras, de los textos

# Proyecto Elba

Dabne – Tecnologías de la información

## 2.- Edición y almacenamiento

- **txt** es la extensión de un documento en texto plano o plain text, se refiere a datos textuales en formato ASCII. Este formato es el más portable ya que es compatible con prácticamente cualquier ordenador y cualquier aplicación.
- **rtf** es el acrónimo de Rich Text Format, lenguaje de descripción desarrollado por Microsoft que usa marcas para enriquecer el aspecto tipográfico de los documentos. Es un pobre formato estandarizado con diversas incompatibilidades no es recomendable su uso para guardar documentación.

# Proyecto Elba

Dabne – Tecnologías de la información

## 2.- Edición y almacenamiento

- **html** (Hypertext Markup Language) es una especie de documento de texto que es usado por los navegadores para visualizar textos e imágenes. Los documentos HTML suelen ser llamados páginas web.
- El formato **pdf** (Portable Document Format), ideado por Adobe, es la herramienta pública empleada en empresas con estándares mundiales para una distribución e intercambio seguros y fiables de documentos electrónicos. En casos de uso de software como ABBYY es interesante el uso de una opción que es pdf con texto bajo la imagen.

# Proyecto Elba

Dabne – Tecnologías de la información

2.- Edición y almacenamiento

## Repositorio

Un *Repositorio* de objetos digitales, y en particular los usados en las Bibliotecas Digitales, es el lugar electrónico donde se alojan los objetos digitales que corresponden a los diversos materiales bibliográficos que se vayan incorporando a dicha biblioteca para posterior consulta privada o pública.

# Proyecto Elba

Dabne – Tecnologías de la información

2.- Edición y almacenamiento

## Repositorio

Organización de un repositorio

- Establecimiento categorías
- Creación directorios correspondientes a cada categoría
- Nomenclatura de los objetos digitales

articulo-xxxxx.pdf

libro-xxxxx-pdf

imagen-xxxx.jpg

audio-xxx.wav

video-xxxx.mpeg

# Proyecto Elba

Dabne – Tecnologías de la información

2.- Edición y almacenamiento

## Deposito

*Deposito*, lugar electrónico donde se almacenarán, parte de los materiales pueden sufrir alguna modificación o si no cumplen todos los requisitos quedarán “congelados” antes de pasar al repositorio. Reproducirá la misma estructura que el repositorio

# Proyecto Elba

Dabne – Tecnologías de la información

## Estándares para la descripción de los recursos digitales

Posibilidades de recuperación de los materiales

- Boca a oreja

Difusión a través de los medios tradicionales (publicación en boletines impresos, ruedas de prensa, etc.)

- Acceso público

Recuperación actual : Google, yahoo, etc.

Recuperación futura: web semántica, *metadatos*

# Proyecto Elba

Dabne – Tecnologías de la información

Estándares para la descripción de los recursos digitales

## Metadatos

Información descriptiva sobre las peculiaridades o atributos de cada objeto digitalizado

Los metadatos proporcionarán información sobre la localización del objeto digital, los formatos en los que se encuentra, materia a la que pertenece, autor, lengua

Su función es facilitar al usuario final, el descubrimiento y navegación por la información



# Proyecto Elba

Dabne – Tecnologías de la información

Estándares para la descripción de los recursos digitales

## **¿por qué usar metadatos?**

Algunos de los motivos serían:

Realizar intercambio de información sobre recursos existentes

Evitar repetición de objetos digitales

Localización de los recursos por parte de agentes, robots,  
etc(inteligencia artificial)

Permiten la optimización de las búsquedas

# Proyecto Elba

Dabne – Tecnologías de la información

Estándares para la descripción de los recursos digitales

## ¿qué metadatos usar?

### **Dublin Core (<http://dublincore.org>)**

- Simplicidad, pensado para que pueda ser utilizado tanto por bibliotecarios como por cualquier autor que desee describir sus documentos y aumentar su visibilidad.
- Consenso internacional en el número y definición de los elementos.
- Flexibilidad, nada en el DC es obligatorio, todos los elementos son opcionales y repetibles, así el usuario elige la profundidad de una descripción.

