

# *Digitalización y archivos*

AURELIO PROL CASTRO

## **1. Introducción**

La tecnología digital abre una perspectiva totalmente nueva. La web mantiene millones de sitios web, e Internet es el lugar para la investigación, la enseñanza, la expresión, la publicación y la comunicación de la información. Las bibliotecas y los archivos son proveedores de información primaria para la sociedad y fueron usuarios de la nueva tecnología digital primero en relación con la catalogación y la gestión de los procesos, y más tarde para proporcionar información sobre sus colecciones a la comunidad de la web. Además de preservar y proporcionar acceso al “material digital”, en la actualidad un gran número de archivos y bibliotecas han iniciado también la creación de copias digitales de sus recursos ya existentes (microfilm). Para dichas bibliotecas y archivos se han redactado estas directrices. (IFLA e ICA para la UNESCO)

Estas son Directrices para emprender proyectos de digitalización de colecciones y fondos, en particular para aquellas colecciones custodiadas en bibliotecas y archivos. Tratan del patrimonio documental en papel, manuscritos, libros impresos y fotografías, y en los tiempos en que estamos se pueden englobar los registros sonoros y de películas, objetos o monumentos. Hacen referencia a la planificación y establecimiento de proyectos, es decir, a la selección, gestión y procesos de producción implicados en dichos proyectos en actividades bien definidas.

Muchas bibliotecas y archivos desearían planificar proyectos de digitalización pero carecen de experiencia. Es necesaria una guía práctica como herramienta de trabajo para la planificación de proyectos de digitalización.

### 1.1. ¿Por qué digitalizar?

Las razones de un proyecto de digitalización, o más exactamente de la conversión digital de documentos son variadas y pueden solaparse. La decisión de digitalizar puede tomarse con objeto de:

- Incrementar el acceso: esta es razón principal y la más obvia, cuando se sabe que hay una alta demanda por parte de los usuarios y la biblioteca o el archivo desean mejorar el acceso a una determinada colección, difusión vía WEB.
- Mejorar los servicios para un grupo creciente de usuarios proporcionando un acceso de mayor calidad a los recursos de la institución en relación con la educación y formación continua, posibilidad de ver el mismo documento por más de un usuario.
- Reducir la manipulación y el uso de materiales originales frágiles o utilizados intensivamente y crear una “copia de seguridad” para el material deteriorado como libros o documentos quebradizos.
- Impulsar el desarrollo de recursos cooperativos, compartiendo intereses comunes con otras instituciones para crear colecciones virtuales e incrementarlas.

Debe estar claro por qué se embarca uno en un proyecto de digitalización: el objetivo determinará el proceso y los costes. Puesto que la digitalización supone un trabajo intenso y caro es importante capturar una imagen de modo que sea posible utilizarla para satisfacer diferentes necesidades.

La base para emprender la digitalización es el reconocimiento de que la tecnología cambiará y que cambiará con frecuencia. Por lo tanto la decisión crucial no tiene que ver con “cuando” comenzar o “si” comenzar. Más bien es cuestión de si la institución puede permitirse la oportunidad de conseguir divulgar en una comunidad global, de la manera que permite la tecnología para mejorar el acceso y la preservación de los recursos culturales y de investigación.

La digitalización será una tarea costosa, que requiere una planificación detallada y el establecimiento de una infraestructura que asegure el acceso continuo a los ficheros digitalizados.

El usuario juega un papel importante en la decisión de poner en marcha un proyecto, pero con mucha frecuencia es difícil definir cuál es ese papel. De hecho las necesidades específicas de los usuarios pueden ser difíciles de conocer.

En la mayoría de los casos existe un supuesto grupo de usuarios, y el objetivo de la institución es incrementar sus servicios y extender su enfoque e influencia. El grupo de usuarios puede diferir, dependiendo del tipo de institución y de la misión de la organización. Las instituciones de educación superior satisfacen las necesidades de su personal docente y de los estudiantes. Las instituciones públicas y nacionales deben satisfacer a una población mayor y más diversificada. Esto influye no sólo en la selección sino también en las formas de presentación y acceso (la interfaz del usuario).

Digitalización no es lo mismo que preservación: la digitalización no es más barata, más segura, ni más fiable que la microfilmación. A diferencia de un fotograma de un microfilm de alta calidad, una imagen digital no es un master de preservación. La única manera en la que la conversión digital contribuye positivamente a la preservación es cuando la copia digital reduce el desgaste y la rotura del original, o cuando los ficheros están escritos como salida de ordenador en microfilm, que cumple los estándares de preservación de calidad y perdurabilidad. Por lo tanto, un proyecto de digitalización no sustituye a un programa de preservación basado en la conversión a microfilm (o de desacidificación, de un tratamiento de conservación o de mejora de las condiciones de almacenamiento).

En general todo ello es cierto. Pero puede haber circunstancias específicas, por ejemplo en los países en vías de desarrollo, que pueden cambiar la base de este planteamiento. Si una institución sin ninguna experiencia ni facilidad para la conservación quiere preservar una determinada colección, puede decidir invertir en equipos de digitalización en lugar de en equipos de microfilmación, evitando de este modo el elevado gasto en cámaras y procesadores de microfilmación y reconociendo que este equipamiento digital y el desarrollo de los conocimientos de su personal satisfarán igualmente otros objetivos. Este cambio de la microfilmación como método generalmente recomendado de preservación por el de la digitalización con sus riesgos a largo plazo no es quizás la solución ideal para el problema de la descomposición del papel, pero puede servir como un modo práctico de proporcionar protección a ciertos documentos.

Las tecnologías digitales ofrecen un nuevo paradigma de preservación. Ofrecen la oportunidad de preservar el original proporcionando acceso a la copia digital; de separar el contenido informativo de la degradación del soporte físico.

La digitalización no produce un ahorro del coste de la gestión de la documentación. Una copia digital nunca puede sustituir al documento u objeto original. El proceso completo, selección, escaneo, creación de registros, etc. requiere una gran inversión y el mantenimiento a largo plazo de los equipos digitales tiene sus propios altos costes.

## ***1.2. Evolución en Archivo Histórico Universitario Santiago de Compostela***

### **1ª FASE:**

El proceso de digitalización en el archivo empezó en el año 1997, en esta primera fase el proyecto estaba compuesto por dos puestos para entrada de datos (cliente), un puesto de digitalización y un puesto para gestionar la base de datos (servidor).

#### **Características técnicas:**

##### **Hardware**

##### **Entrada de datos y servidor:**

- Como mínimo 64MB de memoria RAM en el servidor y 32MB en los clientes
- Monitor de 17"
- Tarjeta gráfica de 2MB como mínimo
- Disco duro de 1GB

##### **Digitalización:**

- Tarjeta SCSI y grabación de CD-ROM
- Tarjeta gráfica de 64 MB.
- Dispositivo de captura de imágenes KODAK y cuerpo cámara reflex NIKON, 3 Megapíxeles escala grises.

##### **Software**

- Sistema operativo OS/2
- Base de datos ODB2
- En esta fase las imágenes se guardaban solo en CD-rom (volumen de consulta)

**2ª FASE:**

Se produce una migración de sistema operativo de OS/2 a WINDOWS NT

**Características técnicas:****Hardware**

son las mismas.

**Software**

- Windows NT 4.0 Server Service Pack 3 en los servidores
- Windows NT 4.0 Workstation Service Pack 3 en los clientes
- ORACLE 8.0.4.0.0 con parche de ODBC 8.0.4.0.2

**3ª FASE:**

Se incorpora un servidor nuevo, Pentium 2 doble procesador, con el cual ya se puede dar servicio a los usuarios. Se incorporan ordenadores para consulta de datos e imágenes, además de un ordenador para el jefe de sala.

La función del jefe de sala es asignar mesa al investigador para que pueda acceder a la base de datos y poder visionar las imágenes de los fondos digitalizados, y si la documentación no se encuentra en la base datos generar unas papeletas de peticiones, además de prestar servicio de reprografía (fotocopias)

Dichas imágenes solo son disponibles en CD-rom, los cuales son introducidos en las respectivas unidades que serán asignadas a cada ordenador desde el puesto de jefe de sala (tantas unidades de CD-rom como puestos de consulta)

**4ª FASE:**

Las imágenes se vuelcan a discos duros, RAID 5, de esta manera las imágenes son servidas al usuario de forma automática sin necesidad de introducir CD-roms. Con el paso del tiempo y en función de las necesidades de crecimiento de la base de datos es necesario el cambio del servidor por uno más potente con lo que se aprovecha para cambiar las versiones del sistema operativo e incrementar el RAID.

Estos cambios también obligan a habilitar una sala especial con control de temperatura y acceso restringido, para proteger tanto al servidor como a los da-

tos. Las imágenes se guardan también en DVDs y discos duros externos, copias de seguridad.

### **Hardware**

#### **Servidor y RAID:**

- Compact Pentium 4
- Disco duro RAID 1Tera

#### **Digitalización:**

- Cámara KODAK con cuerpo NIKON, 14 megapíxeles de resolución, imágenes en formato RAW, color RGB y grises

### **Software**

- Windows 2000 Server Service en el servidor
- Windows XP en los clientes
- ORACLE 9.2

### **5ª FASE:**

La base de datos sale a la web inicialmente con una versión Internet Information Server (Windows), actualmente esta basado en plataforma JAVA (Apache). Se compra otro servidor para poder realizar dicha tarea.

### **6ª FASE:**

Se cambia la cámara por un escáner tipo cenital, COPYBOOK. Se añade otro RAID 5 con una capacidad de 3 Teras. Además de las imágenes digitalizadas se añaden las que se pasa de microfilm a digital.

### **Hardware**

#### **Servidor y RAID:**

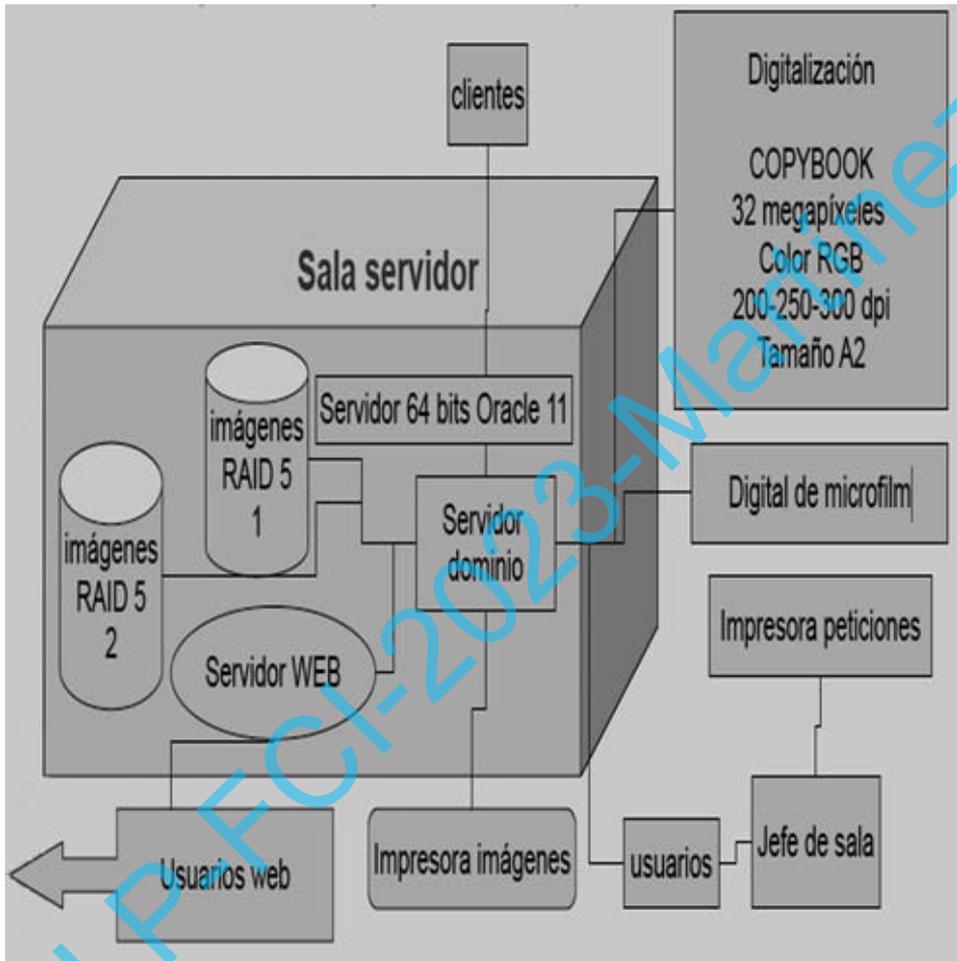
- Servidor 64 bits
- Disco duro RAID 3Teras

#### **Digitalización:**

- Copybook 32 megapíxeles de resolución, imágenes en formato JPG, TIFF, color RGB y grises

### **Software**

- Windows 2003 64 bits
- ORACLE 11



## 2. Requisitos técnicos e implementación

### 2.1. Conversión

Una imagen digital es una “fotografía electrónica” convertida en un conjunto de elementos pictóricos (píxeles) y ordenada de acuerdo con una relación predefinida de columnas y filas. El número de píxeles en una tabla dada define la resolución de la imagen. Cada píxel tiene un determinado valor tonal que depende de la intensidad de la luz reflejada desde el documento original al dispositivo de carga acoplado (CCD) mediante diodos sensibles a la luz. Cuando los documentos se exponen a la luz crean una carga eléctrica proporcional, que genera, a través de una conversión analógica/digital, una serie de señales digitales representadas por un código binario. La unidad más pequeña de información almacenada en un ordenador se denomina bit (acrónimo de binary digit). El número de bits utilizados para representar cada píxel en una imagen determina el número de colores o escala de grises que se pueden representar en una imagen digital. Esto se llama profundidad del bit.

Las imágenes digitales se conocen también como imágenes representadas por bits o imágenes *raster* para diferenciarlas de otros tipos de ficheros electrónicos como los ficheros vectoriales en los que la información gráfica se codifica mediante fórmulas matemáticas que representan líneas y curvas.

Los documentos originales se transforman en imágenes convertidas en bits por medio de un escáner o cámara digital. Durante la captura de la imagen dichos documentos se “leen” o escanean a una resolución y a una profundidad predefinidas. Los ficheros digitales resultantes que incluyen los bits para cada píxel se formatean y etiquetan de manera que se facilite a un ordenador su almacenamiento y recuperación. A partir de estos ficheros el ordenador puede realizar representaciones analógicas para su visualización en pantalla o impresión. Debido a que los ficheros con imágenes de alta resolución son muy grandes puede ser necesario reducir el tamaño de los archivos (compresión) para hacerlos más manejables tanto para el ordenador como para el usuario.