# Proyecto Elba

Dabne – Tecnologías de la información

## Acceso a la recursos digitales

Una vez creado un conjunto de objetos digitales, necesitamos una serie de herramientas e ideas de cómo configurar y definir los sitios web donde se harán visibles los materiales digitalizados

# Proyecto Elba

Dabne – Tecnologías de la información

**Acceso a la recursos digitales**

## **Formas de dar acceso**

- **A nivel local : Intranet**
  - Existencia de servidores destinados al uso interno de los recursos digitales
  - Uso del navegador como interfaz de acceso al material
  - Creación de un sitio web para acceder a los recursos
- **A nivel global: Internet:**
  - Existencia de un servidor de acceso público a los recursos digitales.
  - Uso del navegador como interfaz de acceso al material
  - Creación de un sitio web para acceder a los recursos

# Proyecto Elba

Dabne – Tecnologías de la información

Acceso a la recursos digitales

## ¿Qué tipo de sitio web crear?

- Basado en páginas web estáticas
- Basado en páginas web dinámicas
- Basado en herramientas existentes

# Proyecto Elba

Dabne – Tecnologías de la información

## Acceso a la recursos digitales

- **Páginas web estáticas**

Creación del acceso a los recursos de forma manual

Necesario tener en cuenta la distinción entre contenido (XHTML) y la representación de contenido (CSS)

**XHTML**, eXtensible Hyper Text Markup Language, es el lenguaje de marcado pensado para sustituir a HTML como estándar para las páginas web, tiene, básicamente las mismas funcionalidades, pero cumple las especificaciones más estrictas.

**CSS**, hojas de estilo en cascada (Cascading Style Sheets) son un lenguaje formal usado para definir la presentación de un documento estructurado escrito en HTML

# Proyecto Elba

Dabne – Tecnologías de la información

**Acceso a la recursos digitales**

- **Páginas web dinámicas**

*Elementos necesarios:*

Base de datos, lenguaje de programación (script) y css

*Características:*

El contenido de la página está almacenado en una base de datos

Lenguaje de programación realiza consultas a la base de datos

Se genera al vuelo el html que luego se muestra en el navegador.

# Proyecto Elba

Dabne – Tecnologías de la información

## Acceso a la recursos digitales

- **Páginas web dinámicas**

*Las páginas web dinámicas pueden ser de creación propia pero nos parece más interesante hacer uso de herramientas existentes: SPIP (système de publication pour l'Internet)*

A nivel gráfico es un software muy maleable

A nivel de inclusión de contenidos muy sencillo, NO es necesario ningún conocimiento previo de html, css, programación y sin embargo posee muchas funcionalidades. Existe un interfaz de administración donde aparecen campos de texto donde rellenar, el título de la obra, la descripción y el texto,etc.

# Proyecto Elba

Dabne – Tecnologías de la información

**Acceso a la recursos digitales**

## **Accesibilidad (WAI)**

La accesibilidad Web significa que personas con algún tipo de discapacidad van a poder hacer uso de la Web. En concreto, al hablar de accesibilidad Web se está haciendo referencia a un diseño Web que va a permitir que estas personas puedan percibir, entender, navegar e interactuar de forma efectiva con la Web, así como crear y aportar contenido.

# Proyecto Elba

Dabne – Tecnologías de la información

**Acceso a la recursos digitales**

## **Licencias**

Uno de los grandes escollos que nos encontramos a la hora de realizar una libre difusión de la cultura y de la información en general, es el tema de los derechos de autor, el copyright.

# Proyecto Elba

Dabne – Tecnologías de la información

Acceso a la recursos digitales

## Licencias

Los materiales de libre difusión han de tener una de estas licencias:

**Dominio público:** materiales se encontrarán en dominio público, es decir, que han pasado 70 años después de la muerte del autor

**Creative Commons:** algunos derechos reservados  
<http://creativecommons.org>

# Proyecto Elba

Dabne – Tecnologías de la información

**Gracias a todos :-)**

**Dudas y preguntas**

<http://dabne.net>

[dabne@dabne.net](mailto:dabne@dabne.net)