Cuando un documento fuente se ha escaneado, todos los datos se convierten a un determinado formato de fichero para su almacenamiento. Existe una variedad de formatos de imágenes ampliamente utilizados en el mercado. Algunos de ellos están pensados tanto para el almacenamiento como para la compresión. Los

ficheros de imágenes también incluyen información técnica que se almacena en el área del fichero llamada “cabecera” de la imagen.

El objetivo de cualquier programa de digitalización debería ser capturar y presentar en formatos digitales el contenido informativo más significativo de un único documento original o de una colección de tales documentos. Para capturar las partes significativas, las valoraciones de la calidad de las imágenes digitales deben basarse en la comparación entre esas imágenes digitales y los documentos originales que se van a convertir. Sin embargo, la solución no es capturar una imagen con la mayor calidad posible, sino comparar el proceso de conversión con el contenido informativo del original, ni más ni menos. Esto plantea dos cuestiones: las características de los documentos originales que se van a digitalizar y la calidad de la imagen.

## **2.2. Los documentos**

Al capturar, se deben tomar en consideración tanto los procesos técnicos implicados en la digitalización como los atributos de los documentos originales. Estos atributos pueden ser de distintas dimensiones y nivel tonal (color o blanco y negro). Los documentos originales también pueden caracterizarse por el modo en que se han producido: a mano (tinta), mecanografiados o impresos, o por métodos fotográficos o electrónicos.

La condición física de los documentos originales puede influir de diferentes maneras en la conversión. Los textos descoloridos, las manchas de tinta, las páginas quemadas y otro tipo de daños a veces destruyen el contenido informativo, pero de forma más frecuente imponen limitaciones físicas a las posibilidades de capturar información durante el escaneo.

## **2.3. Características de las imágenes**

### **2.3.a) Píxel**

Las imágenes rasterizadas o de mapas de bits son fotos electrónicas tomadas de una escena o escaneadas de documentos –fotografías, manuscritos, textos impresos e ilustraciones– de las que se confecciona un mapa en forma de cuadrícula de puntos o elementos de la figura (píxeles). A cada píxel se le asigna un valor tonal (negro, blanco, matices de gris o de color), el cual está representado en un código binario (ceros y unos). Los dígitos binarios (“bits”) para cada píxel son almacenados en una secuencia, junto con la dirección numérica que ocupa, en lo

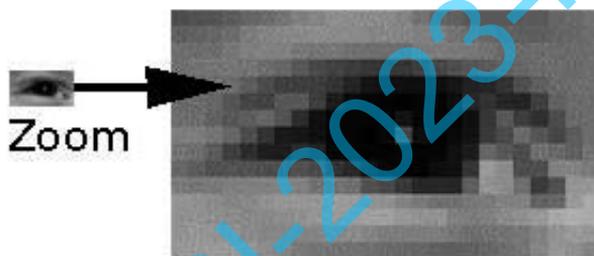
que se conoce como “mapa de imagen”. Luego se interpreta y lee los bits para producir una versión analógica para su visualización o impresión.

1	1	1	1	1	1	1	1	1	1
1	0	0	0	1	1	0	0	0	1
1	1	0	1	1	1	1	0	1	1
1	1	0	1	1	1	1	0	1	1
1	1	0	1	1	1	1	0	1	1
1	1	0	0	0	0	0	0	1	1
1	1	0	1	1	1	1	0	1	1
1	1	0	1	1	1	1	0	1	1
1	1	0	1	1	1	1	0	1	1
1	0	0	0	1	1	0	0	0	1
1	1	1	1	1	1	1	1	1	1

A cada píxel se le asigna un valor tonal, en este ejemplo 0 para el negro y 1 para el blanco

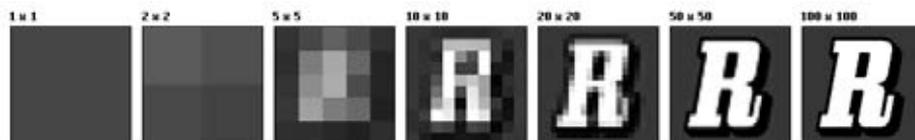
### 2.3.b) Dimensiones de píxeles

Son las medidas horizontales y verticales de una imagen, expresadas en píxeles. Las dimensiones de píxel se pueden determinar multiplicando tanto el ancho como la altura por el dpi.



### 2.3.c) Resolución

La resolución de la imagen es la cantidad de píxeles. La resolución se utiliza también para clasificar casi todos los dispositivos relacionados con las imagen digital ya sean pantallas de ordenador o televisión, impresoras, escáneres, cámaras, etc. La resolución expresa el número de píxeles que forman una imagen de mapa de bits. La calidad de una imagen también depende de la resolución que tenga el dispositivo que la capta. El número de píxeles que contenga una imagen depende de cuántos píxeles utilice el sensor CCD de la cámara para captar la imagen. La resolución de una imagen digital se expresa multiplicando su anchura por la altura en pantalla. Por ejemplo la imagen de 1200 x 1200 píxeles = 1.440.000 píxeles, expresado en Mp megapixel es igual a 1,4 Mp.



### 2.3.d) Profundidad de bits

Es determinada por la cantidad de bits utilizados para definir cada píxel. Cuanto mayor sea la profundidad de bits, tanto mayor será la cantidad de tonos (escala de grises o color) que puedan ser representados. Las imágenes digitales se pueden producir en blanco y negro (en forma bitonal), a escala de grises o a color. Una imagen bitonal está representada por píxeles que constan de 1 bit cada uno, que pueden representar dos tonos, utilizando los valores 0 para el negro y 1 para el blanco. Una imagen a escala de grises está compuesta por píxeles representados por múltiples bits de información, que típicamente varían entre 2 a 8 bits o más. Ejemplo: En una imagen de 2 bits, existen cuatro combinaciones posibles: 00, 01, 10 y 11. Si “00” representa el negro, y “11” representa el blanco, entonces “01” es igual a gris oscuro y “10” es igual a gris claro. La profundidad de bits es dos, pero la cantidad de tonos que pueden representarse es 22 ó 4. A 8 bits, pueden asignarse 256 (28) tonos diferentes a cada píxel. A cada píxel se le asigna un valor tonal, en este ejemplo 0 para el negro y 1 para el blanco.



Una imagen a color está representada por una profundidad de bits entre 8 y 24 o superior a ésta. En una imagen de 24 bits, los bits por lo general están divididos en tres grupos: 8 para el rojo, 8 para el verde, y 8 para el azul. Para representar otros colores se utilizan combinaciones de esos bits. Una imagen de 24 bits ofrece 16,7 millones de valores de color. Pero cada vez más, los escáneres están capturando 10 bits o más por canal de color, llegando a digitalizar imágenes en 32 bits.

Cálculos binarios para la cantidad de tonos representados por profundidades de bits comunes:

- 1 bit = 2 tonos
- 2 bits = 4 tonos
- 3 bits = 8 tonos
- 4 bits = 16 tonos
- 8 bits = 256 tonos
- 16 bits = 65.536 tonos
- 24 bits = 16,7 millones de tonos
- 32 bits = 4.294 mill. De tonos

### 2.3.e) *Tamaño del archivo*

Depende del número de píxeles de la imagen y de su profundidad de bits. Se calcula multiplicando las dimensiones de la imagen en píxel verticales y horizontales por la profundidad de bit para determinar la cantidad de bits presentes en un archivo de imagen. Por ejemplo, si se captura una imagen de 24 bits con una cámara digital con dimensiones de píxel de 2.048 x 3.072, entonces el tamaño de archivo es  $(2048 \times 3072 \times 24) / 8 = 50.331.648$  bytes.

	<b>DPI</b>			
	<b>100</b>	<b>200</b>	<b>400</b>	<b>600</b>
<b><i>TIFF</i></b>	1.93 Mbytes	7.81 Mbytes	31.3 Mbytes	70.3 Mbytes
<b><i>JPEG</i></b>	114 Kbytes	230 Kbytes	660 Kbytes	1.2 Mbytes

### 2.3.f) *Compresión*

La compresión se utiliza normalmente para reducir el tamaño del fichero con el fin de procesar, almacenar y transmitir las imágenes digitales. Los métodos utilizados son, por ejemplo, la eliminación de información repetida o la de difícil visión para el ojo humano. De este modo la calidad de una imagen puede verse afectada por las técnicas de compresión utilizadas y por el nivel de compresión aplicado. Las técnicas de compresión pueden ser tanto “sin pérdida”, lo que significa que una imagen descomprimida será idéntica a su estado original debido a que no se ha perdido ninguna información cuando se ha reducido el tamaño del fichero, como “de pérdida” cuando la información menos significativa se reduce o desecha en este proceso.

En general, la compresión “sin pérdida” se utiliza para los ficheros maestros y las técnicas de compresión “con pérdida” para los archivos de acceso. Es importante tener en cuenta que las imágenes pueden responder a la compresión de diferentes modos. Algunos tipos concretos de características visuales como leves variaciones de tono pueden producir efectos visuales no deseados.

### 3. Captura de imágenes copybook

#### 3.1. Características técnicas.

El CopyBook es un escáner cenital que nos permite digitalizar formatos DIN A2. Tiene la capacidad de trabajar únicamente con luz ambiente sin ningún tipo de luz integrada. Incorpora varios ajustes para facilitar el trabajo (detección automática de formato, resolución, autofocus, tiempo de exposición automático, corrección de luz, balanza de blancos y negros) además de procesos para la mejora de las imágenes.

Permite escala de grises y color RGB, cabezal de cámara de 35 Mpíxeles, resolución óptica máxima de 300 dpi, formato de imágenes soportados TIFF, JPEG, BMP. Previsualización de la documentación para una mejor colocación.



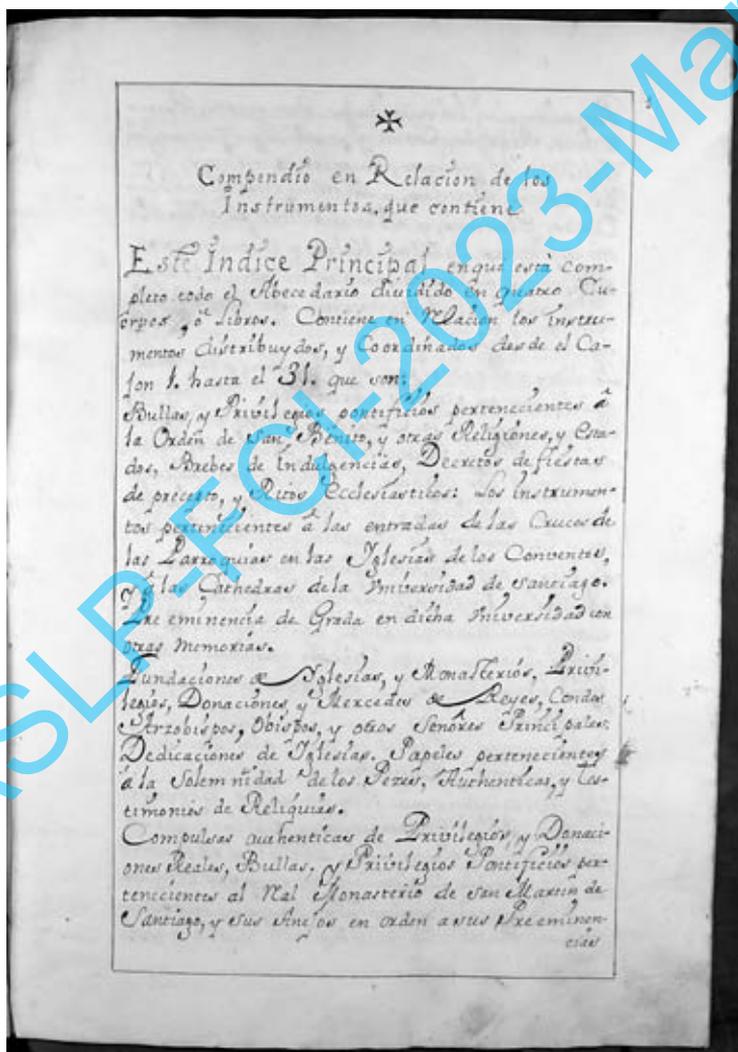
#### 3.2. Proceso de captura.

Las imágenes se capturan en escala de grises y en formato JPEG con un ajuste de compresión de 90 con una resolución de 200 dpi. También se pueden hacer en color, dichas imágenes se hacen por petición de los usuarios sobre todo para publicaciones.

Dichas imágenes se pueden hacer de dos en dos, utilizando el cristal, siempre que la altura de la documentación este dentro de los límites. Estos casos casi no se dan ya que la documentación que se maneja es mucho más grande. Por lo general se digitalizan primero las rectas y luego las verso.

Las imágenes se guardan en una carpeta en el ordenador, sin importar la numeración, para posteriormente ser insertadas en la base de datos.

*Clero 737, 1 r.*





## 4. Base de datos.

### 4.1. Introducción.

Archidoc-archigés es una solución informática que proporciona las herramientas *software* necesarias para llevar a cabo la gestión y el control de los fondos de archivos y bibliotecas. Facilita el acceso rápido a todos los documentos y su clasificación por distintos criterios, permitiendo de esta manera su pronta localización. Sirve para hacer inventario de archivos y bibliotecas y para añadir en fichas informáticas una síntesis del contenido de cada documento archivado. Además garantiza el control y la disponibilidad de recursos con los que cuenta cualquier centro de documentación.

Las ventajas que proporciona podríamos resumirlas en las siguientes:

- Mejora la gestión de los fondos disponibles y no disponibles.
- Facilita el acceso a las fuentes de información a través de medios informáticos, sin necesidad de tener físicamente el documento (Internet).
- Evita el deterioro por el uso de los fondos archivados.
- Evita movimientos de documentos archivados y reubicaciones físicas.
- Facilita la agrupación de usuarios en distintos perfiles, en función de las tareas que cada uno de ellos tenga asignadas.

Incorpora módulos de gestión de bases de datos; digitalización de imágenes; control y gestión de usuarios internos y externos; gestión de fondos (conservación y restauración incluidas) y depósitos; etc.

Opera sobre plataforma Windows (NT y 2000); incluye también las normas de descripción archivística ISAD(G) e ISAAR(CPF), y posee gran flexibilidad en su configuración.

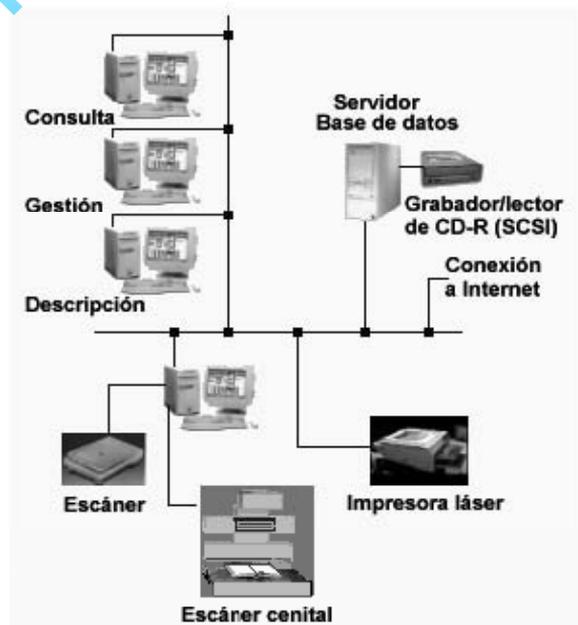
Archidoc incorpora funciones relacionadas con el tratamiento y consulta de la documentación que permiten al usuario generar gráficamente el cuadro de clasificación de los fondos; estructurar los elementos de descripción según las normas ISAD(G), ISAAR(CPF) y otras fichas particulares; generar masivamente objetos basados en patrones y rangos de firmas; normalizar descriptores; generar tesauros; digitalizar documentos y crear bases de datos de imágenes que den la posibilidad de consultar fondos sin que el usuario los manipule y con ello salvaguardarlos de posibles deterioros; almacenar las imágenes en formatos estándares (TIFF, JPEG, CCITT...); retocar en la pantalla las imágenes para

que puedan verse mejor; realizar todo tipo de búsquedas de fondos: utilizando descriptores, signaturas, texto libre de campos designados a tal efecto, etc; gestionar las solicitudes que se puedan producir de servicios tales como impresión, fotocopias, originales, etc; preparar publicaciones, catálogos y guías, ordenadas según distintos índices; etc.

Archigés facilita la gestión de usuarios y fondos en sus distintos aspectos y fases de su ciclo vital. Esto permite al usuario realizar tareas de control y seguimiento de la utilización de los fondos documentales que hacen los usuarios; distinguir entre usuarios internos y externos de la institución; controlar las solicitudes de acceso a los fondos; expedir tarjetas de investigador; generar informes internos, estadísticas, memorias, etc; gestionar los puestos de trabajo y la distribución de mesas en las distintas salas de la institución: salas administrativas, de investigación, de consulta de fondos por ordenador, etc; controlar los movimientos de los fondos en el depósito; controlar el estado de conservación de fondos y llevar el seguimiento de los procesos de conservación y restauración; planificar y controlar las transferencias, ingresos y préstamos de documentos; registrar las altas y bajas de fondos; definir derechos de acceso a distintos grupos de usuarios que garanticen la confidencialidad de fondos y la seguridad de la institución; etc.

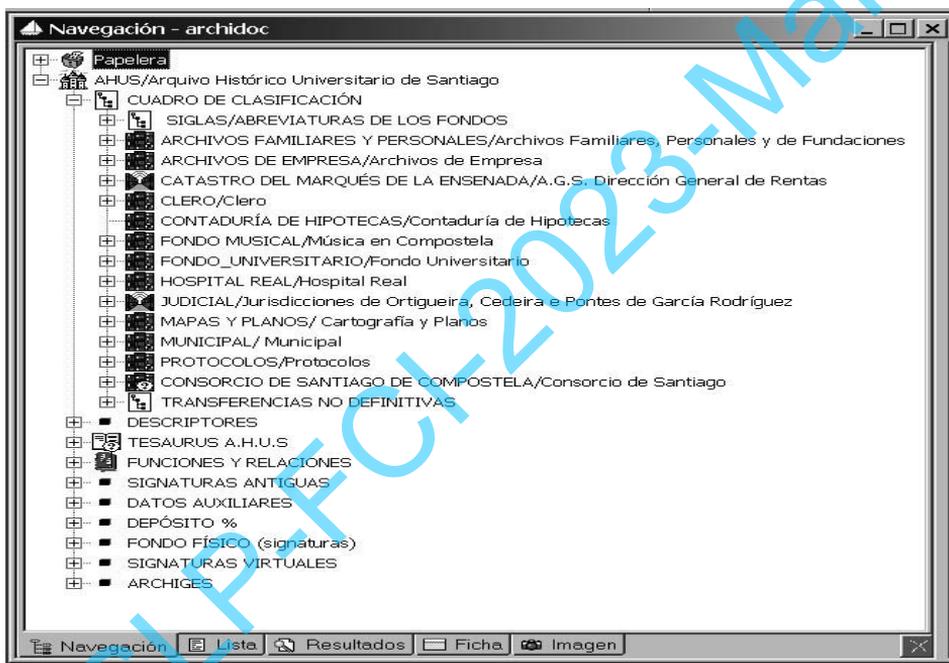
Para gestionar las personas que tienen acceso a los fondos el sistema distingue entre personal interno, empleados, y personal externo, investigadores. Cada uno de ellos puede a su vez agruparse en distintos perfiles de usuarios que utilizarán la aplicación de acuerdo a las funciones que tengan asignadas.

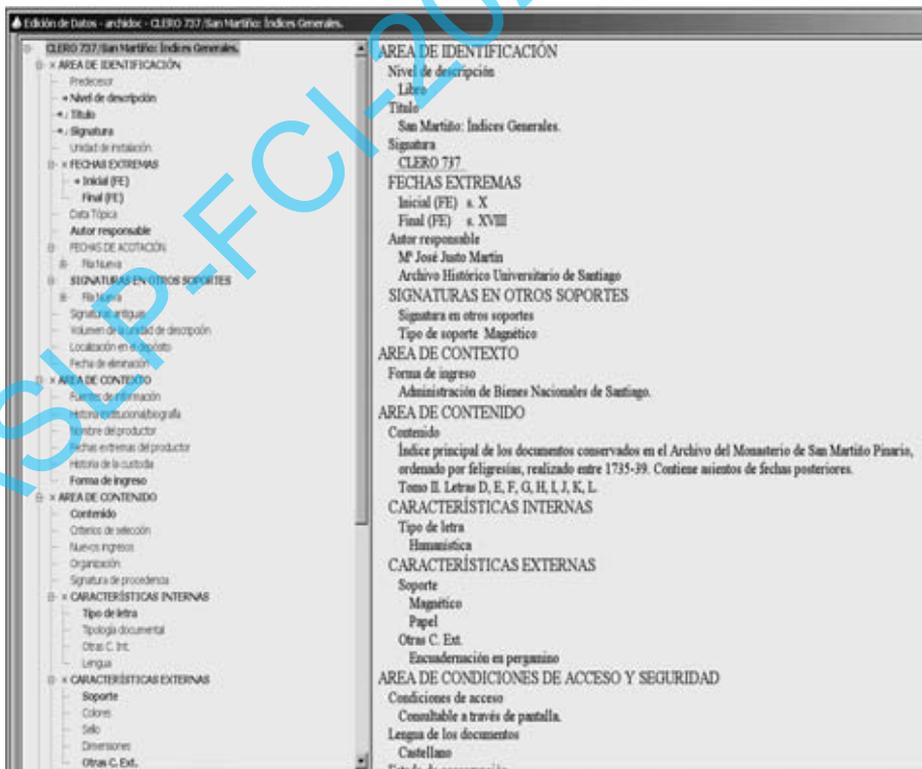
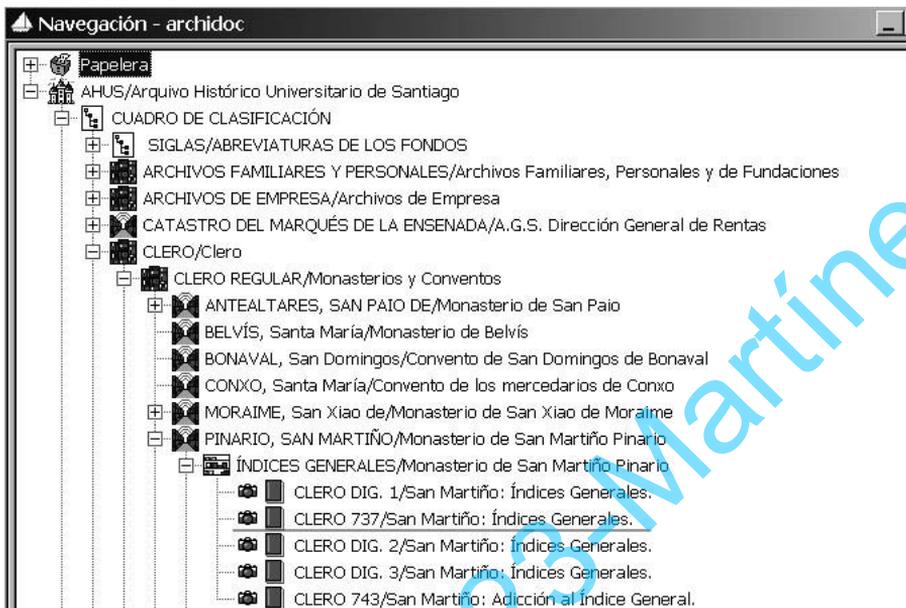
Archidoc-archigés se implementa en una arquitectura cliente-servidor que permite crecer desde un pequeño sistema inicial en monopuesto



hasta grandes sistemas con varias estaciones para los distintos usuarios en el mismo centro o a través de Internet.

La base de datos de ARCHIDOC está organizada en forma de árbol. Cada elemento de información de la base de datos es un objeto; a las ramas o elementos principales los llamamos nodos (nodos padres) y a su vez éstos pueden tener descendientes (nodos hijos), es decir, nodos que están un nivel por debajo de él, o hermanos (nodos hermanos), es decir, los que están al mismo nivel. Esta estructura permite enlazar jerárquicamente toda la información del archivo de forma dinámica y crecer en la dirección que sea necesaria.





## 4.2. Digitalización.

Este proceso sirve para convertir en imágenes los fondos del archivo o biblioteca. Normalmente el usuario encargado de realizar esta tarea es el digitalizador. El procedimiento consta de varias etapas:

1. Previamente el administrador debe haber configurado en archidoc los volúmenes de digitalización, dónde se van a guardar las imágenes. Esta labor se realiza en el momento de instalación y configuración del sistema y salvo que se añadan nuevos dispositivos para almacenar imágenes o nuevos usuarios de digitalización no deberá volver a realizarse.
2. A continuación, seleccionar el nodo en el que se van a guardar las imágenes digitalizadas y crear la estructura en la base de datos.
3. En tercer lugar, tiene que establecer las opciones de digitalización.
4. En cuarto lugar, configurar el escáner que va a utilizar para digitalizar las imágenes y después seleccionar el modelo del mismo.
5. En quinto y último lugar, indicar el formato de las imágenes en el que se guardarán al ser digitalizadas.

### 4.2.1. Crear la estructura en la base de datos.

Es la siguiente etapa antes de proceder a digitalizar las imágenes. De lo que se trata es de crear el nodo en el lugar adecuado de la base de datos para que se guarden las imágenes del fondo físico del archivo.

Cómo crear una estructura en la base de datos para las imágenes:

1. Sitúe el cursor en el NODO PADRE del fondo físico en el que va a guardar las imágenes. Fondo Físico SIGNATURAS, CLERO, CLERO REGULAR, SAN MARTIN PINARIO, INDICES GENERALES, CLERO 737



2. Crear un nodo nuevo descendiente y rellenar los campos del objeto. Puede dejar en blanco, si lo desea, el campo Etiqueta ya que al digitalizar la imagen le pondremos un nombre. Asignamos el icono Imagen a este objeto.

Con esto ya tenemos creada la estructura para proceder a digitalizar los fondos que necesitamos. Es el momento de establecer las opciones de digitalización y la configuración del escáner.

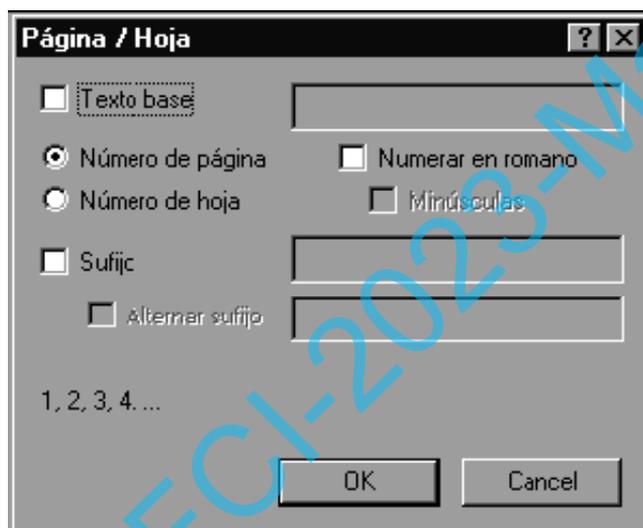
#### 4.2.2. Establecer las opciones de digitalización

Con el objeto que acabamos de crear seleccionamos Opciones de digitalización.

Contiene lo siguiente:

-“Número de imagen”. Es el contador de la imagen digitalizada y salvada. Archidoc-archigés utiliza este número para dar nombre a los objetos con imagen que se van creando. La forma de asignar este nombre se configura con el

botón “Secuencia”. El número que aquí se indica es el primero de la secuencia de digitalización. En una secuencia normal de digitalización este contador va incrementándose de uno en uno, pero puede alterarse si marca las casillas de verificación “Doble cara” y “Hacia atrás”, explicadas más tarde. “Secuencia”. Si lo pulsa se abre el cuadro de diálogo que muestra la figura y que permite configurar el nombre que se asigna automáticamente a cada uno de los objetos imagen que se van generando durante la digitalización. En principio, si no se configura nada, archidoc-archigés asigna a cada objeto el número del campo Número de imagen.



“Texto base” se habilita el cuadro de texto de su derecha en el que puede escribir el texto que se va a anteponer al número de la imagen que se digitaliza. Por ejemplo, si escribe PÁGINA se numerarán: PÁGINA1, PÁGINA2, PÁGINA3.

“Número de página” o “Número de hoja”. Ambas se excluyen mutuamente y actúan sobre el número que se incluye en el nombre del objeto. Con la opción Número de página se entiende que el documento digitalizado está paginado, es decir, el número se va incrementando de uno en uno con cada imagen. La opción Número de hoja asume que el documento está foliado y por lo tanto el número se incrementa de uno en uno cada dos imágenes.

“Numerar en romano”. Marque esta casilla si quiere que se utilice este tipo de numeración en vez de la arábiga. Tiene asociada la casilla “Minúsculas”, si la marca los números romanos se escriben en minúscula.

“Sufijo”. Si marca esta casilla puede introducir el texto que se escribirá después de la numeración de las imágenes. Puede elegir además un sufijo alternativo y se numerará de la siguiente manera: sufijo principal, sufijo alternativo, principal, alternativo.

En la parte inferior del menú se muestra un ejemplo de cómo quedarán los nombres de los objetos digitalizados. Pulse “Ok” para guardar esta configuración y volver al cuadro de diálogo anterior donde continuar con las opciones de digitalización.

“Aumentar contraste” tiene que introducir un valor entero positivo que puede variar de 0 a 49 y que regula la intensidad con la que actúa el aumento del contraste. Fuerza un incremento del contraste en la imagen tras ser explorada por el escáner. En líneas generales esto mejora la calidad y legibilidad de la imagen final. Suele utilizarse para ver las imágenes que se han tomado con cámaras digitales.

“Rotación” Puede seleccionar también si desea que las imágenes se giren o no al digitalizarlas. Tiene las siguientes posibilidades: 0°, 90°, 180° y 270°.

“Doble cara” y “Hacia atrás”. En ocasiones, ya sea por las características del documento que se va a digitalizar o por el tipo de escáner, es más cómodo explorar primero páginas pares y después páginas impares. Si marca “Doble cara”, archidoc-archigés genera los objetos impares incrementando el contador de imagen de dos en dos y dejando hueco a las páginas pares. Si marca las dos casillas, archidoc-archigés genera los objetos pares disminuyendo el contador de imágenes (también de dos en dos) y situando los objetos pares en sus lugares correspondientes. Recuerde que para fijar estas opciones debe tener en cuenta el modelo de escáner, y que lo que es válido para un modelo no puede serlo para otro.

“Configurar”. Nos abrirá una ventana de Windows para seleccionar las imágenes que previamente hemos digitalizado con el escáner.

“Seleccionar”. Si pulsa este botón se abre el cuadro de diálogo Selección de fuente para elegir el driver *Twain* correspondiente al dispositivo que va a utilizar; en nuestro caso es la opción ArchidocFileImport.

“Formato”. Este botón sirve para fijar el formato en el que se van a guardar las imágenes digitalizadas (también puede hacerlo directamente con la opción Formato de imagen del menú Imagen, o del correspondiente menú contextual, habilitada únicamente si tiene abierta y activa la ventana Imagen). Si lo pulsa se abre el cuadro de diálogo Selección de formatos que permite elegir el formato de compresión de las imágenes que se van a salvar.



Este cuadro de diálogo se divide en cuatro áreas correspondientes a las cuatro posibles formas de captura: Grises, Binaria, 256 Colores y *True Color* (16 millones de colores). Cada opción es almacenada en memoria y no es necesario volver a invocar este diálogo; si se vuelve a este cuadro de diálogo se divide en cuatro áreas correspondientes a las cuatro posibles formas de captura: Grises, Binaria, 256 Colores y *True Color* (16 millones de colores). Cada opción es almacenada en memoria y no es necesario volver a invocar este diálogo si se vuelve a configurar el escáner o se cambia de dispositivo. Cada campo contiene las opciones posibles. Con los valores de la figura anterior, si el escáner estuviera configurado para explorar en binario, la imagen se salvaría en formato TIFF Grupo Cuatro. En caso de que estuviera explorando en 256 colores, la imagen sería salvada con formato BMP de OS/2. Una imagen explorada en *True Color* se salvaría JFIF-JPEG al 75% de calidad, y si es el caso de 256 niveles de gris se salvaría como JFIF-JPEG al 85% de calidad. En el caso de la compresión de imágenes binarias en TIFF G4, si el original presenta un fondo tramado, es posible que la imagen comprimida ocupe más que la original. Si esto sucede y está seleccionada la casilla “No comprimir si expande” la imagen será almacenada como TIFF sin comprimir. Si no se marca esta casilla, archidoc-archigés dará un mensaje de aviso y no salvará la imagen a la espera de que se elija un nuevo formato de compresión.

Pulse “Aceptar” para guardar esta configuración, volver al cuadro de diálogo Opciones de captura y continuar con el proceso de digitalización.

“Captura”. Presenta la primera imagen.

“Ocultar configuración”; esta opción evita que se abran las propiedades del escáner cada vez que capturemos una imagen; en nuestro caso lo que evita es abrir el explorador de Windows.

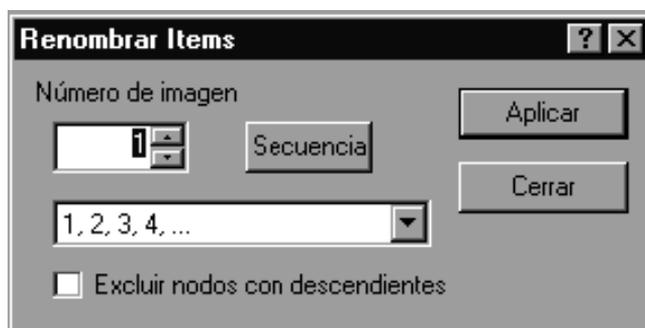
“Ciclo” Si pulsa este botón el sistema salva la imagen previamente leída, crea un nuevo objeto y ordena una nueva exploración al escáner. Por lo tanto es como si pulsara primero el botón “Salvar” y luego “Captura”.

“Lectura automática”. Si pulsa este botón el sistema salva la imagen previamente leída, crea un nuevo objeto, explora, asocia y salva imágenes con lo que evita la necesidad de pulsar los botones “Captura” y “Salvar” (o el botón “Ciclo”). No es necesario por tanto la intervención del operador de digitalización, salvo, claro está, para cambiar las páginas que se van a digitalizar. Tras pulsar “Lectura automática” se habilita el botón “Parar Lectura” para detener este proceso.

“Cerrar”. Si lo pulsa cierra el cuadro de diálogo Opciones de captura y vuelve a la ventana de la aplicación.

#### 4.2.3. Renombrar objetos

Acabamos de decir que las imágenes son objetos de la base de datos. En ocasiones se necesita renombrar un bloque de imágenes u objetos de otras características. El sistema ha previsto esta situación y dispone de la opción Renombrar objetos. Al elegir dicha opción se abre el cuadro de diálogo Renombrar items que muestra la figura:



Su explicación es la siguiente:

“Número de imagen”. Indica el valor numérico por el que se va a iniciar la reenumeración. La lista de secuencias permite elegir una de las que previamente haya configurado desde el botón “Secuencia”

“Excluir nodos con descendientes”. Si se activa esta casilla de verificación indica que aquellos nodos seleccionados que tengan descendientes no se reenumerarán.

“Secuencia” Al pulsarlo se abre el cuadro de diálogo Página / Hoja en el que puede indicar la forma en la que quiere renombrarlos.

#### ***4.2.4. Copias de seguridad***

Una vez que las imágenes están incorporadas, se verifican y si no hay problema se procederá a crear las copias de seguridad. Dichas copias se hacen grabando las imágenes en volúmenes de consulta, CD-rom o DVD.

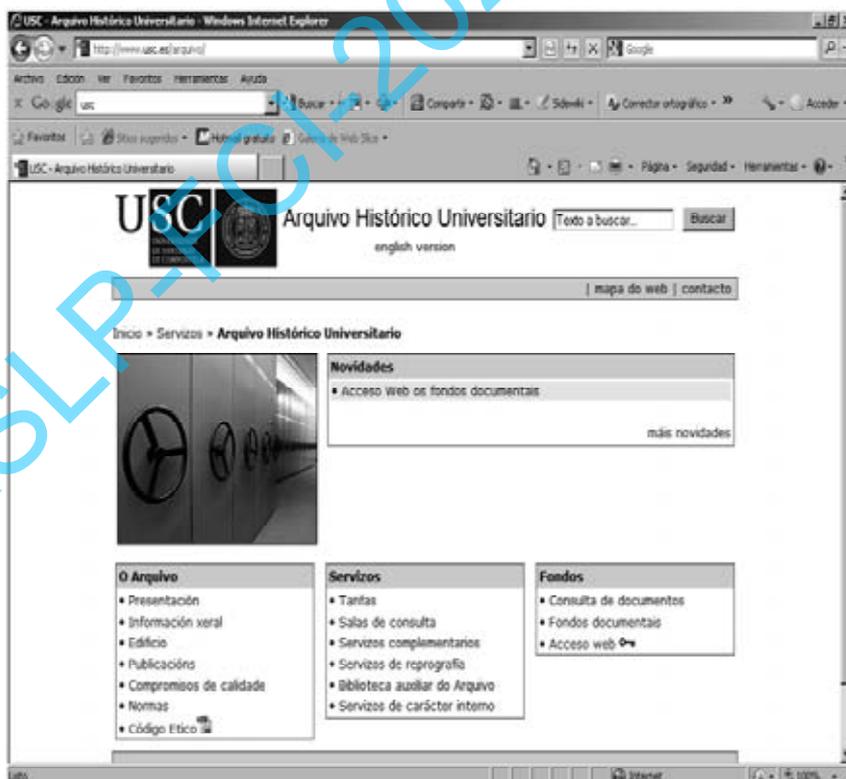
Estos volúmenes de consulta se guardan en el RAID para que los usuarios pueden ver las imágenes automáticamente.

## 5. La base de datos en la web

Los fondos documentales están a disposición de todos los usuarios vía web gracias al modulo online de archidoc, dicho modulo esta basado en plataforma JAVA y el acceso es gratuito.

Los pasos a seguir para conectarnos a la base de datos son relativamente sencillos: entrando en la página principal de la universidad de Santiago de Compostela ([www.usc.es](http://www.usc.es)), en el apartado de servicios y entrando en el Archivo Histórico Universitario ([www.usc.es/archivo](http://www.usc.es/archivo)). Allí tendremos toda la información para poder acceder a la base de datos, además de un tutorial en forma de video donde nos explica como movernos por las distintas ventanas y como efectuar búsquedas. Hay que puntualizar que para poder ver las imágenes es necesario descargarse un pequeño programa, para descargarlo existe un enlace dentro de la propia aplicación.

La base de datos online utiliza la misma estructura jerárquica que el cuadro de clasificación utilizado en el propio archivo.





Archivos Web - Windows Internet Explorer  
 https://archivos.uc.es/imagenes/actor/actorForma.html?id=120717133ba334=1

Archivos web archDOC

Simple Galería Avanzada Salir

CLERO 737

**Elemento del cuadro**

TÍTULO: San Martín: Índices Generales.  
 SIGNATURA: CLERO 737  
 FECHA INICIAL: s. X  
 FECHA FINAL: s. XVIIII  
 NIVEL: Libro  
 ARCHIVO: CLERO/Clero

**Contenido de la Carpeta**

Resultados de 1 a 10 de 510

ÍCONO	TIPO	TAMAJÑO
1 r.	jpg	378,92 KB
1 v.	jpg	397,85 KB
2 r.	jpg	497,8 KB

Archivos Web - Windows Internet Explorer  
 https://archivos.uc.es/imagenes/actor/actorForma.html?id=120717133ba334=1

■ Imágenes Cerrar

Elemento del cuadro

1 r. [1/10] Siguiente >>

1

\*  
 Compendio en Relacion de los  
 Instrumentos, que contiene  
 Este Índice Principal, en que está com-  
 plecto todo el Albeceario dividido en quatro Cu-  
 erpos, ó Libros. Contiene en Relacion los Instru-  
 mentos Cúbrubuydos, y Coordinados desde el Ca-  
 non 1. hasta el 31. que son  
 Bullas, y Privilegios pontificios pertenecientes á  
 la Orden de San, Benito, y otras Religiones, y Cta-

## 6. Proyectos de colaboración (URBE)

En el proyecto URBE participan: Universidad de Santiago, Consorcio de Santiago y Archivo de la Catedral.

El proyecto URBE se creó con el fin de ayudar al estudio sobre la historia de la ciudad de Santiago. Para ello se ha recurrido a los protocolos notariales, los cuales son una de las mayores fuentes de estudio, testimonio e información de la historia de nuestra ciudad.

Este proyecto de investigación tiene como objeto el estudio de los protocolos notariales, que en nuestro caso, se conservan en el Archivo Histórico Universitario de Santiago.

Un protocolo notarial es un conjunto de documentos generados por un notario durante un periodo de tiempo determinado (mensual, anual, etc.). Existen diversas clases de notarios en la ciudad de Santiago: notarios de Número y Concejo y de Número y Cabildo, escribanos de número de la Audiencia del Asistente o Justicia Mayor, escribanos de la Audiencia Eclesiástica y seglar del Juez de Quintana, además de otros escribanos reales. El proyecto se centra en los protocolos de los notarios de Número y Concejo y de Número y Cabildo de la ciudad de Santiago. El porqué de esta elección se debe a que estos notarios mencionados aportan una mayor representación de la vida compostelana.

La tarea comienza con un análisis del protocolo, y dependiendo de su estado de conservación se procederá a su digitalización y descripción. El siguiente paso será el conocer si el protocolo presenta foliación, numeración o ambas y además si hay un índice en el que se refleja el contenido de las escrituras que conforman dicho protocolo. A continuación se procederá a la digitalización del protocolo, respetando el orden natural del dicho conjunto documental y siguiendo unas pautas que nos servirán para la descripción posterior. Para la descripción se utilizan las normas ISAD (G): General International Standard Archival Description. Para esta tarea empleamos el módulo de descripción del sistema integrado de gestión archivística ArchiDOC-ArchiGES. Dicho sistema nos permite el visiado conjunto de la descripción y las imágenes, así como las búsquedas simples y avanzadas y la reproducción de originales, etc.

La descripción de las escrituras sigue el siguiente esquema:

- Tipo documental (venta, arriendos, testamentos, donaciones, etc.).
- Otorgantes (personas físicas o jurídicas), incluyendo los datos de vecindad y oficio si los tuviesen.
- Destinatarios (personas físicas o jurídicas), incluyendo los datos de vecindad y oficio si los tuviesen.
- Número de folio.
- Fecha de la producción de la escritura.

Con este trabajo y su producto final, la difusión de la información contenida en estos protocolos hará que la administración, los ciudadanos e investigadores dispongan de una gran ayuda y de un valioso instrumento para el conocimiento y estudio de la evolución social y económica de la ciudad.

## 7. Referencias bibliográficas

- Directrices para proyectos de digitalización *UNESCO* (IFLA e ICA).
- Manual Archidoc-archigés, *Servicios Informáticos El Corte Inglés*
- Imágenes del *ARQUIVO HISTÓRICO UNIVERSITARIO SANTIAGO DE COMPOSTELA*